Graphs: an introduction

1 Introduction

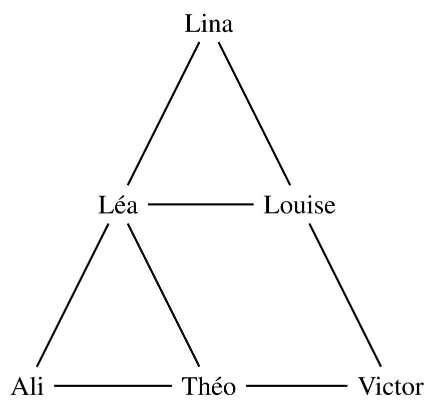
From our point of view, a **graph** is mainly a formalism enabling the representation of **relations** linking various **entities**. These relations might be of different types: social relations, geographical proximity, common points, etc. The advantage of using such a representation is that it allows to forget the semantics associated to each relation (proximity, etc.) in order to only keep the **structure** formed by these relations.
A graph is thus a set of entities called **nodes** or **vertices** (one vertex, several vertices), and a set of **links** between them: **edges** (non directed links) and **arcs** (directed links). And we associate entities to nodes/vertices and relations to edges/arcs.

1.1 A first example

Let us consider a group of persons: Léa, Louise, Lina, Théo, Ali, and Victor and some relations between them:

- Léa, Louise and Lina have a first name beginning with an ”L”
- Théo and Victor are brothers.
- Ali, Léa and Théo live in the same district.
- Victor and Louise have a phone number that ends with number 42.

From this set of information we can represent the relations linking the members of this group, and this may be represented as a **graph**. For obtaining it, we create a vertex for each person and we add an edge between two nodes as soon as they are linked by some relation. This gives the following graph:



1.2 Terminology

In the previous graph, first name are associated to **vertices** or **nodes** (no difference between these two terms), and the relations are associated to **edges**. More formally, we denote $G = (V, E)$ the graph G, with V the set of vertices and E the set of edges.

In this graph, every edge $e \in E$ has two **extremities**, two vertices, thus, each edge can be represented as a **pair of nodes**, $e = \{v_i, v_j\} = \{v_j, v_i\}$

Vertices v_i and v_j are said **adjacent** or just **neighbors**.

In the same context, given an edge $e = \{v_i, v_j\}$, we say that e is incident to v_i (and also to v_j).

The notion of **neighborhood** is central for algorithms, and is even key for building algorithm for walking through the graph starting from one initial vertex without knowing all the vertices at start. The idea consists is walking from one vertex to another one, step by step, according to its neighborhood (we will see that later).

The number of edges incident to a vertex v is called the **degree** of v. In our example graph, the degree of Léa is 4 while it is 2 for Victor.

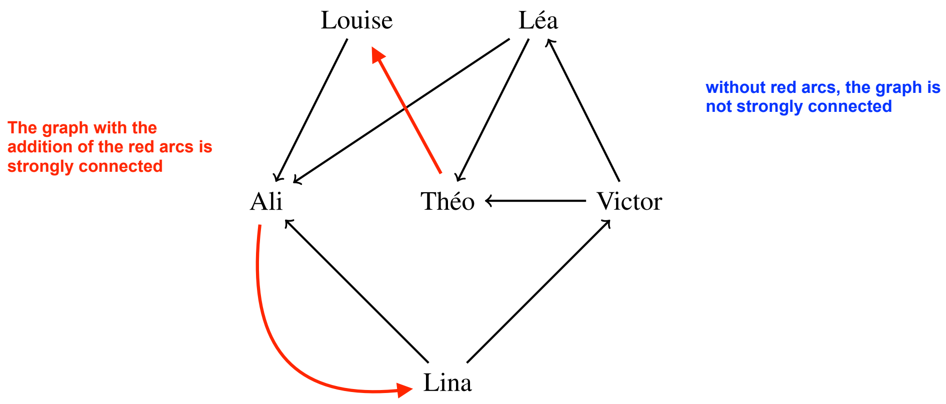
The **order** of a graph is its number of vertices, always for our example graph, its order is 6.

1.3 Another Example

Let us consider the same group but focusing this time on communications between them (just another set of relations). Each group member has an email address and we suppose that during the last 24 hours they exchange some email in the following way:

- Léa sent an email to Théo and Ali
- Lina sent an email to Victor and Ali
- Louise sent an email to Ali
- Victor sent an email to Léa and Théo

We can again represent these relations/communications/exchange using the graph formalism, a little bit different than the previous example, because relations are, this time, **directed** from one person to another one. In the graph this is represented by the use of arcs instead of edges.



1.4 More terminology

In this graph, first names still correspond to **vertices** (or nodes), and the relations are called **arcs**. As for the non-directed case, in a graph $G = (V, A)$ each arc $a \in A$ owns two extremities, but their are ordered this time, so, an arc is a couple of vertices $a = (v_i, v_j) \neq (v_j, v_i)$

Given an arc $a = (v_i, v_j)$, v_i is sometimes called the **source node** and v_j the **destination node**.

As for the non-directed case, vertices v_i and v_j are said **adjacent** or just **neighbors**, and the arc is said incident to vertex v_i and v_j.

The orientation of links enables us to refine the notion of degree that can be decomposed into **in-degree** (number of arcs for which the vertex is the destination) and **out-degree**. The degree of a vertex in a directed graph is equal to the sum of its in-degree and out-degree.

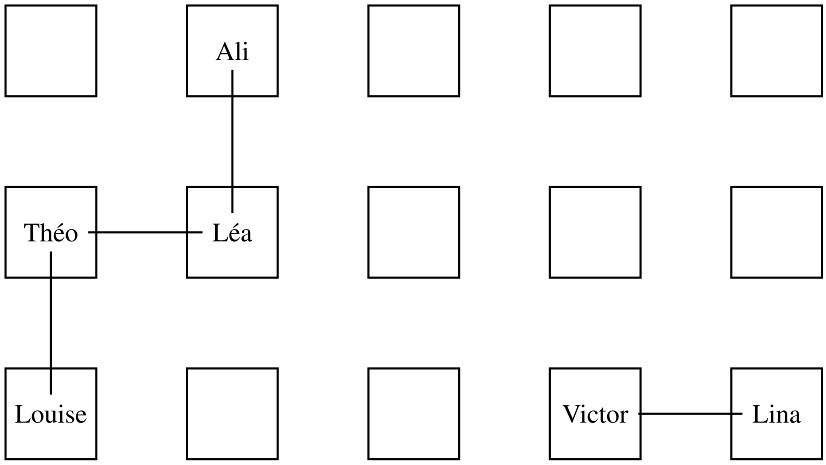
In our example, vertex Léa has an out-degree equal to 2 and an in-degree equal to 1, and thus its degree is 3.

1.5 Third Example

We keep our group and now look at their position in the classroom.

- Léa is next toThéo who is in front of Louise.
- Ali is in front of Léa
- On the same row as Louise but two columns away are Victor and Lina.
- Victor is next to Lina.

We consider that two of them are neighbors if they are direct neighbors either on a row or on a column of the classroom.



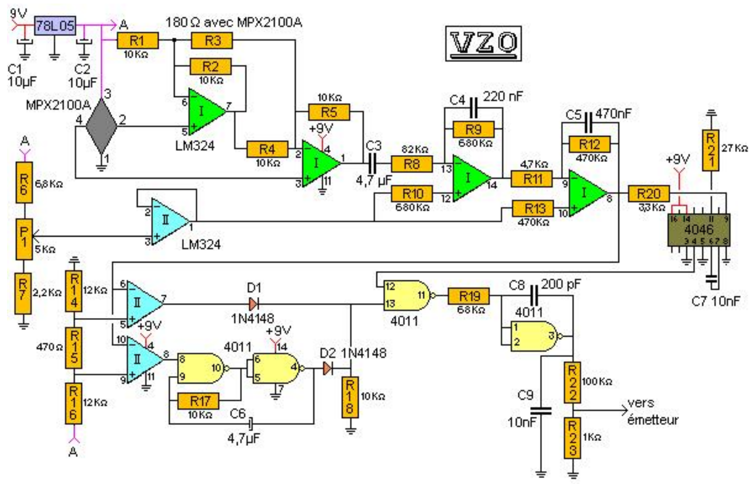
1.6 Terminology again

This is also a graph, but this time it presents a particular property, it is **not connected**. It is composed of several parts that are not linked. In other words, starting from one vertex and following edges, some vertices cannot be reached. On the contrary, in a **non directed connected graph**, starting from any vertex it is always possible to reach any other vertex of the graph by passing through a set of edges. We will come back to this notion latex but just keep in mind that we refer to **graph connectivity** to deal with this property.

2 Graphs as models of the World

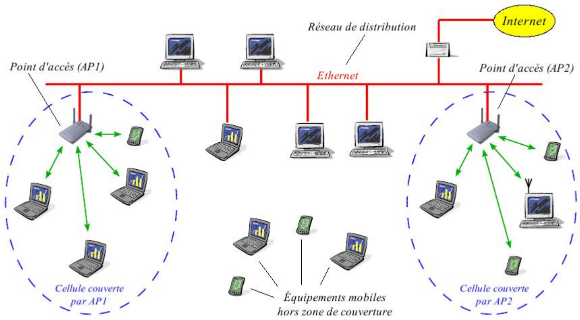
As soon as you are looking to your surrounding world as entities linked by relations (whatever the semantic), you can describe the world as graphs. Here are some examples of that.

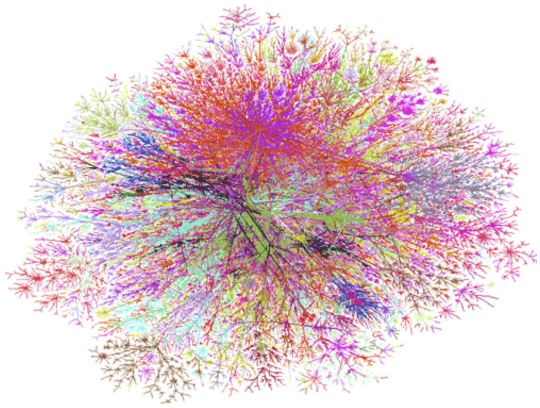
2.1 Electronic



Vertices are electronic components and edges are printed circuit board tracks

2.2 Computer Science Networks

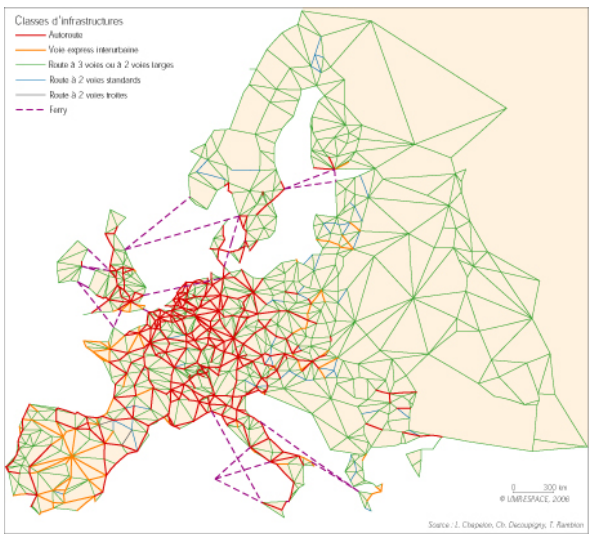


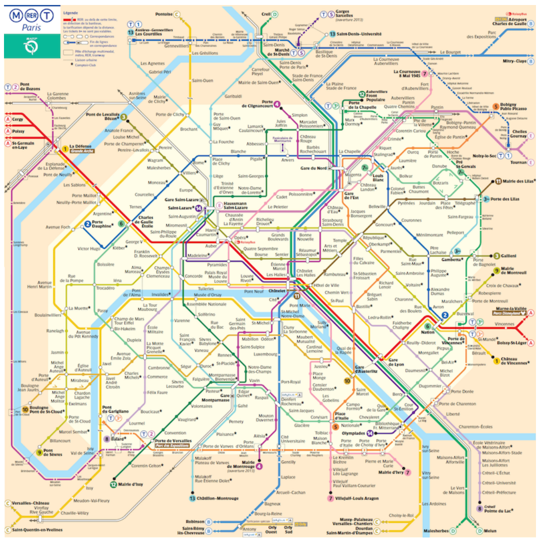
‘


- Left image: vertices are computers, routers and wireless networks access points and edges are communication links between these elements. You can notice that edges can be of different types if the network is hybrid (wired and wireless).
- Right image: image of the Internet Network. Nodes are web sites and each edge is an hypertext link between two sites.

  Remark: this network is directed, since your own web site may have a link to wikipedia but it is unlikely that wikipedia had a link to your own web site.

2.3 Geography

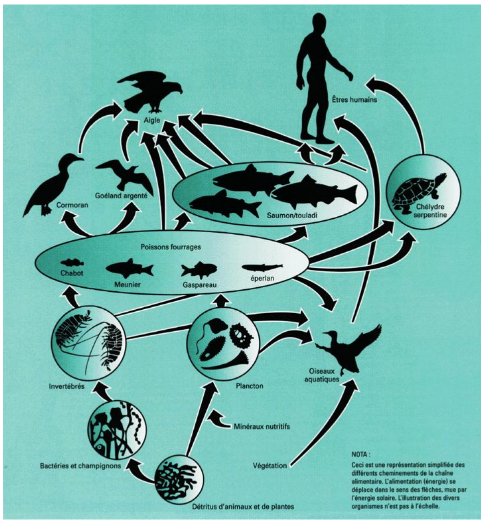


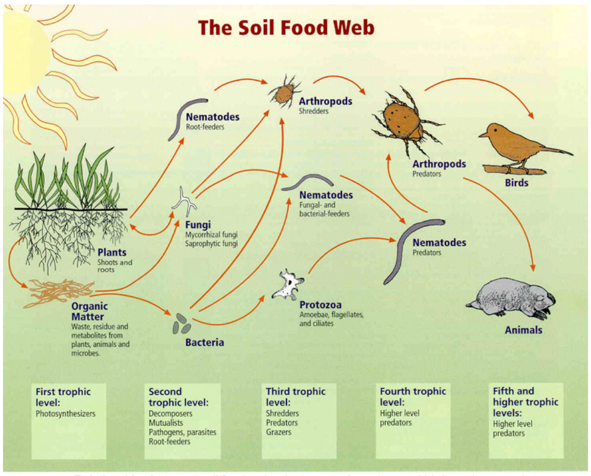
‘


- Left image: at the continent level, vertices are big cities and main roads junctions. Edges are highways.
- Right image: vertices are underground stations and edges are railways/tunnels linking them.

<u>Remark:</u> for a given domain, depending on the scale, vertices and edges might be different.

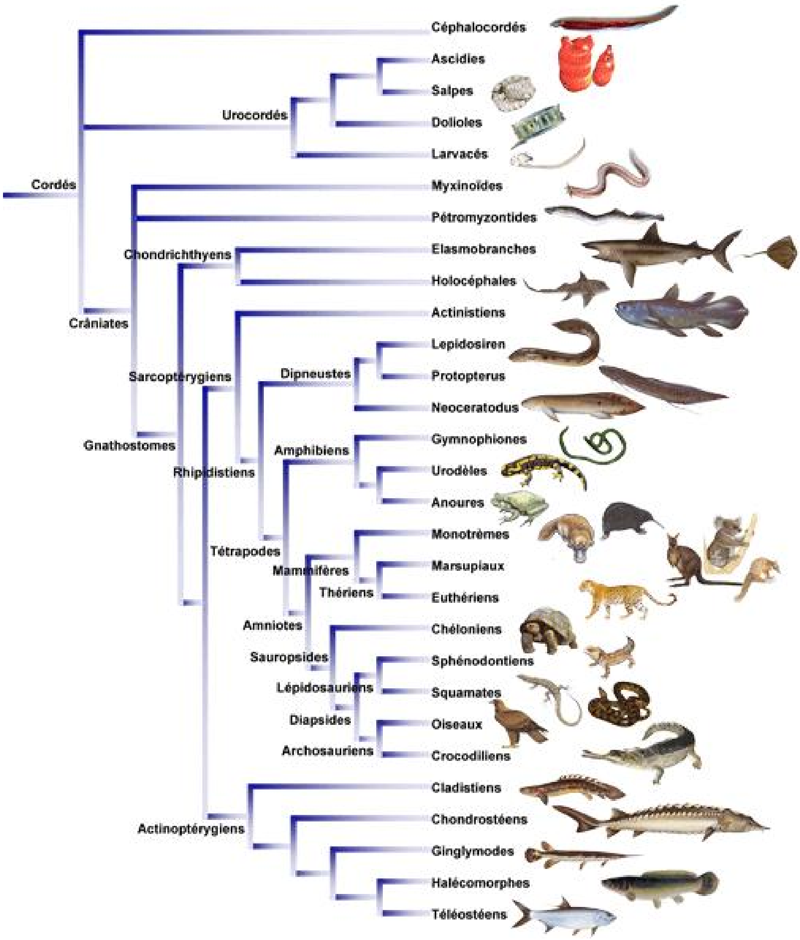
2.4 Biology and Ecology



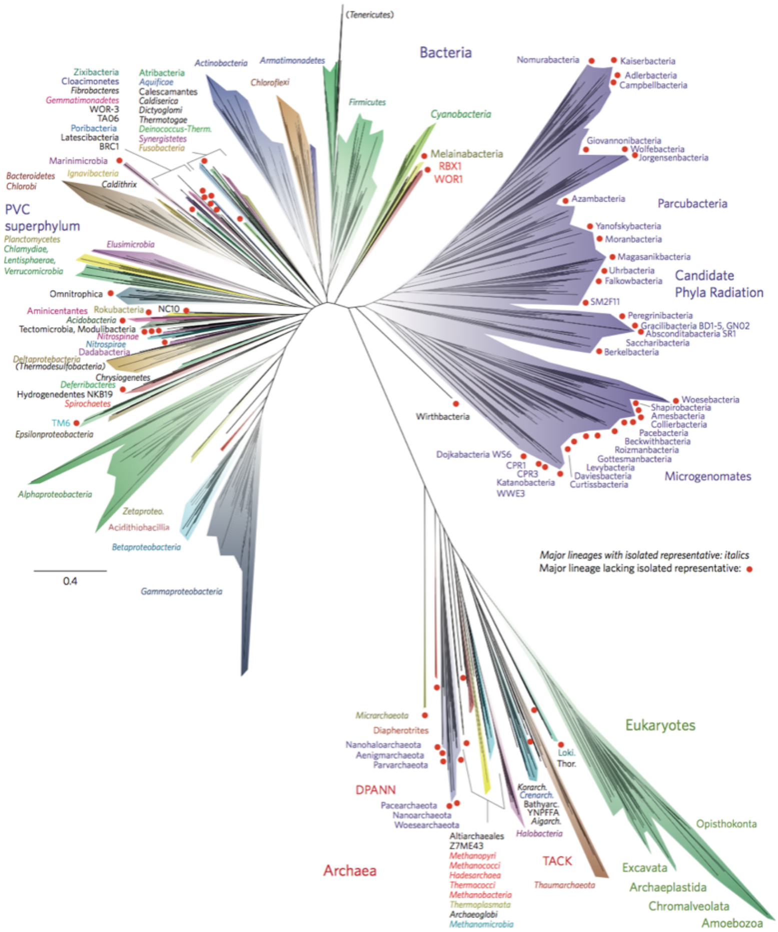


Relationships between soil food web, plants, organic matter, and birds and mammals
Image courtesy of USDA Natural Resources Conservation Service
http://soils.usda.gov/sqi/soil_quality/soil_biology/soil_food_web.html.

Vertices are organisms and predator-prey relations are represented by arcs. This is a food web, a network based on different trophic levels.



Biologists have a tendency to classify living organisms, this is called taxonomy. When a new species is discovered, they try to insert it within their very large book. For that, they compare species with each other, and note their similarities and differences (physical differences, genomes, etc.). From these comparison they obtain specific graph structures called **trees**. The one represented below is called a dendogram and represents taxonomic relationships. In these graphs, currently living species are the **leaves** (animals), and internal nodes correspond to common hypothetical ancestor.

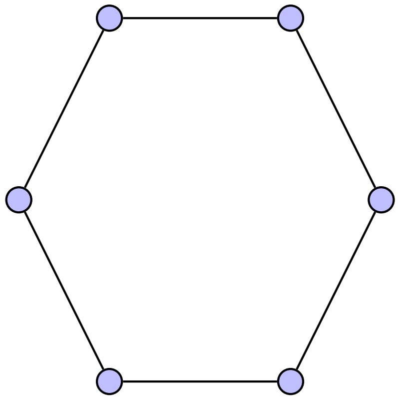


3 Some Particular Graphs

Some graphs present particular properties and are gathered into categories. One way of classifying them relies on their shape, the structure of their links, we call that the **topology** of the graph. Here are some well-known graph topologies.

3.1 Ring

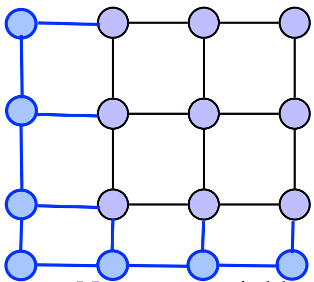
definition : a **ring** is a connected graph in which all vertices have a degree equal to 2.



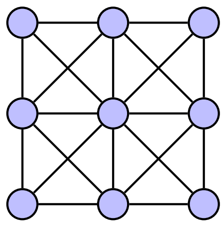
3.2 Grid

very often used for modelling the environment —> mesh
weather forecast models use this type of graphs

The graphs below are called **grids**. The grid on the left is said to present a **von Neumann neighborhood** and the right one a **Moore neighborhood**. These two grids are of size 3×3 and their order is 9.



von Neumann neighborhood



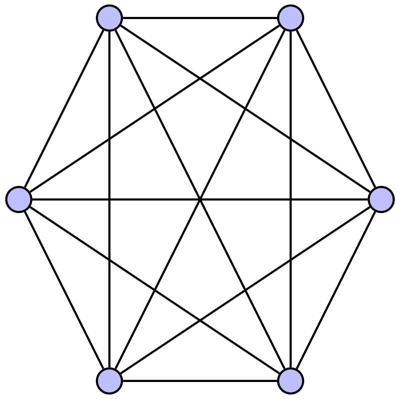
Moore neighborhood

Cellular automata —> Game of Life (John Conway)

3.3 Full-connected Graphs

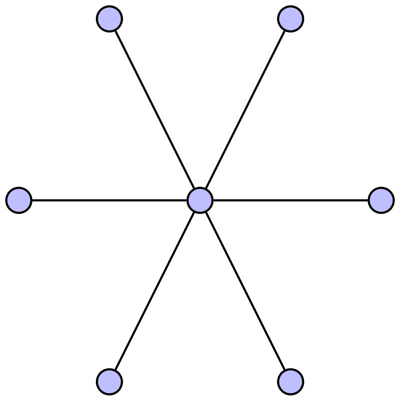
Definition: a **full-connected graph** is a connected graph in which each vertex is connected to every other vertex.

More formally:
Given $G = (V, E)$ a connected graph, if $\forall v_i, v_j \in V \times V, \{v_i, v_j\} \in E$ then G is **full-connected**.



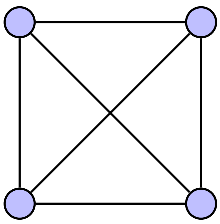
3.4 Star Graph

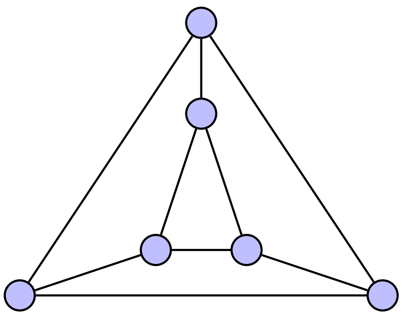
Definition: a **star graph** of order n is a connected graph in which $n-1$ vertices have a degree equal to 1 and one vertex has a degree equal to $n-1$.



3.5 Regular Graphs

Definition: a **regular graph** is a connected graph in which all the vertices have the same degree.

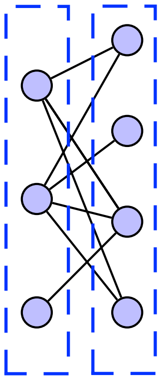




3.6 Bipartite Graphs

Definition: a **bipartite graph** is a connected graph which set of vertices is divided in two subsets such that each edge has one extremity in each subset.

More formally:
Given $G = (V, E)$ a connected graph. If $\exists V_1, V_2$ such that: $V_1 \bigcup V_2 = V$ and $\forall e = \{v_i, v_j\} \in E$ $v_i \in V_1$ and $v_j \in V_2$, then the graph is bipartite.



3.7 Trees

Trees are ubiquitous graphs in computer science. They can be defined in several ways, here is one:

Definition: a **tree** is a connected graph in which the number of edges is equal to the number of vertices minus 1.