NANYANG TECHNOLOGICAL UNIVERSITY

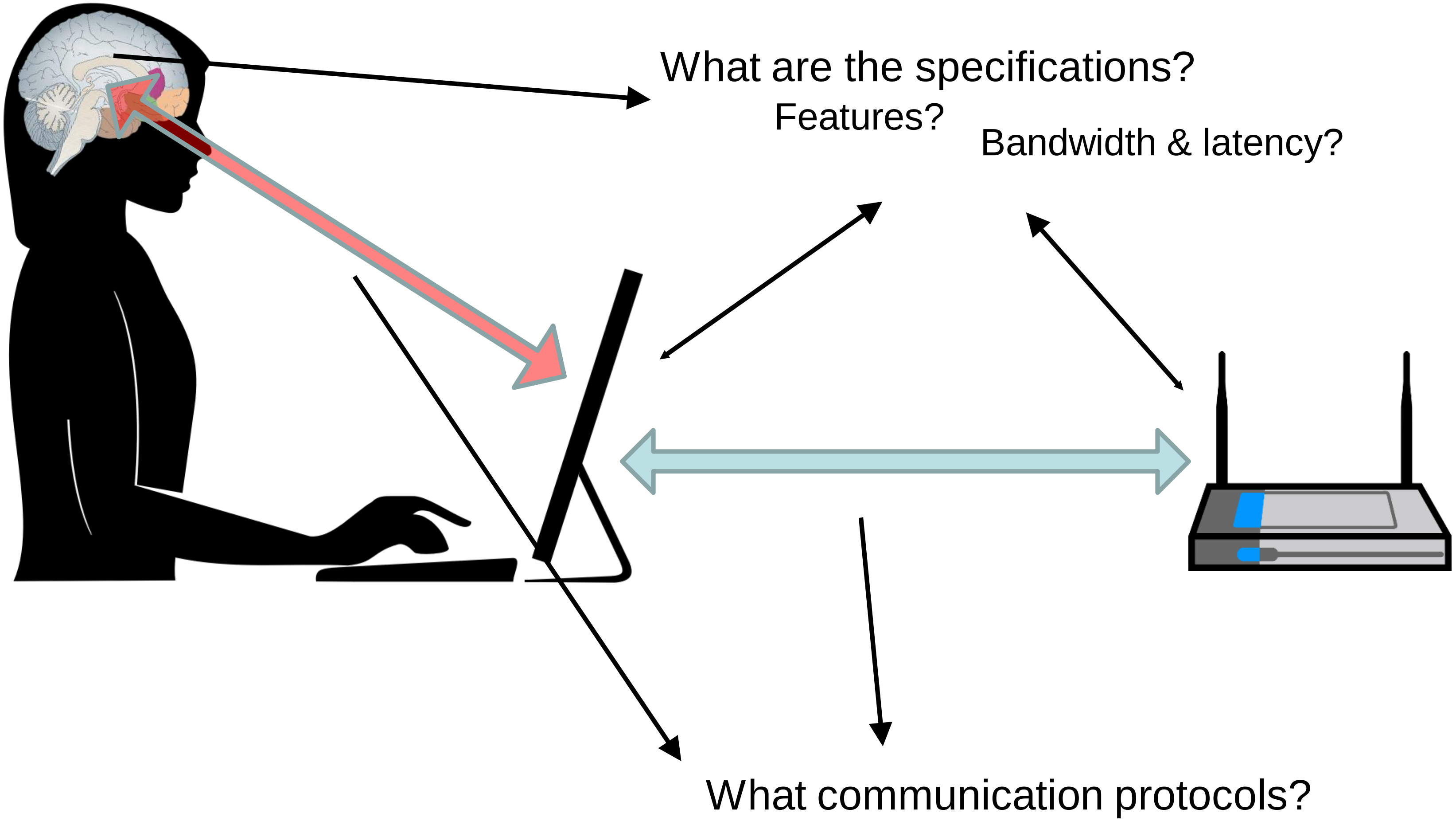
Humans

CZ2004 Human-Computer Interaction

Contents

- Sensing and Perception
  - Senses
  - Perception
  - Attention
- Higher-Level Cognition and Mental Processes
  - Memory
  - Reasoning & problem solving
  - Skill acquisition
  - Mental models
  - Emotions and Personality Traits
- Action and Behavior
  - Motor coordination
  - Fitts' Law
  - Intentional (verbal) communication
  - Non-intentional communication (body language)

Human-Computer Interfacing vs Computer-Computer Interfacing

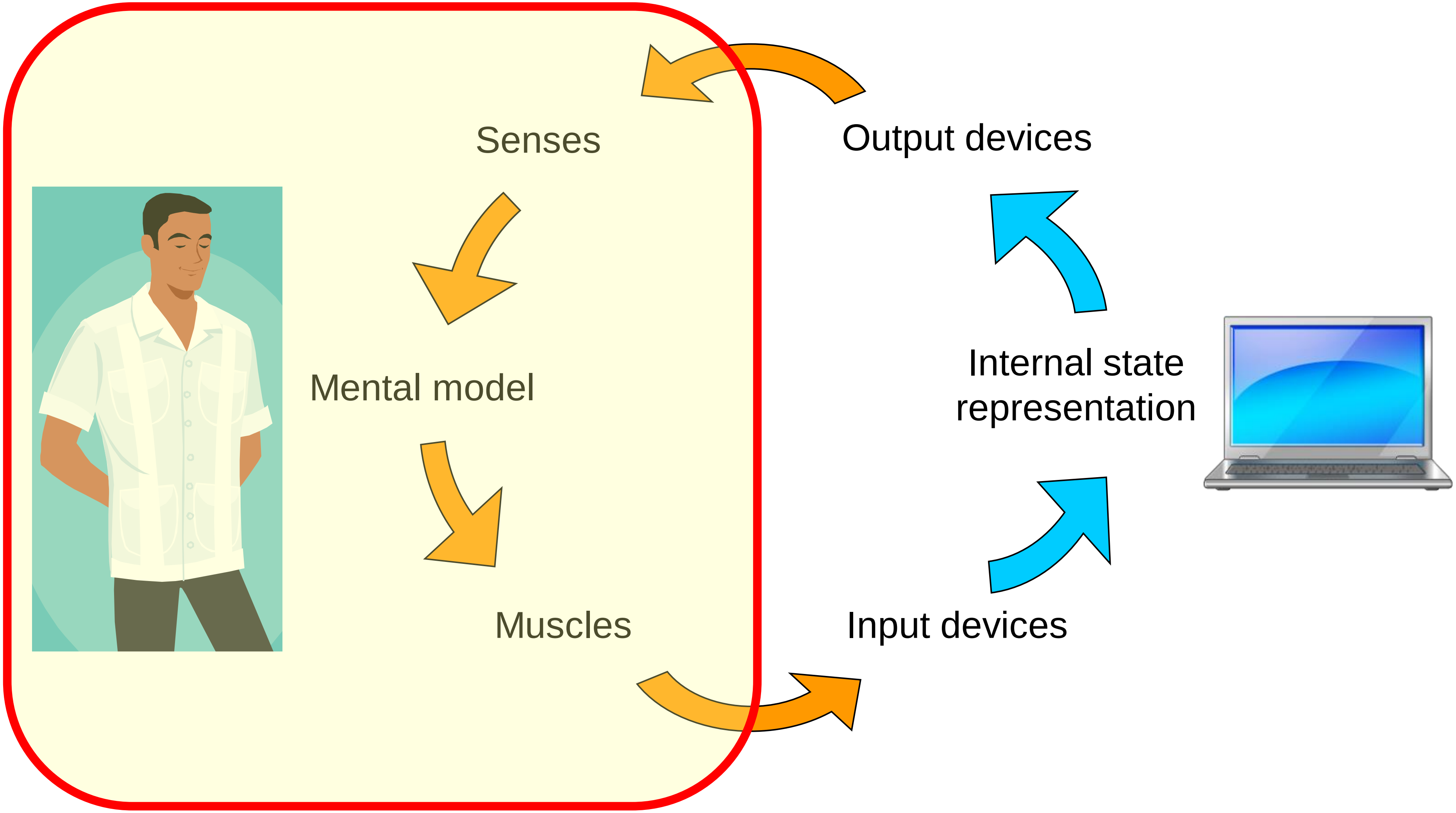


Learning Objectives

- Acquire basic knowledge about:
  - how we humans sense and perceive the world around us
  - how we remember things
  - how we solve problems and learn new skills
  - how we react emotionally
  - how we move
  - how we communicate, deliberately or subconsciously

- Relevance
  - If we have a greater scientific understanding of ourselves, we can:
    - make human-computer interaction better and smoother in a methodical way
    - think up new ways of interacting with computers, beyond what you know today and see in Hollywood

Interfacing Humans and Computers



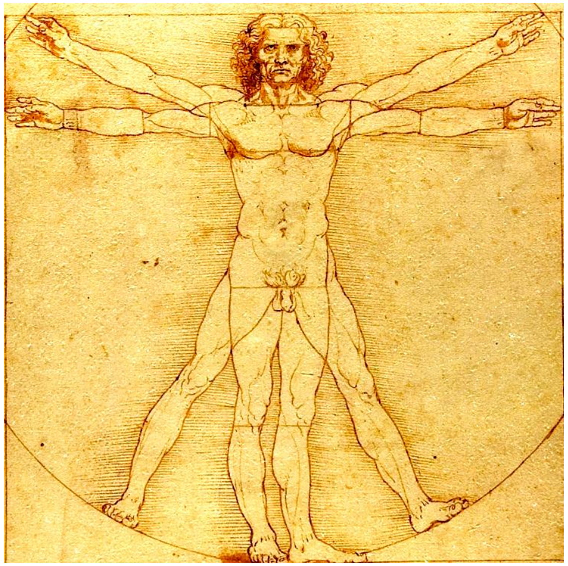
Human as Processing I/O Device

Higher-Level Cognition & Affect

- Memory
- Reasoning / problem solving
- Learning / skill acquisition
- Mental models
- Emotion and Personality

Sensing & Perception

- Senses
- Cognitive perception
- Attention



Action & Behavior

- Motor coordination
- Intentional / verbal communication
- Body language

To build good interfaces *for* humans, we need to know our human capabilities ("human specifications")

Sensing and Perception

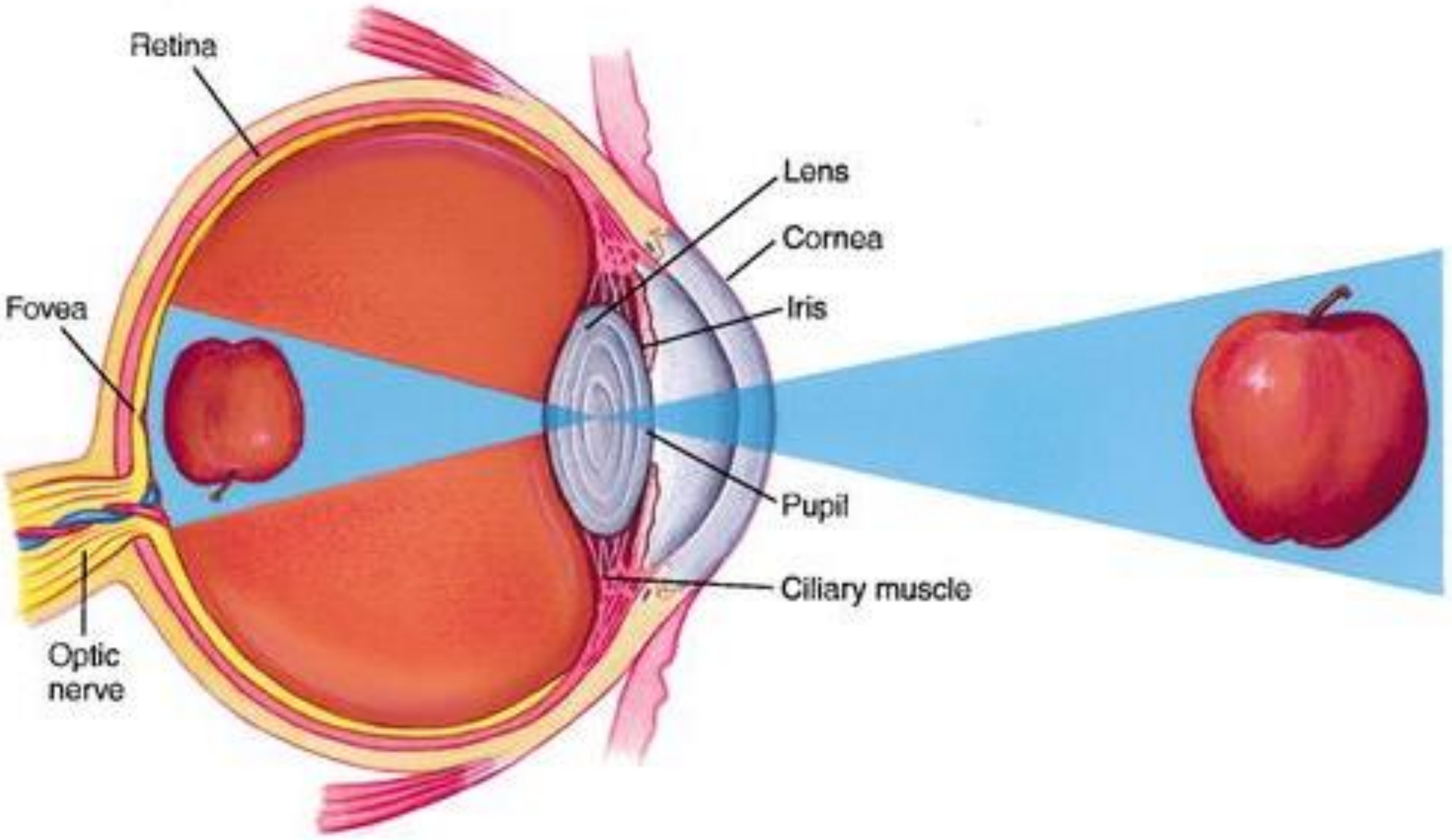
How Many Senses?

- Sight (vision)
- Hearing (audition)
- Smell (olfaction)
- Taste (gustation)
- Touch (tactition)

- Balance (equilibrioception)
- Personal body pose (proprioception)
- Temperature
- Pain

- Others: thirst, hunger, need for breath, need for bathroom break, etc.

Human Visual System

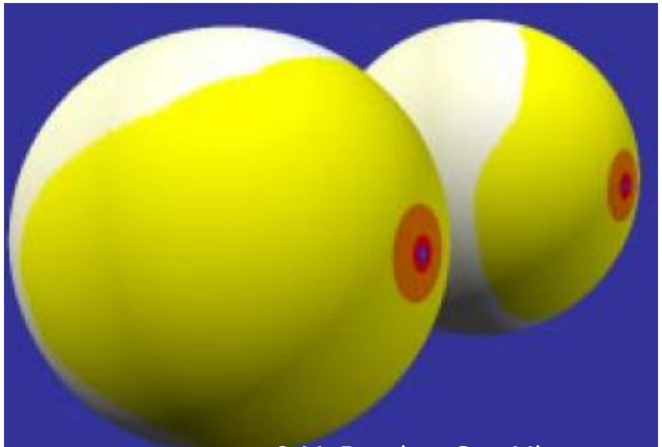


Human Visual System

- Pupil in the Iris
  - Acts as an aperture
  - Adjusts automatically to amount of light
  - Pupil is small in daylight, large in the dark
- Lens
  - Focuses the light onto the retina
  - Accommodation: eye muscles change focal length to keep image sharp
- Retina
  - Senses the incident light intensities
  - Photoreceptor cells on the outermost layer:
    - 10^8 rods (black-white sensitive)
    - $5x10^6$ cones (colour-sensitive)
  - Highest density of photoreceptors in the fovea
- Optic nerve
  - sends encoded light information to the visual cortex in the brain

Visual Field and Visual Acuity

- Human visual field
  - Roughly 220° horizontally, 120° vertically
- There are 10^8 photoreceptor cells on the retina
  - How many megapixels??
- Uneven density
  - 50% concentrated in the fovea (detail sensitive)
  - 50% in rest of retina for peripheral vision (motion sensitive)
- Fovea
  - 1mm wide ➔ 5° of scene in high resolution
  - 6/6 vision ➔ *just* separate bright/dark lines at 0.01° spacing
- Blind spot (scotoma)
  - No photoreceptors where optic nerve passes over retina
- UI considerations
  - Useful limits on resolution of displays? (think iPhone “retina” display)
  - What field of view is appropriate when displaying scene in a first-person shooter?

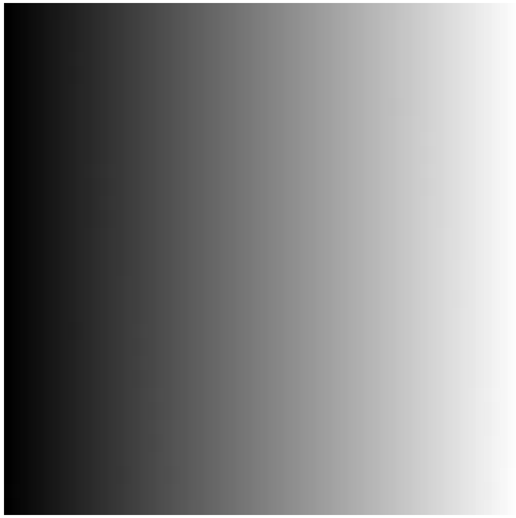


© M. Deering, Sun Microsystem

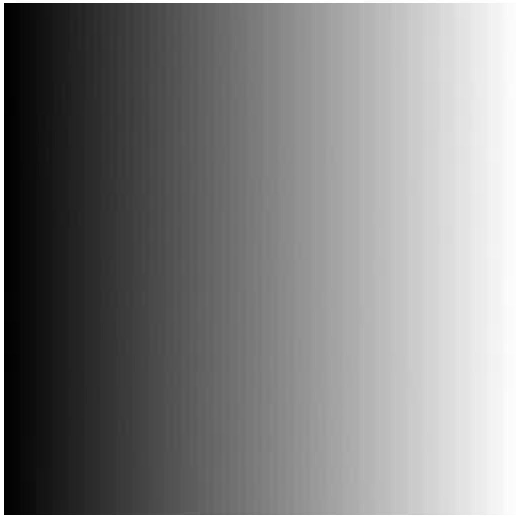
From http://pyroteq.com/field-of-view-first-person-shooters/

Brightness Sensitivity

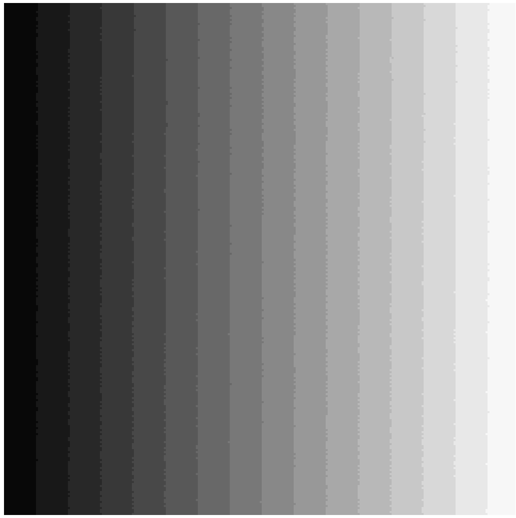
- Brightness "Just Noticeable Difference" (JND)
  - The smallest difference between 2 brightness intensities that is noticeable to humans
  - for normal computer screens, humans can see around 60 different gray shades
- Digital images are typically quantized to 256 intensity levels (8 bits) per RGB color channel
  - Prosumer DSLRs can do 12 bits ➔ allows for post-processing without quantization artifacts



256 gray levels



64 gray levels



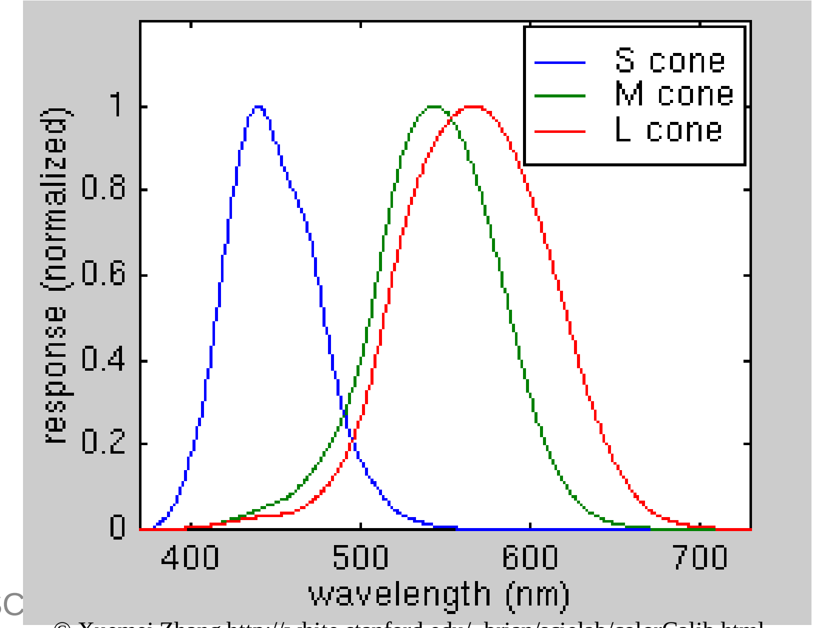
16 gray levels

Persistence of Vision

- Persistence of Vision
  - A *positive* afterimage (ghost image) seen for short time after the original image is removed
  - Duration typically 1/25 – 1/5 s, but depends on intensity and duration of original image
  - E.g. experienced after camera flash
- Negative Afterimage
  - forms after extended viewing of a bright original image e.g. 30 seconds, (greatly depends on original image intensity)
- Flicker-fusion
  - CRT displays need to be refreshed fast >70Hz to avoid flicker and appear constantly bright

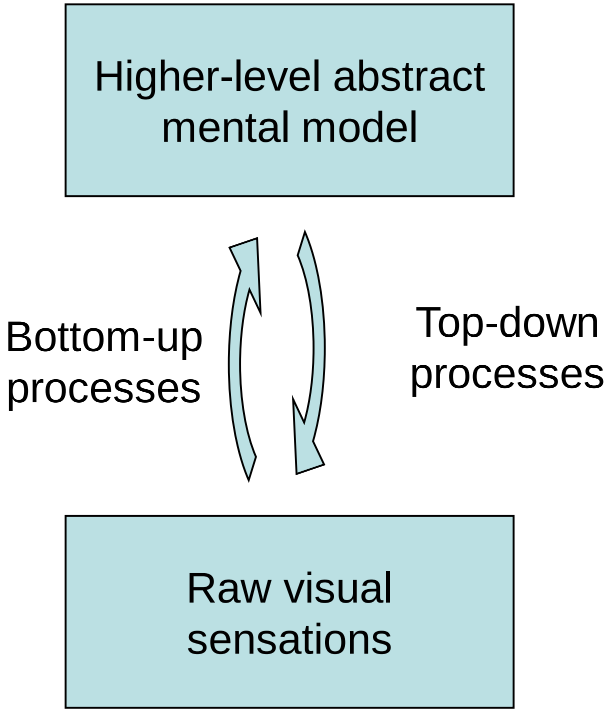
Color Perception

- Visible light is not actually colored
  - Light = combination of EM waves at different wavelengths and amplitudes
  - Color is *purely* human visual sensation
- Humans have 3 types of photoreceptor cones on the retina
  - sensitive to different range of light wavelengths
  - long (L), medium (M) and short (S) -wavelength cones
- Combined stimulation of cones leads to sensation of colors
  - Stimulation of L, M, S cones do not separately lead to sensations of R, G, B
- One goal of display device makers
  - Want to stimulate L,M,S cones in as many different ratios as possible
    ➔ maximize different *color gamut* (i.e. color range)
  - Affects choice of LCD primary colors



Higher-level Visual Perception

- The human mind takes in visual sensations and interprets into a mental model
- Challenging:
  - Raw visual sensations are very incomplete and noisy, even after low-level feature processing in the visual cortex
- Theorized to be carried out a combination of
  - Bottom-up processes
  - Top-down processes
- Bottom-up
  - Combine low-level features or tokens into intermediate and higher-level features
- Top-down
  - Guess or have a prior idea what the correct high-level mental model is, then test to see if it fits the low-level features

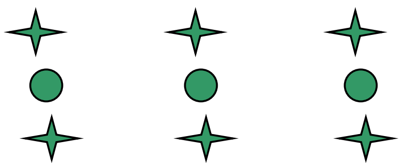


Bottom-Up: Perceptual Grouping

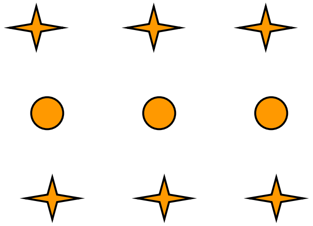
- How to perceptually group low-level tokens into more structured intermediate features?

- **Gestalt Grouping Principles.** Tokens tend to group if they have:

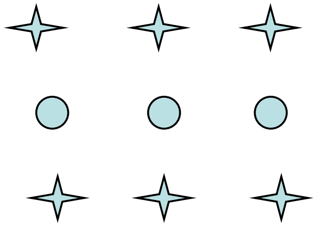
  - Proximity (near each other)



  - Similarity



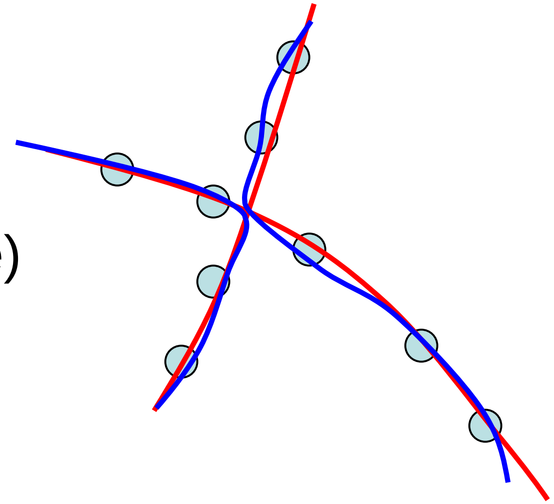
  - Common Fate (moving in the same way)



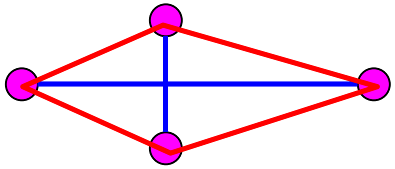
(continued on next slide)

Bottom-Up: Perceptual Grouping

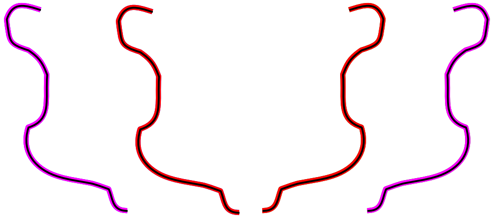
- **Gestalt Grouping Principles (continued)**
  - Good Continuation (lie along a smooth curve)



  - Closure (can form a closed boundary)

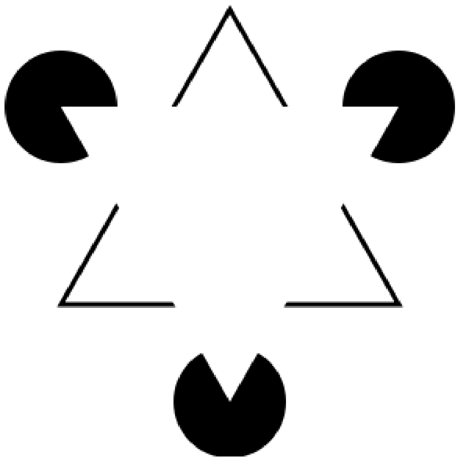


  - Symmetry (can form a symmetrical pattern)



Bottom-Up: Perceptual Grouping

- More examples



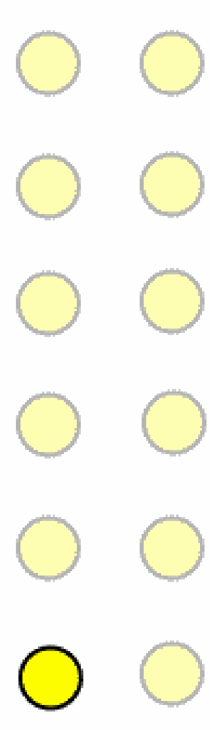
Kaniza's triangle

… illusion

- Relevance of Gestalt Principles to user interface design
  – Guidelines on the spatial layout of user interfaces, so that users can very quickly appreciate which parts of the interface are related to each other

Bottom-Up: Apparent Motion

- Phi Phenomenon
  - Humans will perceive apparent motion from appropriate sequences of *discrete* images
  - Enables:
    - Flip book animation
    - Cinema / television
    - Computer animation



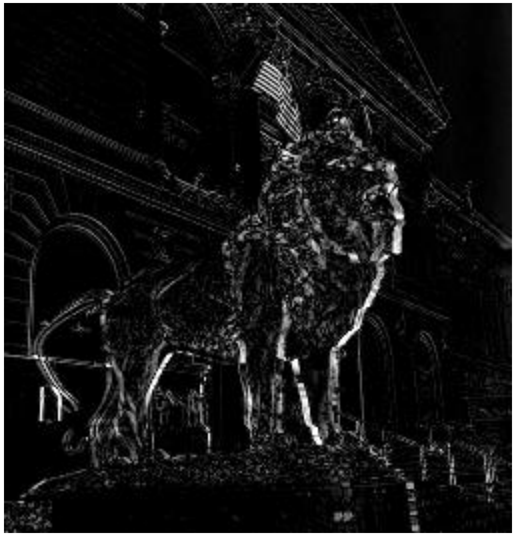
From Phi Phenomenon Wikipedia page, unknown author

- Judder
  - perception of jerky motion due to insufficient frame rate (e.g. in movies, games)
  - Mediated by basic motion interpolation at display end (e.g. this function available in HDTV's), or by applying motion blur to original frames
  - Example: http://www.youtube.com/watch?v=j3ydVKYE9fM#t=48s

Bottom-Up: Stereopsis

From http://2d-3d-movie-tips.blogspot.sg/2015/10/watch-youtube-3d-with-gear-vr.html

- Stereopsis is the ability to perceive depth (i.e. to see in “3D”)
  - due to humans having binocular vision
- Each eye sees a slightly different image
  - as they have slightly different viewpoints
- The mind interprets the disparities (spatial offsets) between the 2 images and fuses into a single 3D mental picture



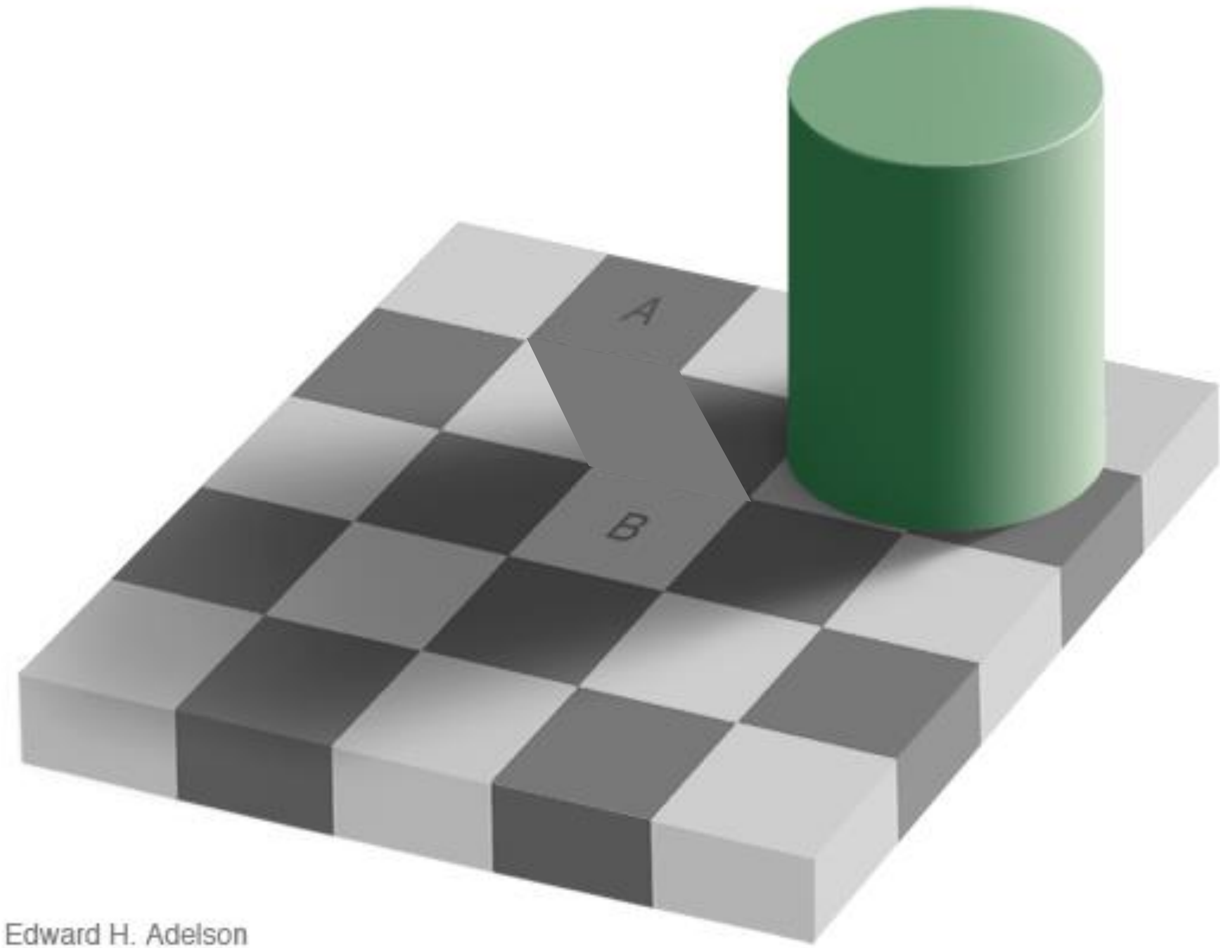
Subtracted image difference

Top-Down: Context in Visual Perception

- Context is very important: used as a prior conditioning to overcome noise and ambiguity in low-level features
  - Part of the top-down visual perception process

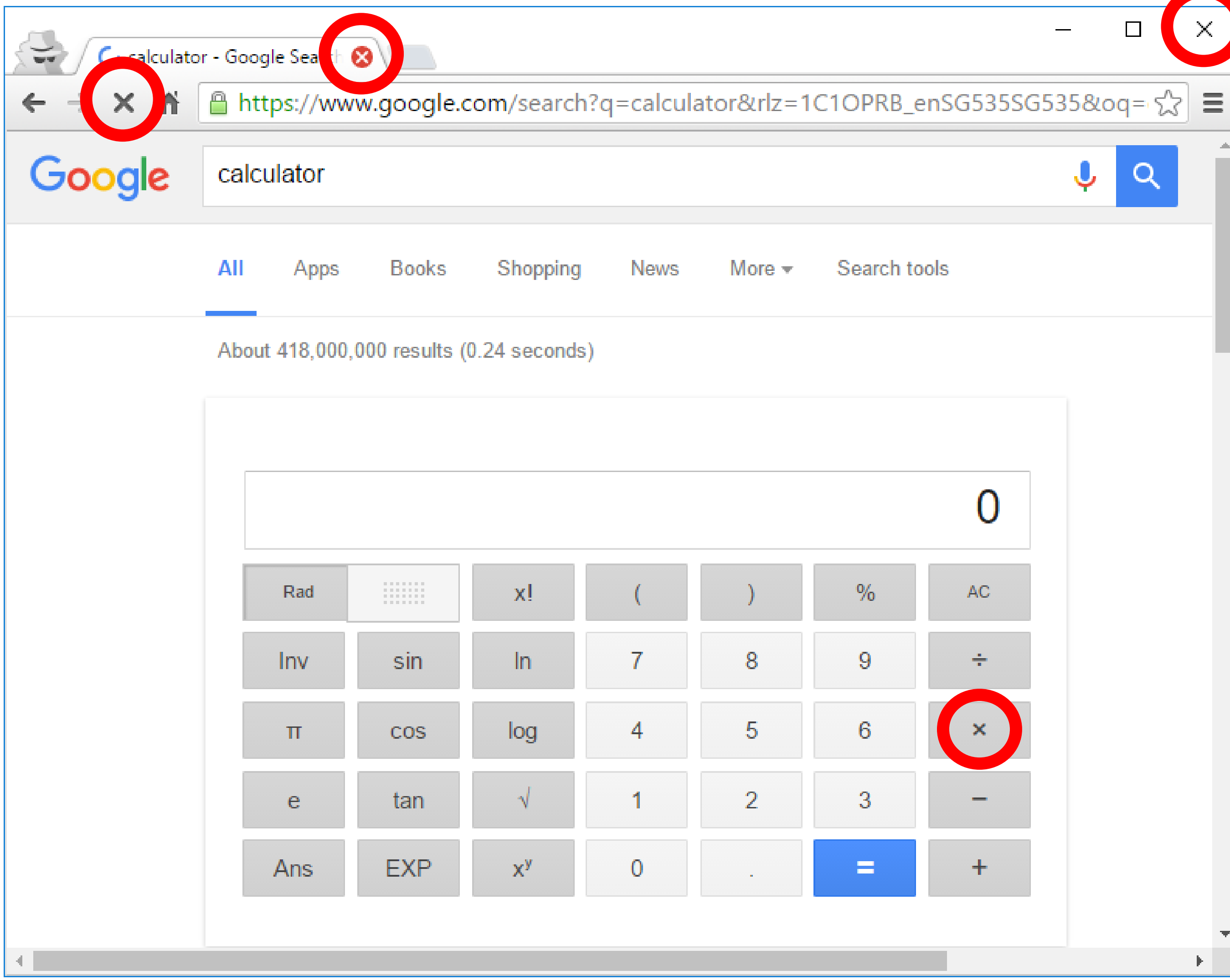
Top-Down: Context in Visual Perception

- Once the higher level scene interpretation is made, the human mind applies heavy context conditioning
  - even on low-level perception



What Does `X' Mean?

- It depends on the **context**



Visual Reading

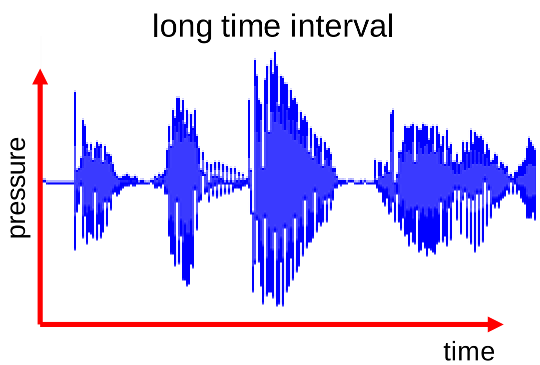
- Adults can read at around 250 words per minute

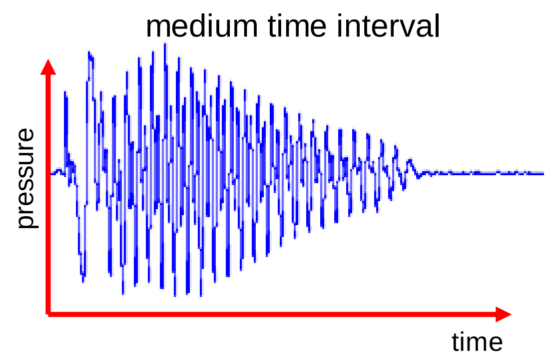
- Text is very heavily structured
  - In English, 26 visual patterns ➔ letters of the alphabet
  - Letters only appear in very constrained sequences when forming (correctly spelled) words
    - For regularly used words, humans can recognize the whole word as fast as letters
  - Sequence of words also heavily restricted by both grammar and meaning

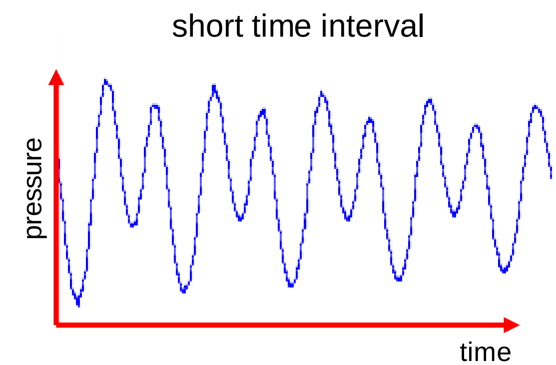
- Enables very strong use of context for understanding text, even those with poor legibility and errors
  - E.g. *"i cdnuolt blveiee taht I cluod aulaclty uesdnatnrd waht I was rdanieg. The phaonmneal pweor of the hmuan mnid, aoccdrnig to a rscheearch at Cmabrigde Uinervtisy, it dseno't mtaetr in waht oerdr the ltteres in a wrod are, the olny iproamtnt tihng is taht the frsit and lsat ltteer be in the rghit pclae.."*

Sound

- Sound is a series of cyclical air (or other medium) pressure changes
- Plotting pressure over time ➔ can be visualized as waves

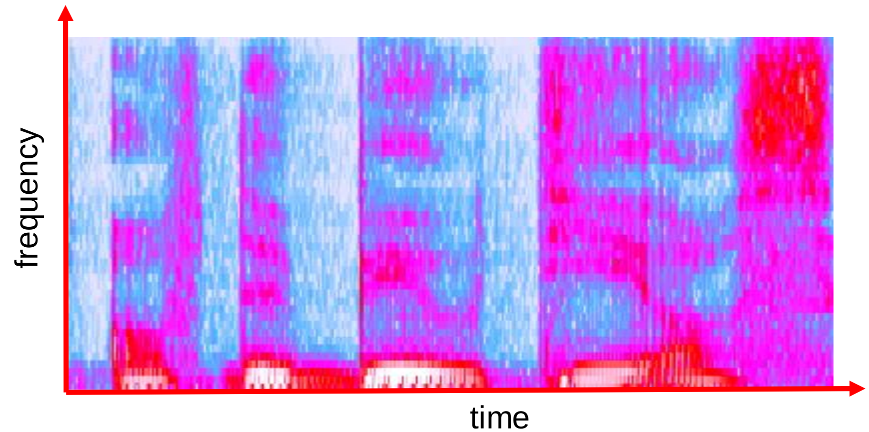






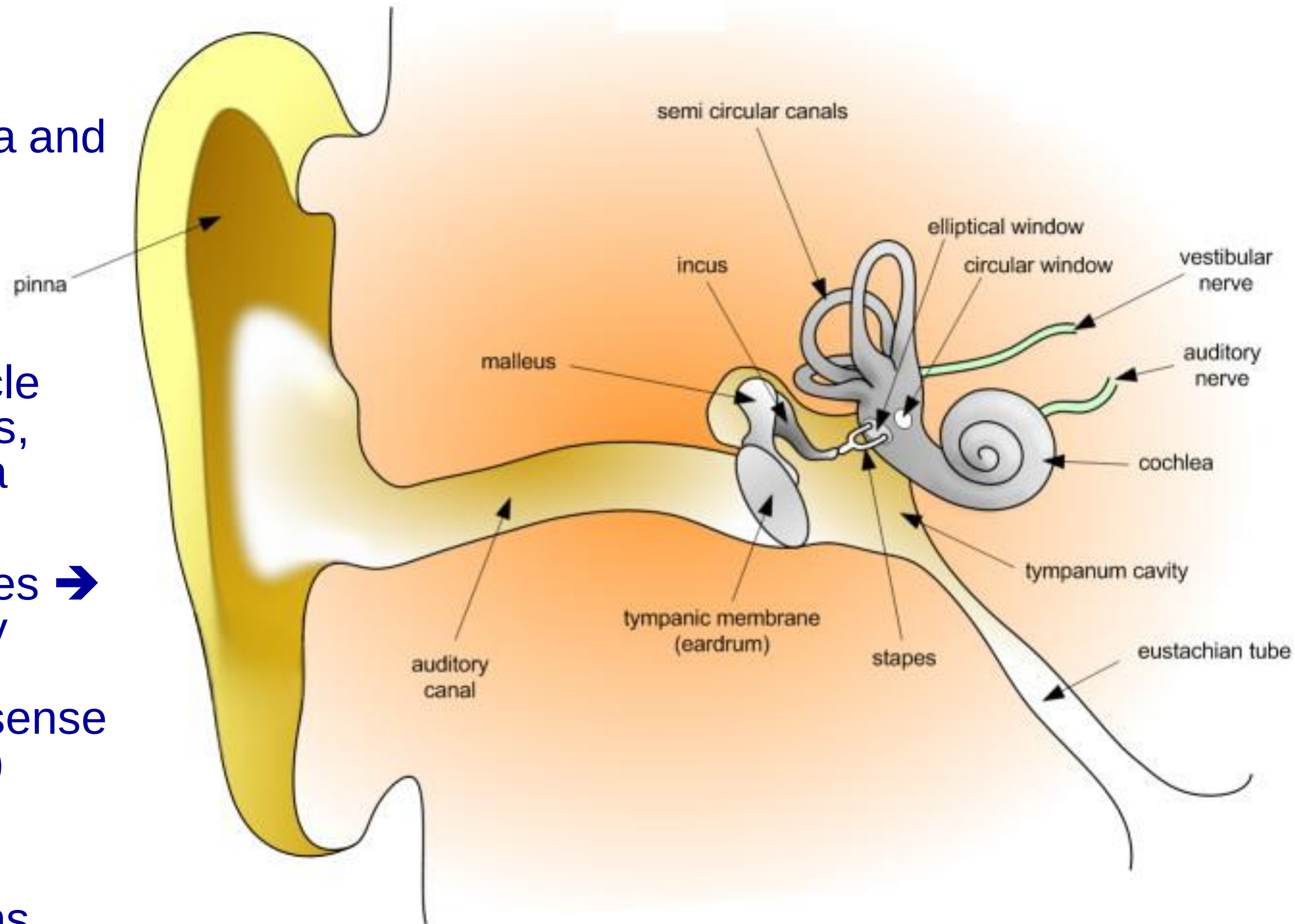
- A more useful visualization: plot frequency components over time

Color denotes amplitude of audio component at a particular frequency



Human Ear and Hearing

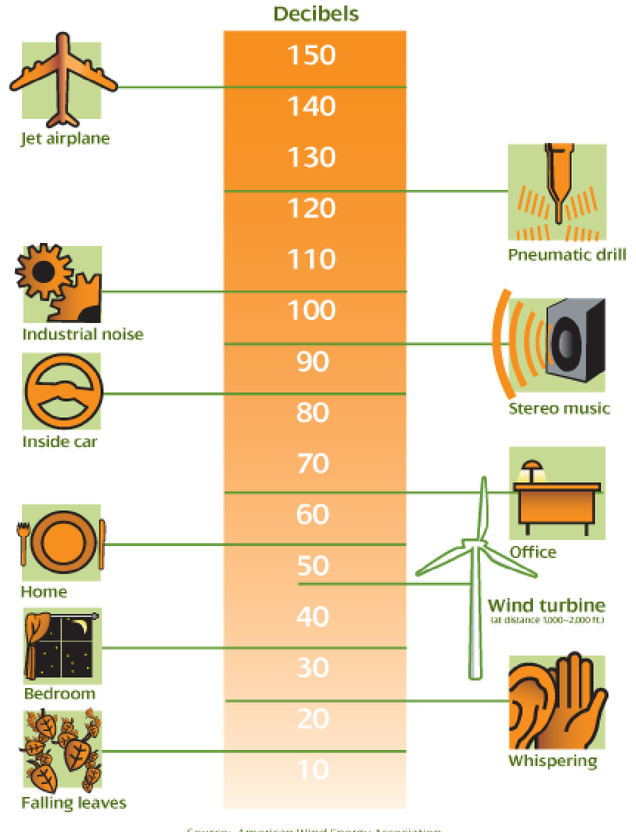
- Sound waves are collected by the pinna and channeled down the auditory canal
- Eardrum vibrates ➔ transmits to the ossicle bones (malleus, incus, stapes) ➔ to cochlea
- Fluid in cochlea moves ➔ movement sensed by auditory hair cells (of different stiffness to sense different frequencies)
- Hair cells send information to neurons



Public domain image, *Wikimedia Commons*

Auditory Perception

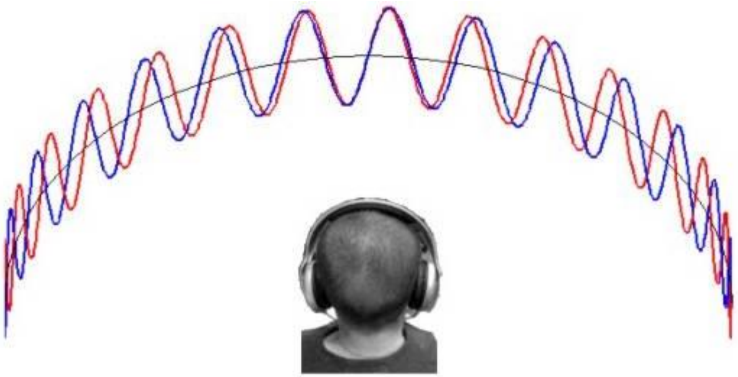
- Sensitivity of human ear: 20Hz to 20kHz (hear http://www.youtube.com/watch?v=fSfc21RDkbw)
  - Upper range depends on age
    - Most adults < 16kHz
- One interesting application
  - "The Mosquito" device (www.compoundsecurity.co.uk/security-information/mosquito-devices)
    - Used to deter teens loitering in public areas by sending out a continuous loud high pitch noise at 17kHz
    - Older adults can't hear this
- Psychoacoustic qualities
  - Pitch – perception of sound frequency
  - Loudness – perception of sound amplitude
  - Timbre – perception of sound "quality", related to overtones and sound envelope
- Loudness is subjective, but the sound pressure level measure (in decibels) is normally used
  - Every 10dB roughly doubles the loudness
  - 0dB is threshold of hearing



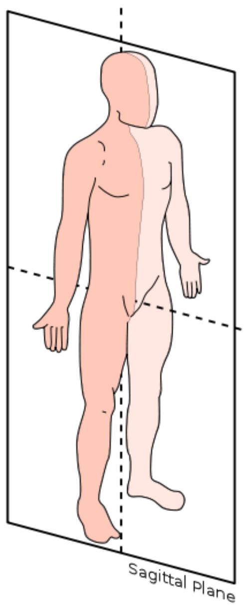
From http://www.iberdrolarenewables.us/images/sound-graphic.png

Sound Localization

From http://mustelid.physiol.ox.ac.uk/drupal/?q=node/59



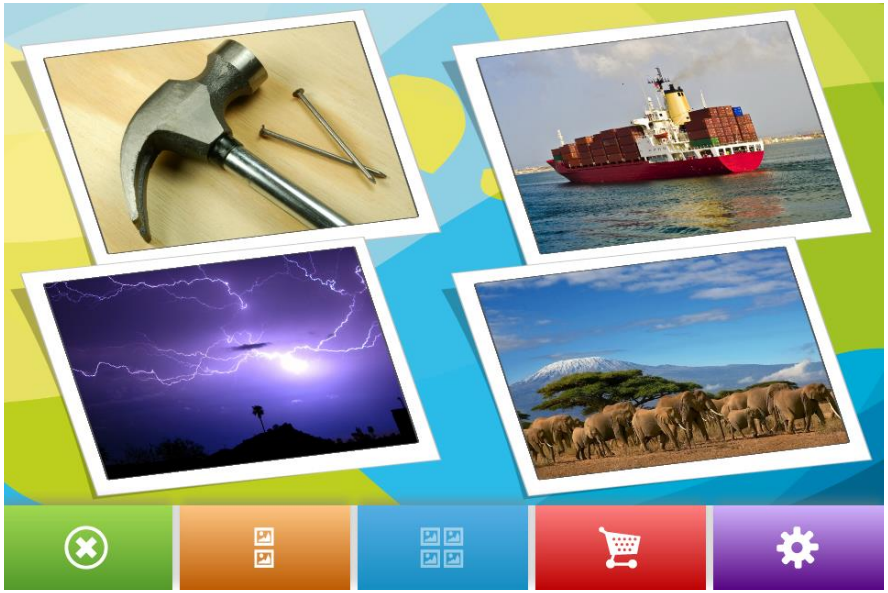
- Lateral (left-right) localization
  - Binaural (using 2 ears)
  - High frequencies: loudness difference between ears
  - Low frequencies: phase / envelope time-shift between ears
- Sagittal plane localization
  - Monaural (individual ear)
  - The folds in the pinna create a frequency notch filter – the ***pinna notch***
  - Different sound directions ➔ different specific frequencies are suppressed
  - Auditory system estimates direction by detecting which frequencies are suppressed



Q: What's the impact on headphone 3D audio systems for user interaction?

Sound Source Recognition

- Humans have ability to recognize environmental sounds
  - Recent research shows two stages of perception:
    - Identifying broad category of the sound source
    - Identifying specific sound source
  - Samples https://www.youtube.com/watch?v=Ii-oNQ2aTmc
- Other forms of sound source recognition
  - Identity of speaker by voice
  - Music

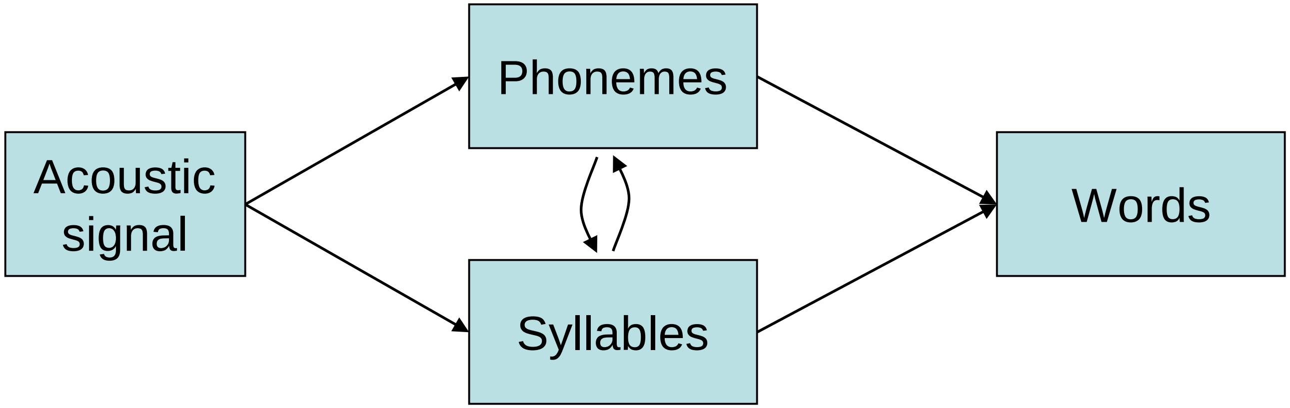


"What's that Sound?" app

Speech Perception

- Humans segment speech audio stream into phonemes and syllables
  - (for phonemes, see http://www.youtube.com/watch?v=BqhXUW_v-1s)

- These are further grouped into higher units such as words and sentences



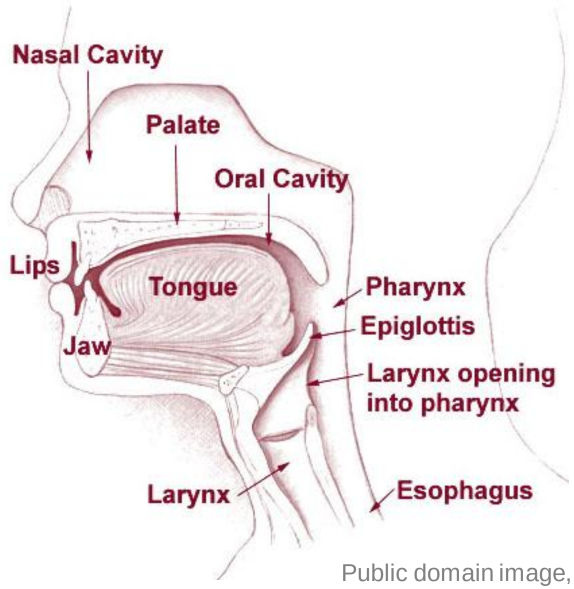
- But there are many other cues used, including co-articulation, intonation, syllabic stress, etc., even visual cues
  - McGurk effect, see http://www.youtube.com/watch?v=G-lN8vWm3m0

Smell and Taste

- Chemoreceptor cells that respond to chemical signals

- Odor chemicals are sensed by olfactory receptor cells in the nasal cavity
- Theory of 7 primary odors (not well established):
  - Camphor (mothballs), musk (perfume), floral, mint, ether (alcohol), pungent (vinegar) and putrid (rotten eggs)
- Scent synthesis https://www.youtube.com/watch?v=oDQpF3NWoJQ

- Taste sensed mainly through taste receptors and ion channels on taste buds
- 5 primary tastes
  - Sweet (receptor), bitter (receptor), salty (ion) sour (ion), savory / umami (receptor)
  - Taste synthesis https://youtu.be/TXIjIyPiXxE?t=1m56s



Public domain image, *Wikimedia Commons*

- Other mouth sensations also contribute to overall gustation experience
  - Temperature, texture, stiffness, etc.

Touch, Temperature and Pain Senses

Public domain image, *Wikimedia Commons*

- Sense of touch is experienced through different types of sensory mechanoreceptors
  - Pressure
  - Tension / stretching
  - Vibrations at different frequencies (for sensing textures)
  - Fujitsu prototype https://www.youtube.com/watch?v=twOHmofWzH0

- Temperature is sensed through thermoreceptors

- Pain is sensed by different receptor types
  - Some are similar to previously described, but with higher activation threshold, e.g. to detect excessive pressure, temperature, light, etc.
  - Other specialized pain receptors, e.g. for detecting skin cuts, and chemicals such as chilli.

Sense of Balance (Equilibrioception)

- Balance and acceleration is sensed through the vestibular system
  - Located in the inner ear
  - Motion sensory cells are found in in:
    - *Semi-circular canals* for rotational acceleration
    - *Saccule* for linear vertical acceleration
    - *Utricle* for linear horizontal acceleration

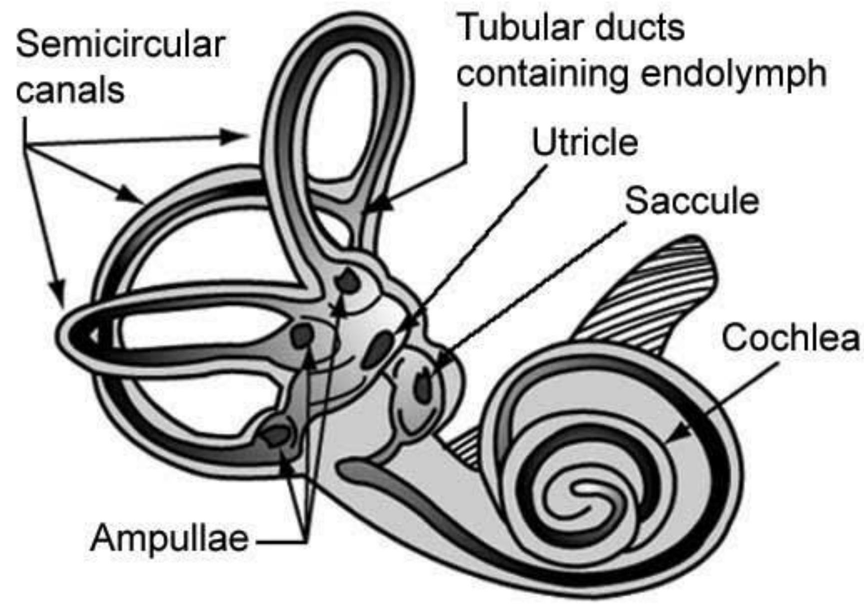


Figure 2: The Vestibular System - semicircular canals and otolith organs

Proprioception

- Sense of one's body position, or pose
  - Includes relative positions of limbs, fingers, etc.
- Proprioception is achieved through various combinations of sensory input:
  - balance
  - muscle stretch

www.nbba.org

- These are specialized receptors for proprioception, rather than shared with other senses

- *Touching fingers experiment*

Human Attention

- Humans have limited mental resources
  - attention is the priority channeling of these resources to specific tasks
- Two forms of attention control
  - Bottom-up, stimuli-driven attention grabbing
  - Top-down, goal-driven focus of attention

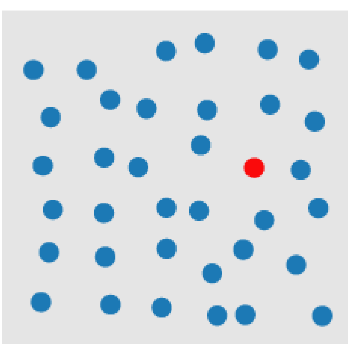
Bottom-Up Attention Control

- Pre-attentive processing in the brain determines what stimuli is relevant to be raised to the conscious mind
  - attention drawn to *changes* and *outliers* in the stimuli
  - carried out subconsciously, fast, with low effort
  - neurons carry out this processing in parallel

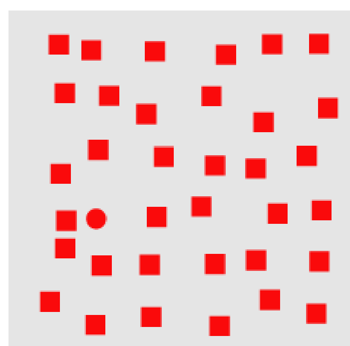
- For visual features, depends on:
  - type of features
  - how much variation there are in the common features
  - how different is the outlier feature

Task: find the red circle in each image

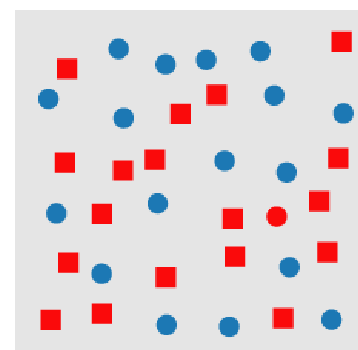
From Chris Healey's page http://www.csc.ncsu.edu/faculty/healey/PP/index.html#Table_1



pre-attentive



pre-attentive



not pre-attentive

Bottom-Up Attention Control

- Some other examples:

Spot the face

pre-attentive

not pre-attentive

- See Stasko's video http://www.youtube.com/watch?v=UFNzATczkDU

Bottom-Up Attention Control

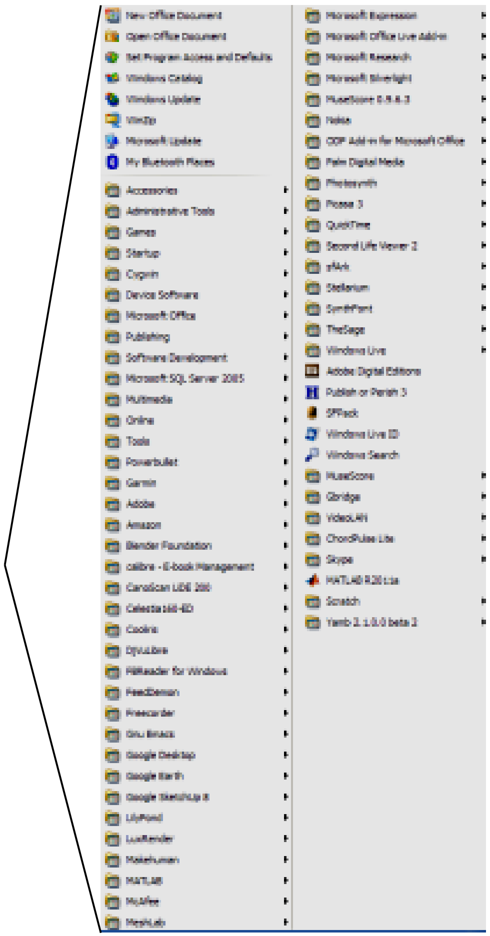
- Significance in UI design
  - good designs need to incorporate visual features where important functions/information stand out *pre-attentively*
  - do not want to overwhelm the user with necessary variation

http://www.lowpriceskates.com/

http://www.hardgraft.com/

Top-down Attention

- Top-down attention requires conscious effort
  - Focused subtasks are carried out sequentially
  - The cost of the effort is felt
- Sometimes this is useful
  - Cocktail Party Effect: ability to follow a single conversation despite other conversations taking place
- For UI design
  - Want to minimize need for top-down attention for *secondary* tasks, e.g. navigating long menus
- During top-down attention focus, other bottom-up attention mechanisms will be diminished
  - known as selective attention
  - See the Monkey Business Illusion
    http://www.youtube.com/watch?v=IGQmdoK_ZfY



Higher-Level Cognition and Affect

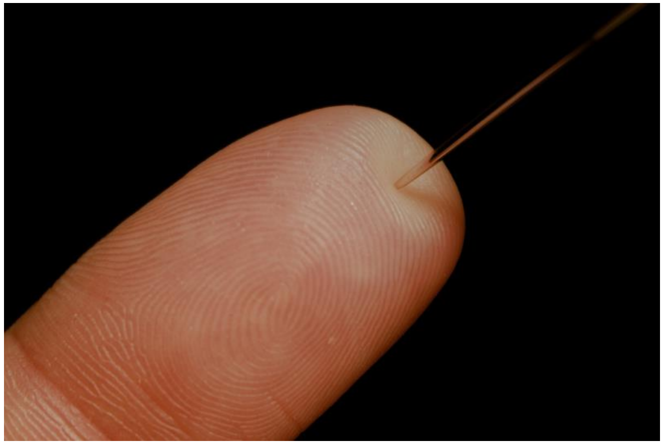
Memory

- Different types of memories, over different time intervals
  - Sensory Memory
  - Short-Term Memory
  - Long-Term Memory

Sensory Memory $\xrightarrow{\text{Attention}}$ Short-Term Memory $\xrightarrow{\text{Practice}}$ Long-Term Memory

Sensory Memory

- Very short-term memory buffer at sensory level
  - Typically < 1s
- Different sensory memories
  - Iconic (sight)
    - e.g. persistence of vision is iconic memory
  - Echoic (hearing)
  - Haptic (touch) memories



Short-Term Memory

- Short-term memory used as working / scratchpad memory
  - Example 1: remembering intermediate products in long multiplication
  - Example 2: remembering one-time password to access online banking system
- Can be quickly formed and accessed, but also quickly lost (in seconds)

- George Miller's "The Magical Number 7, ± 2"
  - Well-known psychology paper published in 1956
  - Humans can remember 7± 2 chunks of information
    - similar for visual, auditory, taste, etc.
  - "chunks" refer loosely to a cluster of items
    - e.g. 1 chunk can be 1 number, or a small group of numbers

Short-Term Memory

- There are ways to help improve short-term memory:
  - Spatial chunking
    - Example: remember "24745038391"
    - Easier to remember "192 – 34 – 934 – 237"
    - Hence phone numbers should be written in chunks
  - Familiar context
    - Example: try remembering

HEC ATR ANU PTH ETR EET

- Easier if you mentally do a circular right shift and rechunk
  - Now easier to remember because the new text is in a *familiar context*: a meaningful English sentence
  - Need reverse procedure to get back original text, but easier to remember the new text *and* the procedure

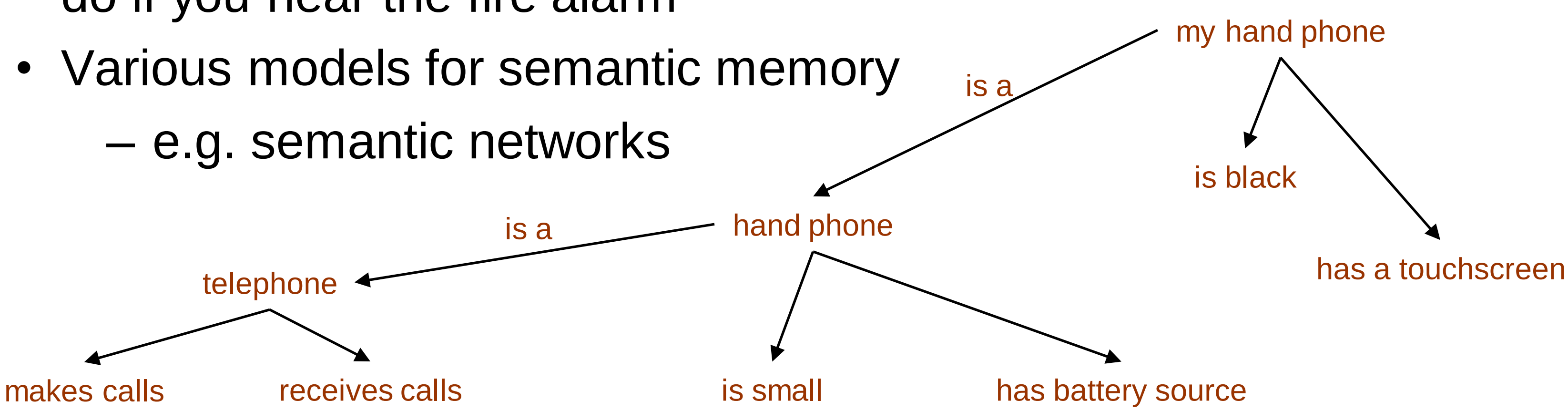
Long-Term Memory

- Long-term memory is slow to form and access, but will last longest
  - Research unclear whether forgetting is loss of memory or loss of index (i.e. loss of ability to retrieve).

- Two types:
  - Declarative Memory
    - Relates to facts, knowledge and experiences
    - “remembering what”
  - Procedural Memory
    - Relates to procedures and skills
    - “remembering how”
    - e.g. how to move around in a first person shooter game, or how to ride a bike

Declarative Memory

- Declarative memory is further divided into two types.
  - Episodic Memory
    - serial memory of events and experiences
    - e.g. events from your first trip to Universal Studios
  - Semantic Memory
    - Structured memory of concepts, attributes, inter-concept relationships and conditional rules
    - e.g. facts like URL of your favorite website, rules like what to do if you hear the fire alarm
    - Various models for semantic memory
      - e.g. semantic networks



Long-Term Memorization

- Long-Term Memorization is done through repetitive practice
- Factors that affect memorization speed / effectiveness
  - Total time
    - more time practicing = stronger memory
  - Distribution of practice effect
    - practice done in intervals is more effective
  - Familiar context
    - Easier to remember things related to familiar systems
    - E.g. harder to remember strong passwords:
      - "*k34!*dl)e2s5*"
    - easier to remember strong nonsensical pass phrases:
      - "*bouncy guts to eat traffic*"

(continued next page)

Long-Term Memorization

(continued from previous page)

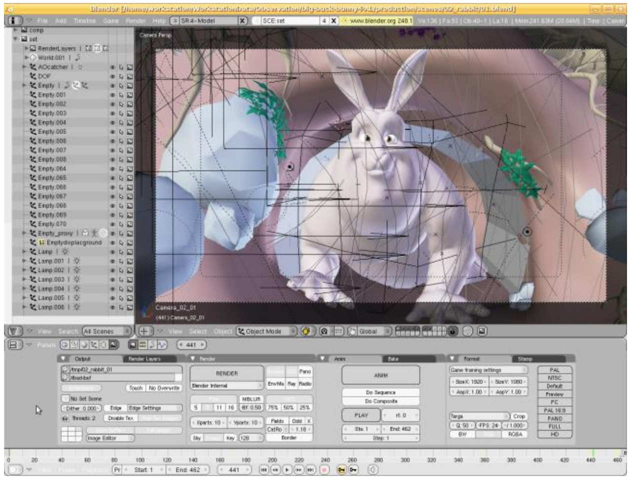
- **Mnemonics** and mnemonic link systems
  - Map concepts to easily remembered constructs
  - Examples:
    - Sentences, e.g. for resistor color codes

      Black Beetles Running On Your Garden Bring Very Good Weather

    - Visual imagery, e.g. 9-multiplication finger trick
    - Sequence of imagined events (exploiting episodic memory)
    - Locations in a virtual environment (exploiting spatial memory)

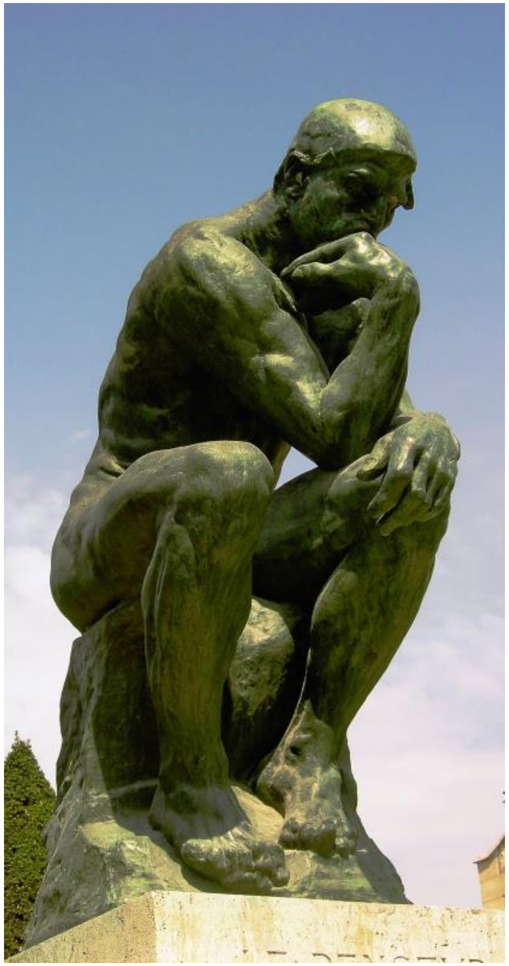
From http://misslepagemaths.primaryblogger.co.uk/

Memory Retention in User Interfaces



- UI's tend to minimize memory load by providing choices, e.g. menu, toolbars, buttons
  - But these may be slower to use in complex apps, e.g. 3D modelers
- Shortcuts can provide speedup for experts, but requires memorization
  - UI's can help users to remember more easily by providing familiar context and/or incorporate mnemonics
- Examples
  - Keyboard hotkeys: trying to use first letter of function name as the shortcut key
    - e.g. in MS Word, Ctrl-B = **bold**, Ctrl-I = *italics*, Ctrl-U = <u>underline</u>
  - Gestural interactions: gestures chosen to relate intuitively to functions
    - e.g. MacBook touchpad gestures

Reasoning



- Reasoning allows conclusions to be drawn
- Deductive Reasoning
  - Reasoning by logic
    Known rule: If it is Sunday, then there is no class.
    Observe: It is Sunday.
    Deduce: Therefore there is no class.
- Abductive Reasoning
  - Guess cause based on effect, i.e. reverse lookup rule
    Rule: If you study hard, you will get good grades.
    Observe: You got good grades.
    Abduce: You studied hard.
- Inductive Reasoning
  - Create general rules or facts based on observations
    Observe: All 5 students I met like this subject.
    Infer rule: If X is a student, X will like this subject.
- Analogical Mapping
  - Create a higher-level *mapping* rule from pairwise observations
    Observe: “cat” is paired with “kitten”.
    Infer mapping rule: adult animal name maps to young animal name.
    (Predictions: “dog” will be paired with “puppy”, etc.).

Reasoning: UI-related Examples

- Deductive Reasoning
  - Reasoning by logic
    Known rule: If you press CTRL-S, the document will be saved.
    Observe: I pressed CTRL-S.
    Deduce: The document has been saved.
- Abductive Reasoning
  - Guess cause based on effect, i.e. reverse lookup rule
    Rule: If you hit the "delete" key, the item will be deleted.
    Observe: I accidentally pressed the keyboard without looking and the item has disappeared.
    Abduce: I must have hit the "delete" key.
- Inductive Reasoning
  - Create general rules or facts based on observations
    Observe: For the past 3 windows, I noticed that when I clicked the "X" button at the top-right corner of the window, it closed.
    Infer rule: For any window, clicking the top-right "X" button will close it.
- Analogical Mapping
  - Create a higher-level *mapping* rule from pairwise observations
    Observe: Hitting "left" key makes the cursor go left.
    Infer mapping rule: Cursor moves in the same direction as arrow on the key.
    (Predictions: Hitting "right" key makes the cursor go right, etc.)

Problem Solving

- Problem solving is how we carry out a complex task without direct instructions
- Problem Space Theory (Newell & Simon, 1963, 1972)
  - A goal state and an initial state
  - Many possible intermediate states
  - Operators to go from an intermediate state to another
- For UI's:
  - "**state**" represents the status or mode of the program
    - e.g. certain checkboxes selected, slide bar in some position
  - "**operator**" is a low-level action
    - e.g. mouse click on a checkbox, click and drag on scrollbar, etc.
  - States and operators can be real or visualized in the user's mind
- Problem solution:
  - *a sequence of operators*
    to go from initial state to goal state

Tower of Hanoi puzzle animation – CC-BY-SA André Karwath

Problem Solving

- Learning the operators
  - Prior known (from manual, instructions, etc.)
  - Trial-and-error
  - Inductive reasoning (how would an operator apply in other states?)
  - Analogical mapping (predicting the functions of different new operators)
- Ways of finding the solution
  - Random or exhaustive exploration of different operator sequences
    - e.g. random walk through a maze problem
  - Heuristics (rule-of-thumb strategies)
    - e.g. hill-climbing approach: “choose next operator that goes to an intermediate state *nearer* the goal state”
  - Planning
    - Decompose the problem into a sequence of subgoals
    - Change plan dynamically depending on whether subgoals turn out to be achievable or not

Problem Solving Example

- *Main goal is to italicize a word in a document. What is the solution operator sequence?*

- Initial state: the word is in normal font style
- Goal state: the word is in italics font style

- Solution operator sequence:
  1. Op: click on left-side of word ➔ state: cursor is next to word
  2. Op: drag cursor across whole word ➔ state: word is selected
  3. Op: click on the *I* button in the toolbar ➔ state: word is italicized

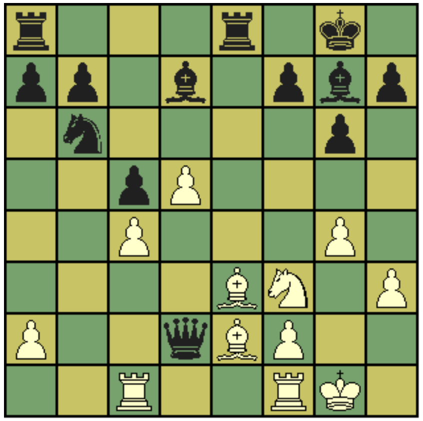
- If you did not know about some operators before hand, e.g. toolbar buttons…
  – you may do trial-and-error clicking of buttons and see what happens, or
  – Click on the **B** button and see that it sets the word in **bold**, so by analogy infer that the *I* button will set the word in *italics*.
  – these processes allow you to learn new operators

Skill Acquisition

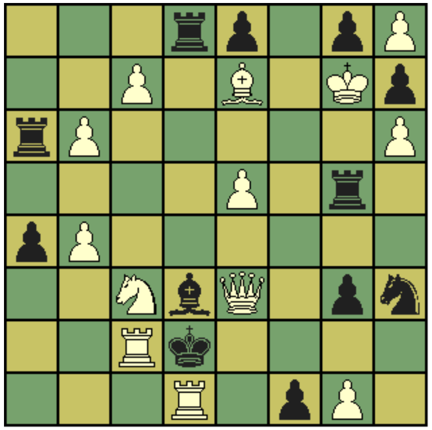
- Ability to problem solve *in a specific domain* becomes better with time and practice
  - Faster and more efficient
  - More accurate

- Chess example (de Groot, 1965)
  - Chess position recall by novices and grandmasters
  - Subjects are shown a board in different states for a short duration
  - Real middle-game configuration:
    - Grandmasters are much better and faster at recalling than novices
  - Random placement of chess pieces:
    - Performances are the same

Real game board: Grandmasters much better at recalling



Random board: Same ability to recall



Skill Acquisition

- ACT-R Theory of Skill Acquisition (Anderson 1983, 1993)

1. Novices learn simple, general purpose rules for choosing operators for different states
    - these rules reside in declarative memory to be constantly looked up – slow access

Compilation

2. Intermediate learners start to compile their knowledge into *domain-specific* composite rules. **Compilation** involves:
    - **Proceduralization** – "hardcode" knowledge into procedural memory
    - **Composition** – make rules for choosing multi-operator subsequences

   Note: these rules only work for initial/goal states similar to those previously encountered, hence "domain specific"!

Tuning

3. Learners become experts as composite rules becomes more tuned for performance and selection of such rules become faster

Mental Models

- A mental model is a user's idea of how a particular device or software works or behaves

- User interfaces should respond in a manner that fits a user's mental model
  - i.e. small (or no) gulfs of execution and evaluation

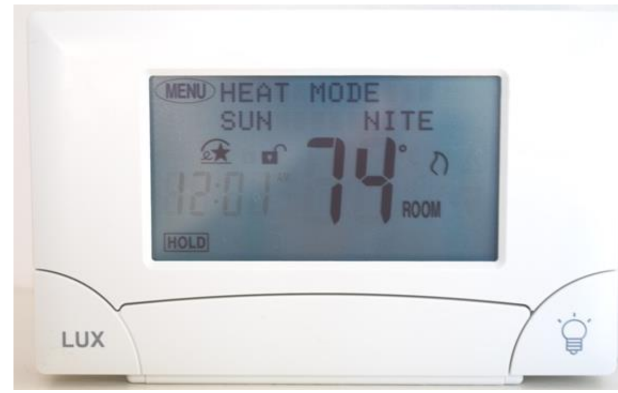
- Examples
  - airbrush function of a paint software
    ➔ colors pixels in a manner identical to user's idea of paint drops landing on the canvas

  - Wii tennis game ➔
    simulation based on user's mental model of how the ball will react when hit

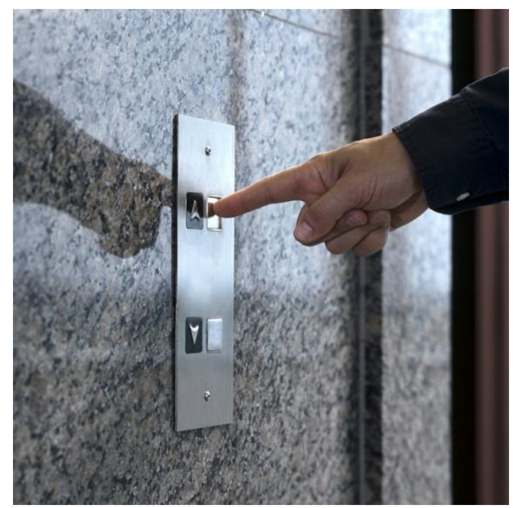
Mental Models

- Examples of errors in mental models:

- Thermostat of room heater (in cold winter)
  - Users like to set thermostats high at the start, then turn down later
    - Mental model of a stove / furnace
  - But heaters are auto full power until preset temperature reached
  - (Opposite situation applies to cold air conditioners)



- Lift call buttons
  - Impatient users press lift call buttons multiple times
    - Mental model of a door knock
    - More presses = more attention



Daniel Dennett's Three Stances

- Humans can have different levels of mental models in their minds:

- Physical Stance
  - Concerned with physical level of detail
  - e.g. "if I move my mouse 1cm at 1cm/s, on the screen the pointer moves 30 pixels"

- Design Stance
  - Concerned with engineering/design level of detail
  - e.g. "drag scrollbar down to read more of the document below"

- Intentional Stance
  - Concerned with behavioral level of detail
  - e.g. "how do I predict and avoid this monster's attack?"

Affect / Emotions

- Affect is a psychology term referring to feeling and emotion
  - Short duration, lasting seconds to minutes
- “Basic” Emotions (Paul Ekman 1999)
  - *Positive*: amusement, contentment, excitement, happiness, pride in achievement, relief, satisfaction, sensory pleasure
  - *Neutral*: surprise
  - *Negative*: anger, contempt, disgust, embarrassment, fear, guilt, sadness/distress, shame
- Each has distinct physiological responses
  - most relevant response is *facial expression*
- There can be composite emotions, e.g.
  smugness = happiness + contempt
- Ekman also proposed the Facial Action Coding System (FACS)
  - decompose facial expression into elemental facial actions, e.g. cheek raiser, lip corner puller, etc.
  - has been partially used to train US airport security personnel

Personality Traits

- Personality traits are permanent / long-term attributes
  - reflect tendency for certain short-term *affect* and/or medium-term *moods*

- Myer-Briggs Type Indicator (related to Jung's model)
  - previously covered

- A later alternative is the Big Five OCEAN Model:
  - **O**penness (curious, creative vs conservative, constant)
  - **C**onscientiousness (self-disciplined, neat, dutiful vs uncaring, sloppy)
  - **E**xtraversion (outgoing, sociable vs reserved, solitary)
  - **A**greeableness (kind, cooperative vs antagonistic)
  - **N**euroticism (anxious, nervous, insecure vs confident, relaxed, secure)

- Determining the personas of target users is an important aspect of UI design

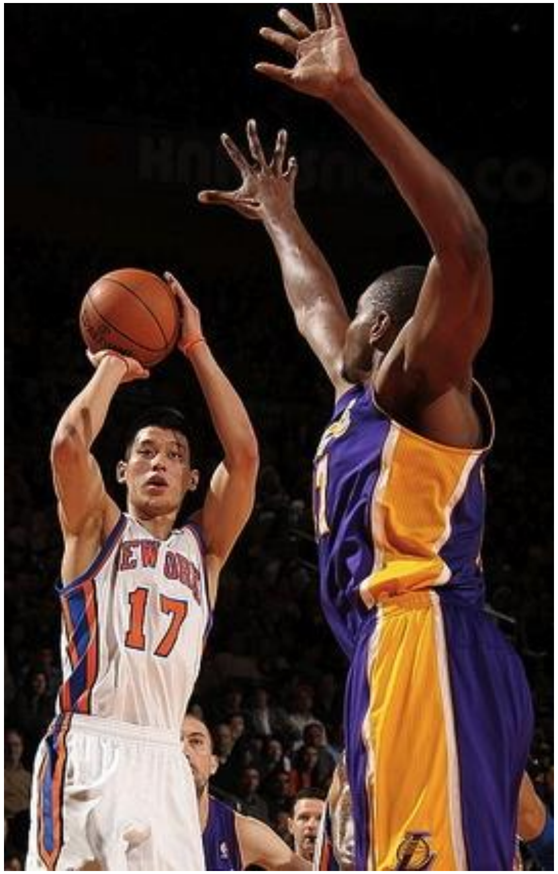
Action and Behavior

Motor Coordination

- A physical action requires activation of multiple muscles
  - (this includes speaking which needs vocal muscles)

- Motor coordination deals with how muscles are synchronously and adaptively controlled
  - Feedback is obtained from the senses
  - Most relevant is *hand-eye coordination*

- Both physical and software UI design need to account for motor coordination ability and limitations of different users

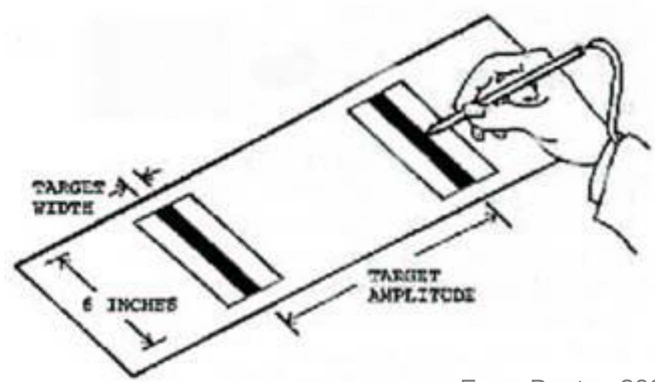


Movement Modeling

- Useful to have models to predict how quickly and efficiently users will be able to use a particular interface design

- Fitts' Law (Fitts 1954)
  - Models how quickly a person can point to a target on a table
  - Has been adapted for computer mouse and other pointing devices

$$T = a + b\log_2\left(\frac{D}{W}+1\right)$$



From Buxton 2003

T = movement time
D = distance needed to move
W = target width
a = start/stop time constant (device dependent)
b = inherent speed constant (device dependent)

- Other models have been used to estimate complex task completion times
  - e.g. the Model Human Processor (Card *et al.* 1986)

Fitts' Law Example

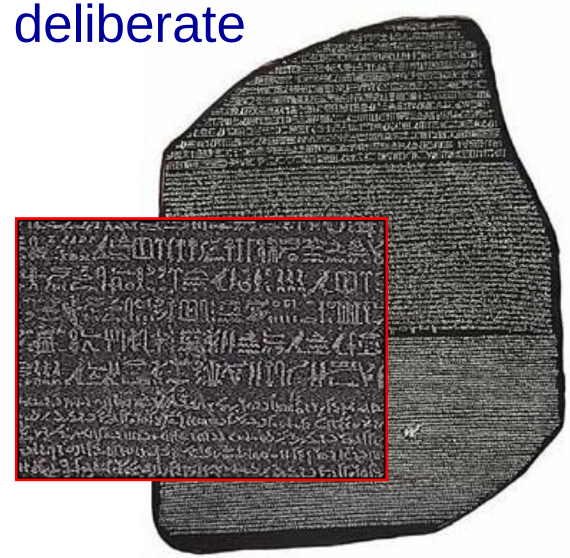
- Suppose a mouse has been empirically measured to have a = 300ms and b = 200ms/bit
- Given a button on the screen of width of 20 pixels, with the initial position of the pointer 140 pixels away.
- Estimate the movement time needed for the user to move the pointer to point at the button

$$T = 300 + 200 \log_2 \left(\frac{140}{20} + 1 \right) = 900\,\text{ms}$$

- Notice:
  - if the pointer was twice as far away, the movement time only increases to 1080ms
  - If the pointer was twice as near, the movement time only decreases to 730ms

Intentional Communication

- Communication is the process of conveying information or meaning
- Intentional communication is communication that is deliberate
  - Linguistic communication
  - Iconic communication
  - Gestural communication

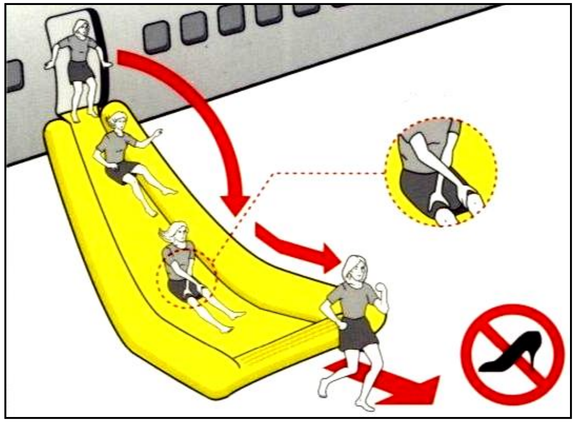


- Linguistic communication uses a language with:
  - Lexicons (symbols), e.g. words
  - Syntax or grammar
    - Rules governing the sequencing of lexicons

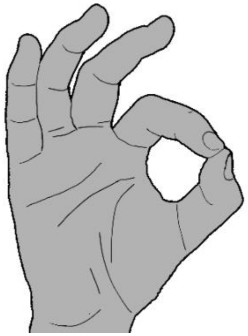
- Types of linguistic communication
  - Speech / oral communication
  - Written text
  - Sign language
    - “Singapore Sign Language” (SgSL) in Singapore

Intentional Communication

- Iconic communication is done via visual imagery, e.g.
  - Sketches, illustrations, visual signs, emoticons
  - Creation, modification or placement of objects
    - e.g. tissue pack to reserve table, threat icons such as human skulls, etc.



- Gestural communication is based on physical action
  - some universal (e.g. pointing, nods, waves)
  - some cultural-dependent (e.g. a-ok sign)
  - may be used in conjunction with speech
  - some tactile (e.g. handshake, tap on shoulder)
  - differs from sign language as there is usually no grammar in normal use of gestures



Non-intentional Communication

- Communicating information subconsciously
  - Body language
  - Vocal paralanguage
- The communication is predominantly about a person's *affective state*
- Body language
  - Head pose, eye gaze and facial expressions
  - Body posture and actions, e.g.
    - crossing arms and legs (aloofness), check watch (impatience), preening actions (romantic interest)
- Vocal paralanguage
  - Intonation, pitch, loudness attributes
    - Can convey emotions, e.g. anger, anxiety, excitement
  - (emoticons can be considered intentional visual paralanguage for text)