Informatica per le Digital Humanities
Lecture 7
March 31 2025
Texts

Slides 2-7
In the previous lectures we have highlighted the differences and analogies between a human memory and a memory inside a computer. We have seen the extreme case of warnings against nuclear waste contamination that need a system that, due to the very long temporal horizon of the waste's radioactivity, trascends the limited boundaries of human communication languages. Let us focus on the present and on the computing devices that make use of a memory, containing operands (data), operations (instructions), and addresses.

Slides 8-12
If we fully embrace the memory metaphor and anthropomorphize a computer, we can say that it remembers data, it remembers what to do with them, and it remembers where both data and instructions are. If remembering where we come from and our cultural heritage is what makes us human, then digital computer memories can be seen as a very simple yet concrete example of digital humanities.

Slides 13-19
A traditional way for us to remember about our past and our culture is to conserve books in a library. Digitizing libraries is a typical endeavor in the digital humanities. A small, related example is given by any text that you might elaborate with a word processing software in your laptop. Such a software offers you the functionality of "find and replace", which can be described in terms of data, instructions, and addresses.

Slides 20-35
However, here we may need to add another color to the color-coding system. We had blue for data, red for instructions, green for addresses, and now we also have yellow for repetition, because the find and replace actions need to be repeated throughout the file we are elaborating. Repetition relies on memory because to repeat something we need to remember that something. Moreover, repetition also needs a reference and retrieval, because we need to know where the instruction to be repeated is stored.

Slides 36-39
Repetition always comes with a risk: if instruction 4 sends us back to instruction 2, we may end up being stuck in an endless loop (2, 3, 4, 2, 3, 4…). Every instruction that prescribes a repetition must include a condition that breaks that loop. Indeed, in this case, we have "until the end of the file". How does a computer know when it has reached the end of a file it is scanning? It relies on the references provided by the file system, which has references to all the addresses of the files.

Slides 40-48
If memory devices make repetition (also known as iteration) possible, iteration, in turn, enables automation. The basic idea is that if instructions are specified in a way that their execution can be delegated to a machine, then the repeated execution will entail that the machine operates on its own for a significant amount of time. This way of working predates digital technologies: it has been a fundamental component of the industrial revolution and the relevant growing role of machines inside factories. This is also the idea behind the desire of

delegating boring and repetitive tasks to machines, which are also able to operate at much higher speed. The “find and replace” functionality in a word processing software, for instance, is a great tool when it comes to scanning documents of notable size (like an entire library).

Slides 49-58
Despite the apparent disappearance of humans in highly automated environments, they are still playing a fundamental role in building such environments in the first place. Even in the simple set of instructions in the “find and replace” task, it is humans who choose what to find, in which file, with what to replace it, etc. The instructions themselves have been conceived and written by humans. The only phase in which humans are absent is when the instructions are executed by a computer. However, in the moment results are obtained humans come back into the picture, otherwise results without anyone checking them out are like a tree falling in a deserted island: soundwaves are created but there is no sound; similarly, results may be shown on a screen, but there is nobody who can take advantage of them.

Slides 59-69
Results from a computer acquire meaning only if there is a person interpreting them. The issue of meaning and interpretation is another phenomenon that predates computer technology and characterizes human culture on a very fundamental level. Communication between two individuals work only when the words that they exchange are understood by both by means of a common background context. Even in such case, there are differences in what entertained in their minds: the word “cat” is the same for everybody involved in the communication, but it is possible, if not very likely, that each individual will entertain a different mental image of a cat. Computers can support communication between humans by digitizing and transmitting texts, images, and sounds, enabling telecommunication. However, computers and other digital devices only work as carriers of those items, whereas the relevant interpretation and generation of meaning still happens in the minds of the human beings who are communicating.

Slides 70-82
Regarding the question whether computers can entertain meaning on their own, one significant negative answer was provided in 1980 by American philosopher John Searle, by means of the thought experiment of the “Chinese Room”.

Slides 83-87
Anyone can experience being in a Chinese Room trying to interpret signs that they don’t know.

Slides 88-92
Philosopher Hubert Dreyfus also sheds light on a difference between communication between people and communication simulated by a computer: in order for the latter to cover all aspects needed in the former, there should be an immense, if not unsurmountable quantity of data that reflect what a human being acquires throughout their life, that is, experience and cultural contexts. With an example of a simple dialog Dreyfus shows how much implicit knowledge is needed for those words to make sense, not only in the literal term of what a word like “gift” means, but also in terms of the actions, the habits, and the agreements among people that are surrounding that concept. This additional and yet fundamental information is known as “common sense”.

Slides 93-102
What these discourses point at is a distinction between signs that can be written, spoken, encoded, shown on a screen, transmitted though cables, etc. and their meaning. It is the well-known divide between syntax and semantics. Ultimately, what philosophers like Searle and Dreyfus tell us is that computers can deal with syntax but not with semantics.

Slides 103-108
Even a simple Web search shows us that a computer is very quick and efficient at providing us with results based on a syntactic search (e.g. all the pages containing the sequence of characters "c-h-i-p"), but it is up to the human being to select the semantically relevant result.

Slides 109-126
The Web is an aspect of the digital world that deserves some further consideration around the relation between syntax and semantics. Data on the Web can be shown in different ways (e.g. in the form of a link, which is what turns texts into hypertexts, or in bold, underlined, etc.) thanks to a markup language that with metadata (i.e. data about data) gives your browser (e.g. Safari, Chrome, Edge…) instructions on how to display data. Metadata are organized in the form of tags that surround data and provide information about it (in the case of the HTML language, about how a browser is supposed to diplay it). Tim Berners-Lee, who first proposed the idea of hyperlinks and the Web in the early 1990s, has more recently suggested the use of HTML tags to add some more information about the tagged data, in an attempt to enable computers to use syntax to provide some indication on the semantics of the words on the Web.