The basics of graphs
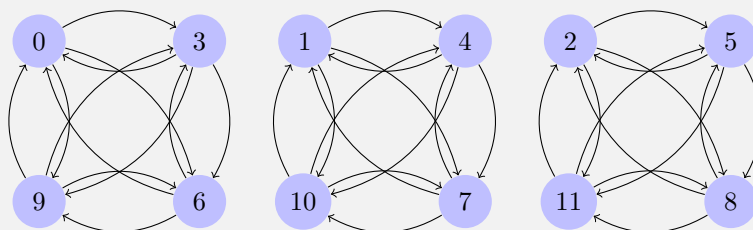
A graph consists of at least two components. The first is the set of nodes. When talking about graphs, every node is usually called a **vertex**, so instead of the graph's set of nodes we talk about the graph's set V of **vertices**. In addition, there is an edge relation $E \subseteq V \times V$ that connects vertices to each other. When drawing graphs, the edges in E are represented by arrows.
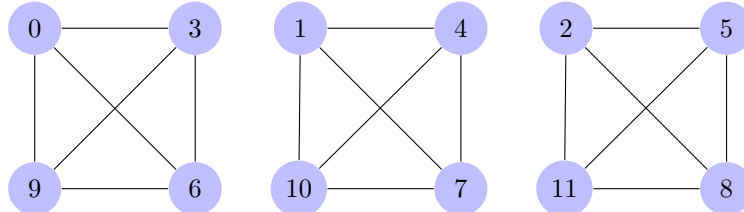
EXAMPLE 1.

The graph depicted below has $\{0, 1, 2, \dots, 11\}$ as its set of vertices, and E consists of the following tuples:

- $\langle 0, 3 \rangle$
- $\langle 0, 6 \rangle$
- $\langle 0, 9 \rangle$
- $\langle 3, 0 \rangle$
- $\langle 3, 6 \rangle$
- $\langle 3, 9 \rangle$
- $\langle 6, 0 \rangle$
- $\langle 6, 3 \rangle$
- $\langle 6, 9 \rangle$
- $\langle 9, 0 \rangle$
- $\langle 9, 3 \rangle$
- $\langle 9, 6 \rangle$
- $\langle 1, 4 \rangle$
- $\langle 1, 7 \rangle$
- $\langle 1, 10 \rangle$
- $\langle 4, 1 \rangle$
- $\langle 4, 7 \rangle$
- $\langle 4, 10 \rangle$
- and so on

By default, we take the edges of a graph to be *directed*: aEb does not imply bEa. This is why we depicted edges with arrows. You may think of them as a one-way roads.

In the special case where E is symmetric, we call the graph **undirected**. It is customary to draw undirected graphs with undirected edges rather than separate arrows in both directions. Since the edge relation in our example graph is symmetric, we can draw it in a much simpler fashion using undirected edges.

If the edge relation is transitive, it is also customary not to draw any edges that can be inferred by transitivity. This convention further simplifies how we depict our example graph.
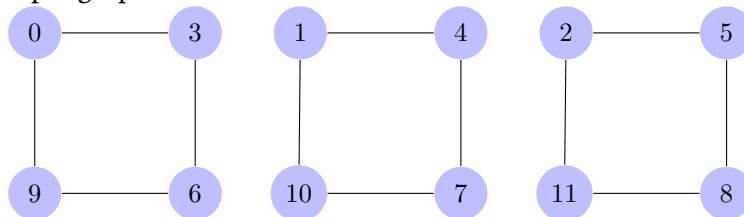
But keep in mind that these three pictures all define the same graph, we are merely using various notational conventions to remove clutter from our pictures. Mathematically, how we draw graphs has nothing to do with their definition.

> **DEFINITION 1.** A *graph* is a pair $\langle V, E \rangle$ where V is a set of **vertices** and $E \subseteq V \times V$ is the **edge** relation.

Note that the definition above treats undirected graphs as the special case of directed graphs where E is symmetric. Some authors instead assume that graphs are undirected by default and use the term **digraph** to refer to a directed graph. Either one is fine, of course, but for linguistics directed graphs are usually more important than undirected ones, which is why we will continue to treat undirected graphs as a special case of directed graphs.

> **EXERCISE 1.**
> For each on the hierarchies listed below, define its corresponding graph.
>
> - the person hierarchy
> - the person-person hierarchy used for the PCC
> - the case hierarchy for noun stem allmorphy

> Each edge you posit should correspond to a line in the figure for the relevant hierarchy.

1 Paths and Reachability

One of the central notions for graphs is **reachability**. A vertex a is reachable from vertex b iff we can follow a sequence of edges from a to b. This sequence is called a **path**.

EXAMPLE 2.
Consider one more the following graph, and assume that the edge relation contains a pair $\langle a, b \rangle$ iff there is a line from a to b in the picture:
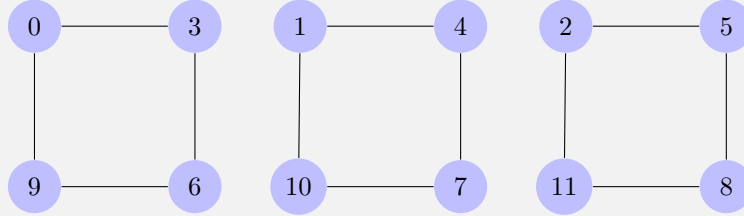
Here 3 is reachable from 0 because there is an edge taking us from 0 to 3. But 6 is also reachable from 0 as we can first move to 3 and from there 6. The path from 0 to 6 can be written as $\langle 0, 3, 6 \rangle$.

On the other hand, 1 is not reachable from 0. No matter which sequence of edges we take from 1, there is no way to ever get to 0.

DEFINITION 2. A **path** through a graph $\langle V, E \rangle$ is a sequence $\langle v_1, \ldots, v_n \rangle$ ($n \geq 2$) such that $\langle v_i, v_{i+1} \rangle \in E$ for all $1 \leq i < n$. For all $a, b \in V$, we say that b is reachable from a iff there is a path $\langle v_1, \ldots, v_n \rangle$ with $v_1 = a$ and $v_n = b$.

As always in mathematics, there are multiple ways of saying the same thing, and the same is true for reachability. A node b is reachable from a in graph $\langle V, E \rangle$ iff $\langle a, b \rangle$ is member of the transitive closure of E. The transitive closure of E is the smallest E' with $E \subseteq E' \subseteq V \times V$ such that E' is transitive.

EXAMPLE 3.
Consider the graph below.
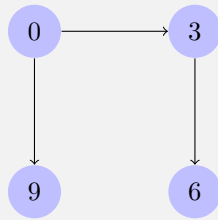
Its edge relation E consists of the pairs $\langle 0, 3\rangle$, $\langle 0, 9\rangle$, $\langle 3, 6\rangle$. The transitive closure E' of E also contains $\langle 0, 6\rangle$. This is the smallest superset of E that is transitive. The transitive closure of E' is E' itself — since E' is already transitive, nothing needs to be added. The symmetric closure E'' of E' must also contain $\langle 3, 0\rangle$, $\langle 9, 0\rangle$, and $\langle 6, 3\rangle$. Now if we take the transitive closure of E'', every node is connected to every node, so the edge relation is identical to $V \times V$. Thus the transitive, symmetric closure of E is $V \times V$ because there is no smaller extension of E that is both symmetric and transitive.