1 HCI Foundation 1 - Human

The first, important thing we have to mention when it comes to Human-Computer Interaction is the **human**. We can talk about the human as a system where several information are sent outside and received inside (I/O like visual, auditory, movement...), stored (sensory, short and long term memory...) and processed and applied (reasoning, problem solving...). Emotions can influence human capabilities and such influence depend on the single human being as each person is different. Let's see some of these capabilities.

1.1 Vision

Vision is divided in two stages:

1. physical reception of stimulus
2. processing and interpretation of stimulus

The eye is the physical receptor for vision, obviously. It is a mechanism for receiving light and transforming it into electrical energy for the brain, in fact images are focused upside-down on retina and correctly processed by ganglion cells (present in our brains) which detect pattern and movement. Some parameters that define the characteristics of what we are seeing are:

- Size and depth (visual angle, acuity, cues like overlapping objects which help perception...);
- Brightness (subjective reaction to its levels, affected by luminance of objects...);
- Color (made up of hue, intensity, saturation, different acuity for colors, color-blind people problem...).

However, the visual system is also able to compensate for movement or changes in luminance, but sometimes there is overcompensation which causes optical illusions on some images or videos and so on. Context is resolved to solve visual ambiguity.

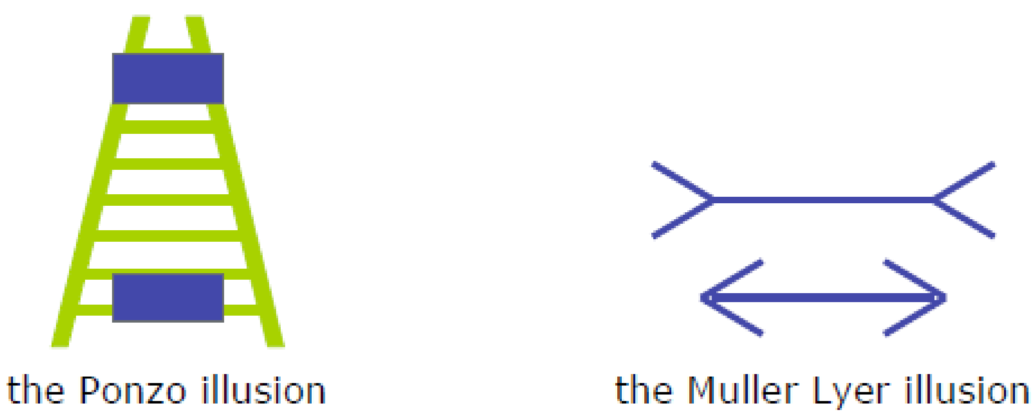

Figure 1: Optical illusions examples.

Reading is something specific we can do with our visual system: in fact we perceive a visual pattern then we give it a meaning with respect to a specific internal representation of language (knowledge of syntax, semantics, grammar...). Word shape is important to recognition and negative contrast is necessary to perceive words and symbols on a computer screen (in fact we always set a light or dark theme with respect to our tastes: the important is to be able to distinguish letters and numbers on the screen we are using!).

1.2 Hearing

Hearing is done thanks to our ears and provides information about environment, like distances and directions of the heared sound. The physical apparatus interested is so composed:

- outer ear, which protects inner and amplifies sound;
- middle ear, which transmits sound waves as vibrations to inner ear;
- inner ear, where chemical transmitters are released and cause impulses in auditory nerve.

The signal we receive is called *sound* and is composed by pitch (sound freq.), loudness and timbre. However, humans can hear frequencies only from 20Hz to 15kHz, so not all the frequencies. Auditory system is capable of filtering sound, distinguishing the sound we are interested in from background noise and other useless stuff (cocktail party phenomenon!).

1.3 Touch

Touch provides important feedback about environment, more than other senses. May be the main sense for environment for someone who is visually impaired! Stimulus are received via receptors in the skin and we have different sensibility with respect to the body part (fingers are more sensible!).

1.4 Movement

Time taken to respond to stimulus is composed by reaction time + movement time. This time depends on age, fitness and other factors and also on stimulus type (visual = 200ms, auditory = 150 ms, pain = 700 ms...). *Fitts' Law* describes the time taken to hit a screen target:

$$Mt = a + b \log_2(D/S + 1) \tag{1}$$

where:

- a and b are empirically determined constants;
- Mt is movement time
- D is distance
- S is size of target

That's why we have to project UIs with targets as large as possible and distances as small as possible.

1.5 Memory

There are three types of memory that we'll now examine:

- Sensory memory: buffers for stimuli received through senses (iconic, echoic, haptic...), continuously overwritten;
- Short-Term memory (STM): scratch-pad for temporary recall with rapid access and decay, and limited capacity;
- Long-Term memory (LTM): repository for all our knowledge (slow access and decay, huge or unlimited capacity). Two types: episodic and semantic. The latter derive from the former. Episodic is related to events we remember, semantic instead provides access to information and is something we use during exams, etc. Semantic LTM represents relationships between bits of information and follows a specific model, called semantic network presented in the next image.

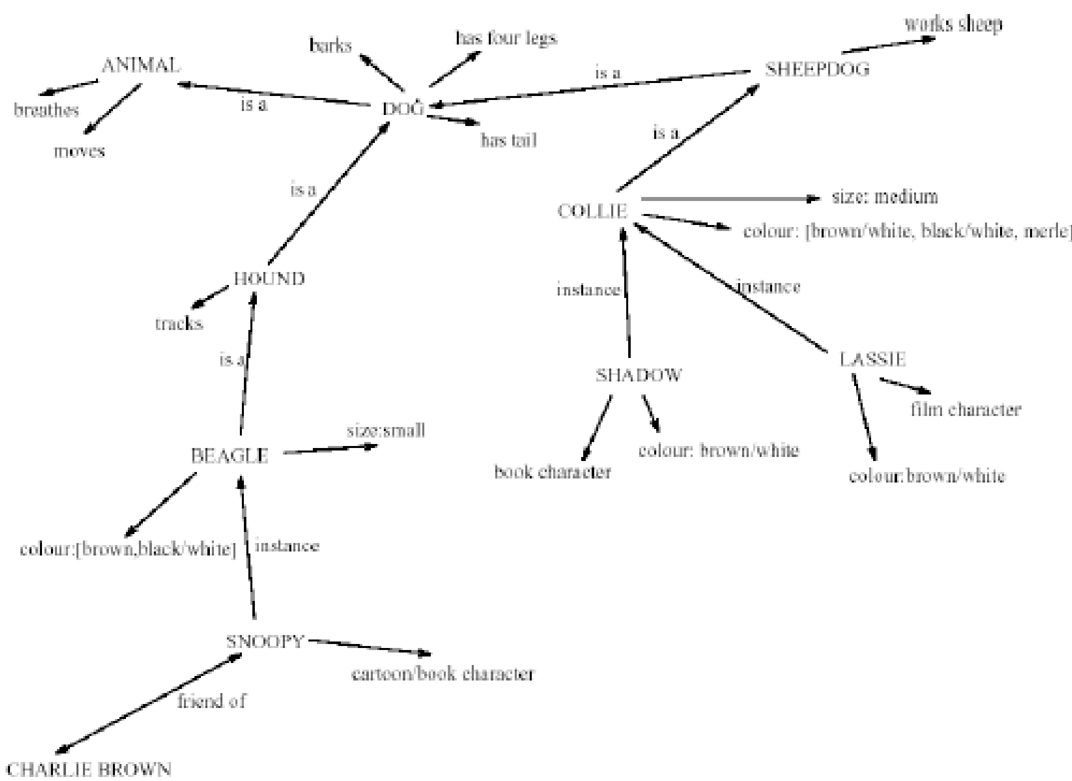

Figure 2: Semantic network.

Information in LTM are organized in data structures called *frames* with some fields as we can see here:

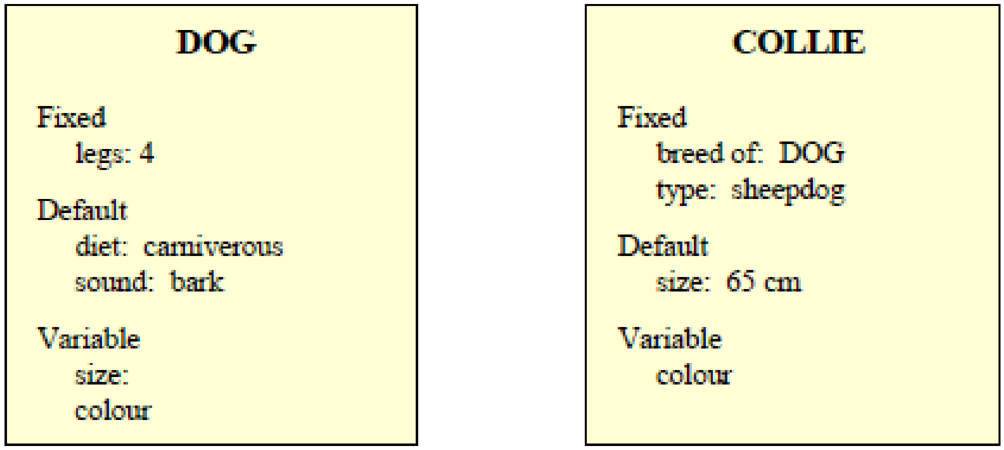

Figure 3: LTM frames.

We can also use *scripts*, which are richer than frames because they not only model knowledge but also the process in which its components are involved.

Figure 4: LTM scripts.

Production rules are instead representation of procedural knowledge, i.e. condition/action rules.

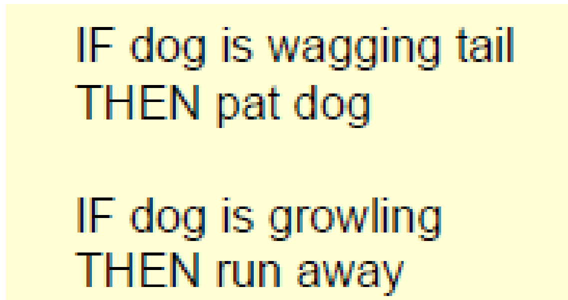

Figure 5: LTM production rule.

From the point of view of actions, what we can do with LTM? We can store information, moving them from STM to LTM and also distributing the learning over time, and giving it a structure to improve such a movement. The immediate consequent action is the retrieval of such information, through recall (remember) or recognition of something we see/hear/experience and so on. Last but not least, we can forget information, for decay (it happened long time ago) or interference (overwriting of info in our brain).

1.6 Thinking

Thinking involves reasoning (deduction, induction, abduction) and problem solving. Let's see them more in detail.

- *Deduction* means derive logically necessary conclusion from given premises/experiences. Logical conclusions not necessarily true (think about syllogisms...). As human knowledge is different from logical deduction algorithms, we bring such knowledge in order to understand if an info is real or not;
- *Induction* consists of generalizing from cases seen to cases unseen, i.e. from particular to general. However it can prove only false and not true, so it is unreliable but still useful. Humans are not good at using negative evidence (Wason cards example!).

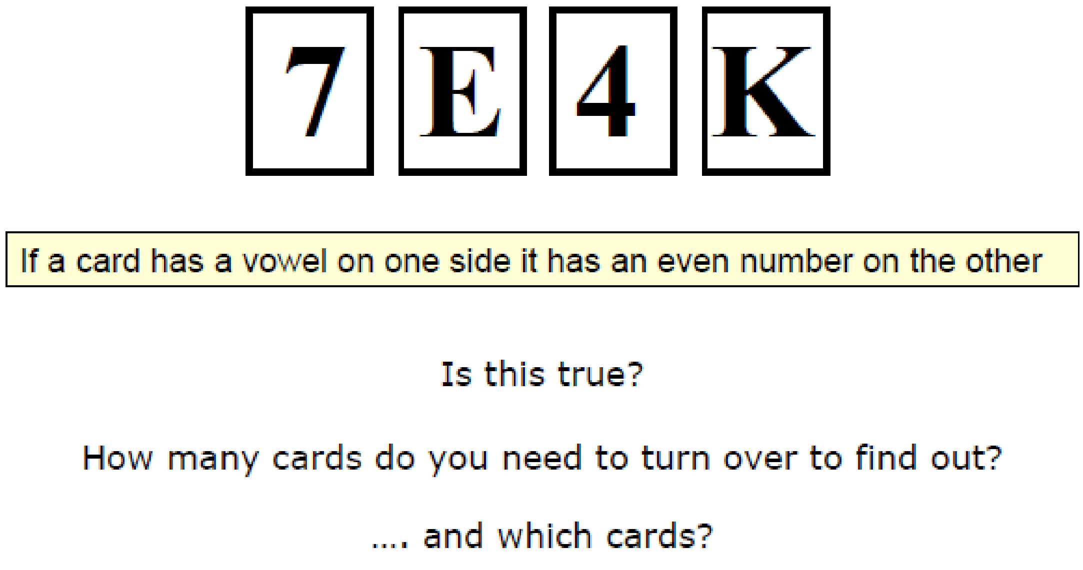

Figure 6: Wason's cards.

The answer is 7 and E, where 7 is the negative evidence. This test shows that humans are not good in using negative evidences (we would have turned over E only, probably).

- *Abduction* is reasoning from event to cause and it is unreliable because it can lead to false explanations.
- *Problem Solving* is the process of finding solution to unfamiliar tasks using knowledge and it is supported by several theories. The problem space theory says that the problem space comprises problem states and the process of problem solving involves generating states using legal operators. Heuristics may be employed to select operators and operates within human information processing system. This model is largely applied to problem solving in well-defined areas. We use *analogy* and *skill acquisition* when we have to solve a problem. However, this is not a perfect method: we can do errors (slips, i.e. right intention but failed attempt, or mistakes, i.e. wrong intention caused by incorrect understanding).

1.7 Emotion

Emotion largely affects human capabilities and it is also a capability itself. There are several theories on how emotion works, like "emotion is our interpretaion of a psychological response to stimuli" by James-Lange. Emotion clearly involves both cognitive and physical responses to stimuli so if we design a system for being emotion-resilient, we have to recreate these emotions also during testing phase. The biological response to stimuli is called *affect*, which influences how we respond to situations (positively or negatively). Stress will increase the difficulty for problem solving and relaxed users will be more forgiving of shortcomings in designs, so the more the user will be stressed, the easier our UI should be in order to give a better UX (aestethically pleasing and rewarding interfaces...).

1.8 Conclusions

There are several differents in each human being like sex, physical and intellectual abilities, age, stress effect and so on. Hence, we have to consider the psychological aspect when designing a system, modeling it with respect to these differences.

2 HCI Foundation 2 - Computer

Nowadays, computers pervade our lives and we regularly interact with them. A computer is made up of various elements such as input devices (text entry and pointing), output devices (screen, digital paper...), memory (RAM, ROM, HDD...) and processing (speed of processing, networks...). In order to understand human-computer interaction we need to understand computers, i.e. what I can do with them. Computers we use nowadays are quite similar, in fact they have a screen, keyboard and pointing device. Some variations can affect the whole system concept such as desktop PCs, laptops and so on. However, each type supports different style of interaction and purpose. There are also small computers in our washing machines, microwaves and so on. Some years ago there was not so much interactivity, but nowadays almost every computer system is interactive. Let's see how we can interact with a computer.

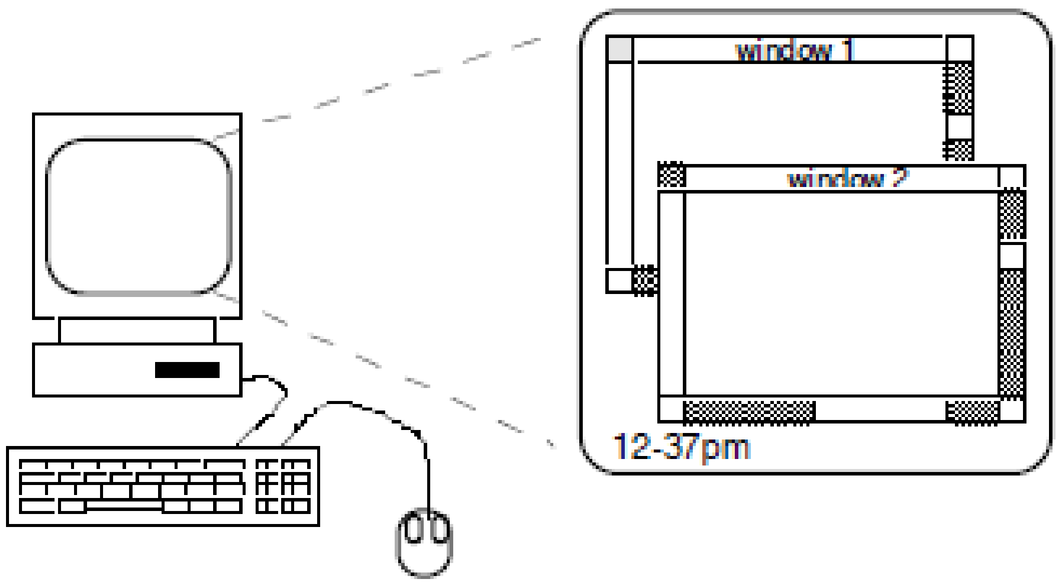

Figure 7: Computer scheme.

2.1 Writing Devices

2.1.1 Keyboards

Keyboards are the most common input devices and allow rapid entry of text by experienced users. Keypress closes connection, causing a character code to be sent and it is usually cable connected but can also be wireless. The most famous layout is the QWERTY one that changes with respect to the country. However, there are also other layouts such as alphabetic orders, etc. or special keyboards for left handed use, phone keyboards with T9 criterion, etc. There are also numeric keypads, i.e. with numbers only.

2.1.2 Handwriting Recognition

Text can be input into the computer, using a pen and a digesting tablet. There are technical problems in this approach, like capturing all useful informations, interpreting individual calligraphies and individual letters from words. This is an approach used on tablets and PDAs.

2.1.3 Speech Recognition

Speech recognition is most successful when there is a single user and limited vocabulary, for obvious reasons. Problems here are related to external noise interfering, pronunciation errors, large vocabularies and different voices.

2.2 Pointing Devices

2.2.1 Mouse

Mouses are handeld pointing devices, very common and easy to use. They permit a planar movement on a display and in the simplest version they have two buttons and a small wheel. They require little physical space and do not cause arm fatigue. Pointers also do not hide screen portions so it is something really useful for human-computer interaction. They work in two main methods:

- Mechanical, older mouses with a sphere on underside of mouse, but usable in most surfaces;
- Optical, with a light emitting diode on underside of mouse and less susceptible to dust and dirt, but sometimes causing issues with certain surfaces.

There have been also experiments with feet (footmouse) but it is something not really common, obviously :D however, foot controls are common elswhere, like in cars or piano pedals.

2.2.2 Touchpad

Touchpads are mostly present in laptop computers and they need a stroke to move mouse pointer. They are small touch sensitive tablets and need stroke settings to work properly (fast or slow).

2.2.3 Joystick

They are more common for computer games and they use pressure of stick for movement and buttons for selection.

2.2.4 Touch-Sensitive Screen

This kind of screen detects interactions through the presence of finger or stylus on the screen.
Advantages:

- fast
- requires no special pointer
- suitable for use in hostile environment

Disadvantages:

- finger can mark the screen
- imprecision
- lifting arm can be tiring

Stylus and light pens are devices which help this kind of interaction and are heavily used in tablets. These pens can also write (handwriting) and perform other interesting tasks, although they can obscure the screen a little bit.

2.3 Display Devices

2.3.1 Bitmap Displays

Bitmap displays are the more common, the ones composed by a vast number of colored dots. We list the main characteristics of such displays:

- Resolution (pixels)
- Aspect Ratio (16:9,...)
- Color Depth

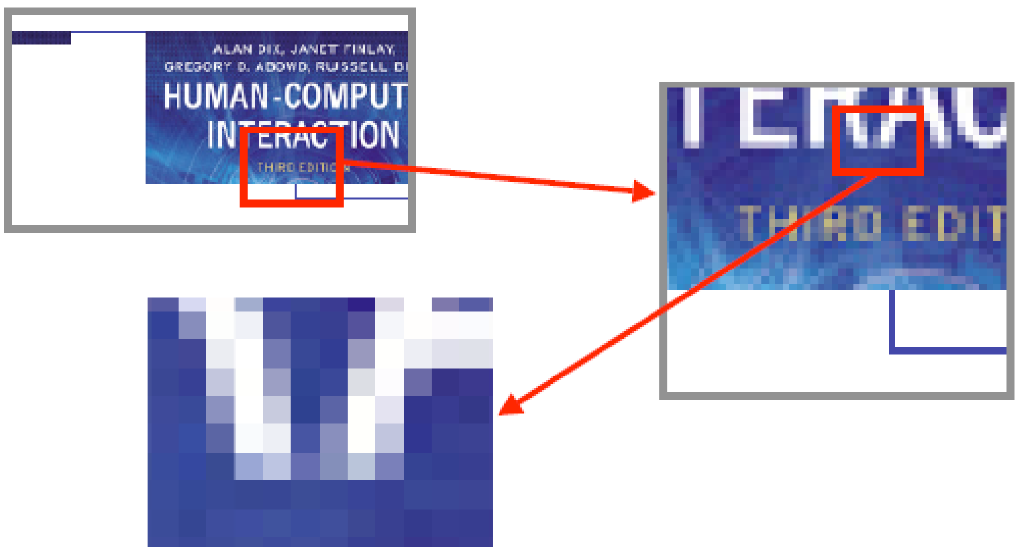

Figure 8: Bitmap display.

2.3.2 Cathode Ray Tube

The system of old TVs: stream of electrons emitted from electron gun, focused and directed by magnetic fiels, hit phosphor-coated screen which glows.

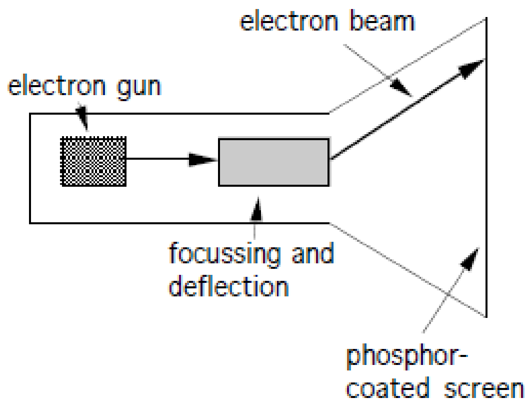

Figure 9: Cathode ray tube.

This technology declined because of the monitors size. It also had some health hazards, like radio frequency emissions and electrostatic and electromagnetic fields related problems.

2.3.3 Liquid Crystal Displays

They replaced CRT monitors because of the smaller size, lighter images and no radiation problems. We find them everywhere nowadays.

2.3.4 Large Displays

They are used for meetings, lectures and so on. Technologies for these displays range from video walls, projected, back-projected etc.

2.3.5 Situated Displays

These displays are put in public places and are sometimes interactive, sometimes not. In all the cases when implementing these displays, the location matters. Think about displays in train stations, etc.

2.4 VR and 3D Interaction

2.4.1 Positioning in 3D Space

In 3D space, positioning is done differently:

- cockpit and visual controls like steering wheels, knobs and dials, like in real life with cars and planes etc.
- the 3D mouse is a mouse allowing 6 directions of movement
- data glove uses fiber optics to detect finger position
- VR helmets
- whole body tracking

Movement is done in three directions through pitch, yaw and roll.

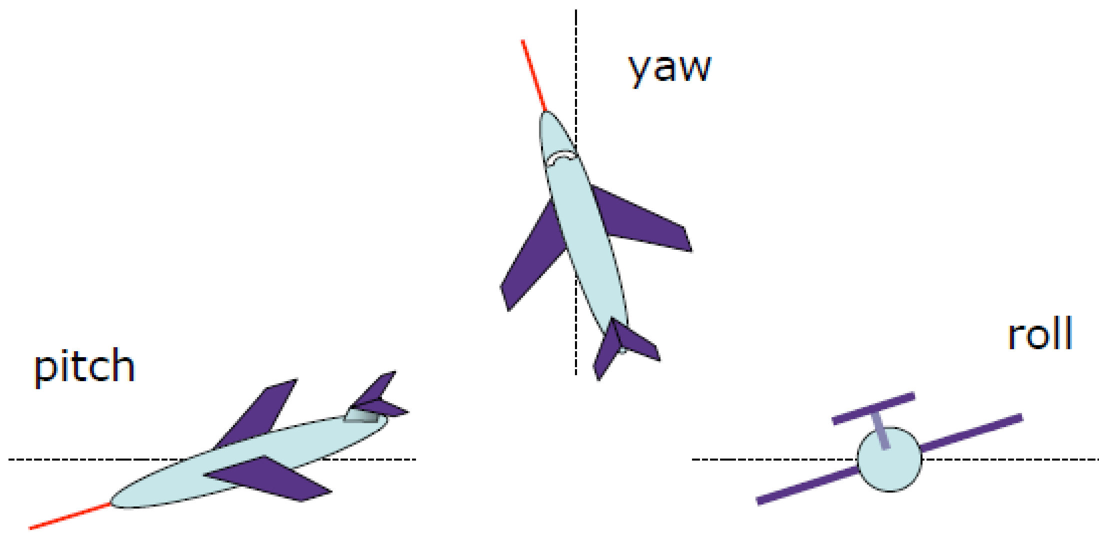

Figure 10: Pitch, yaw and roll.

2.4.2 3D Displays

3D displays are different too with respect to the previous ones, in fact we have:

- Desktop VR: ordinary screen, mouse or keyboard control but perspective motion give 3D effect
- Seeing in 3D: through helmets or special glasses etc.

Note that sometimes these devices can bring to sickness (like head spinning) so technology is trying to improve in this sense.

2.5 Physical Controls, Sensors, Etc.

2.5.1 Dedicated Displays

In dedicated displays, representations can be analog or digital. Head-up displays instead can be found in aircraft cockpits and show most important controls depending on context.

2.5.2 Sound

Beeps, bongs, clonks, whistles and whirrs used for error indicatons or confirmation of actions.

2.5.3 Touch, Feel, Smell

They are used by humans to interact with AR/VR systems, like in games and simulation. Touch and feeling are important, while smell and taste experiment very limited technology currently.

2.5.4 Environmental and Bio-Sensing

This part include all the sensors around us (car courtesy light, ultrasound detectors...) and even our own bodies (body temperature, heart rate...).

2.6 Paper, Printing and Scanning

2.6.1 Printing

Printing is the action of representing something on paper or physical entity, from a digital one. Images made of small dots are the most common ones and their critical features are resolution, speed and cost. Dot-based printers are the most common and are dot-matrix printers (line of pins that can strike the ribbon, dotting the paper), ink-jet and bubble-jet printers (tiny blobs of ink sent from print head to paper), laser printers (like photocopier, where dots of electrostatic charge are deposited on drum which picks up toner rolled onto paper which is then fixed with heat. Thermal printers instead are more common in workplace.
Every letter we see printed or also on the screen has a font which can change the readability of text through its features (size, type,...).

2.6.2 Scanners

Scanners are devices which take paper and convert it into a file (bitmap) so it is the contrary of printing. They are used in desktop publishing for incorporating photographs and other images, document storage and retrieval systems, special scanners for slides and photographic negatives. Note that paper can also be used as input for certain programs, thus we call it paper-based interaction.

2.7 Memory

2.7.1 Short-term Memory (RAM)

Random Access Memory is on silicon chips and is accessed and used at runtime, then erased when turning off the PC, so it is volatile. Some non volatile RAMs are used to store basic set-up information. They can be found on typical desktop computers.

2.7.2 Long-term Memory (Disks)

Magnetic disks of every kind and optical ones. They can also be found in nowadays PCs and computer systems.

2.7.3 Flash Memory

Silicon based but persistent, usually for USB devices.

There are some relevant features related to memory, like methods of access, format, size, compression which can affect the speed and capacity of the memory.

2.8 Processing and networks

Designers tend to assume fast processors and make interfaces more and more complicated. However, that's not the right thing to do as there are a lot of old processors still running in production use nowadays. There are also problems if system is too fast, like excessive text scroll speed. On interactive performance, we always have to assume some limitations like in computation, storage channel, graphics and network. In this way we will be able to design an UI correctly.
Networked computing is instead everything which allows access to large memory and processing, other people (email, chats...) and shared resources. Anyway, it has some issues too, like delays, conflicts and unpredictability. Internet is a great example of such a thing.

3 HCI Foundation 3 - Interaction

What is the so-called interaction? *Interaction* is the communication between user and system, basically. It is quite simple, but it contains several concepts and components that we will analyze in this section.

3.1 Models of Interaction

There are several models that describe interaction modes and situations, but now we will analyze the two most relevant ones.

- *Donald Norman's Model*: it's a model which consists of seven stages

  1. user establishes the goal
  2. formulates intention
  3. specifies actions at interface
  4. executes action
  5. perceives system state
  6. interprets system state
  7. evaluates system state with respect to goal

  This model concentrates on user's view of the interface. A fast-and-easy way to see such a scheme is the execution/evaluation loop:

  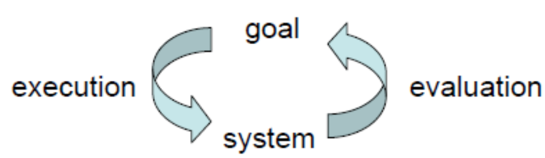

  Figure 11: Execution(2-4)/evaluation(5-7) loop.

  Remember that some systems are harder to use than others, i.e. user's formulation of actions is different from actions allowed by the system, and user's expectation of changed system state is different from actual presentation of this state.
  For fixing human errors, we improve interface design to avoid slips, and improve system understandability for avoiding mistakes.

- *Abowd and Beale framework*: an extension of Norman, composed by four parts:

  - user
  - input
  - system
  - output

Each part has its own unique language and the interaction consists of translation between languages.

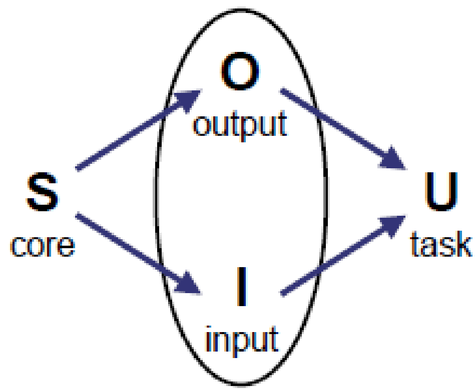

Figure 12: Abowd and Beale framework scheme.

In this model, user intentions are translated into actions at the interface, translated into alterations of system state, reflected in the output display, interpreted by the user.

3.2 Ergonomics

Ergonomics is the study of the physical characteristics of interaction; it is good at defining standards and guidelines for constraining the way we design certain aspects of systems. Some examples of ergonomics are: use of color, surrounding environment, health issues,...
Industrial interfaces are different from office ones, so we have to think about which environment we are interested in. Glass interfaces are used nowadays in industrial contexts: these interfaces are cheaper and more flexible, also because they contain multiple representation of same info.

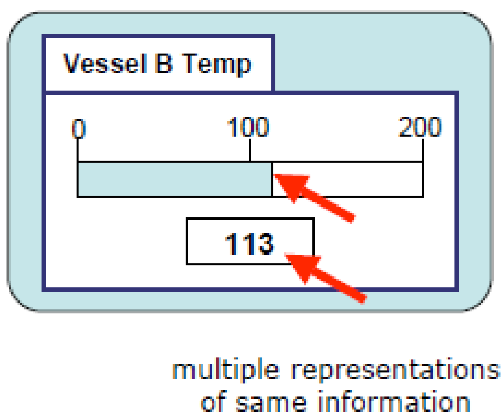

Figure 13: Glass interface example.

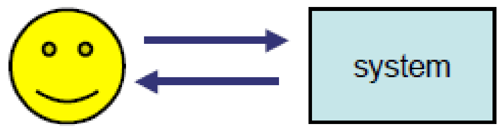

Figure 14: Direct manipulation: user interacts with artificial world.

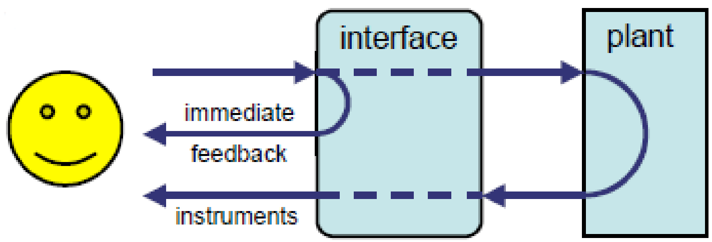

Figure 15: Indirect manipulation: user interacts with real world through interface.

3.3 Interaction Styles

There are several interaction styles such as command line interface, menus, natural language and so on. Now we describe them in detail.

- *Command line* is a way of expressing instructions to the computer directly, it is suitable for repetitive tasks but better for expert users than novices. It offers direct access to funcionality and it is usually more powerful than a standard UI. Commands and abbreviations should be meaningful for its use.

- *Menus* are a set of options displayed on the screen, that can be selected. Visible options are selectable by numbers, letters, arrow keys, mouse and so on. Often, options are hierarchicallly grouped.

- *Natural language* is familiar to user, so we mean speech recognition or typed natural language. Problems are that natural language can be vague, ambiguous and solutions could be trying to understand a subset, picking key words.

- *Query interfaces* are question/answer interfaces suitable for novice users but with limited functionality. Query languages are used, such as SQL when dealing with databases.

- *Form-fills* are primarily for data entry or data retrieval and screen looks like paper form. There are several labels and text entry fields.

- *Spreadsheets* are a sophisticated variation of form-filling, like grid of cells that contain a value or formula, etc.

- *WIMP interfaces* are interfaces composed by Windows, Icons, Menus and Pointers, that's why they are so called. M and P can also be Mice and Pull-down menus. This is the default style for majority of interactive computer systems, especially PCs and desktop machines.
- *Point and Click interfaces* instead, are used in multimedia, web browsers and hypertext. We just have to click something to get it work. Typing is minimal.
- *3D interfaces* are the ones used in VR but also the UIs where light and occlusion give depth, simply.

3.4 Context

Context affects interaction and includes other people, motivation and inadequate systems. These are its components.

3.5 Experience, Engagement and Fun

When we design a UI, we have to think about User eXperience, UX. Let's see a small difference between a real cracker and a virtual one, just to understand a simple experience.

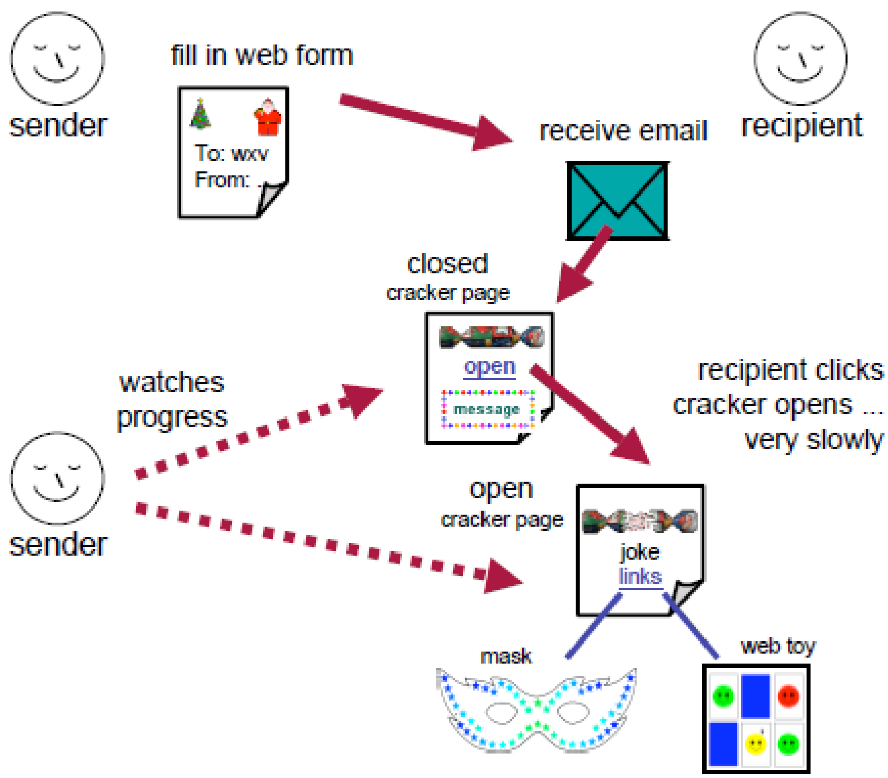

Figure 16: How crackers work.

	real cracker	virtual cracker
Surface elements		
design	cheap and cheerful	simple page/graphics
play	plastic toy and joke	web toy and joke
dressing up	paper hat	mask to cut out
Experienced effects		
shared	offered to another	sent by email message
co-experience	pulled together	sender can't see content until opened by recipient
excitement	cultural connotations	recruited expectation
hiddenness	contents inside	first page - no contents
suspense	pulling cracker	slow ... page change
surprise	bang (when it works)	WAV file (when it works)

Figure 17: Crackers experience.

In physical design, there are many constraints (ergonomic, physical, context related, aesthetic...) and they are often in contrast, so we have to find a tradeoff (let's think about the contrast between ergonomics and physical aspect!).
When designing a UI we have to make sure that physical aspects reflect logical ones (e.g. on/off buttons) and put inverse actions, like open call/close call on old phones keyboard and on washing machines buttons.
General lesson: if you want someone to do something (and to improve your product's UX), make it easy for them and understand their values!