Non-euclidian data and Graphs

Raoul Grouls, 10 juni 2025

A strange new universe

What is Euclidian geometry?

Euclidian data follows the axioms of euclidian geometry

1. A straight line may be drawn between any two points.
2. Any terminated straight line may be extended indefinitely.
3. A circle may be drawn with any given point as center and any given radius.
4. All right angles are equal.
5. For any given point not on a given line, there is exactly one line through the point that does not meet the given line

What is a vectorspace?

Let V be a set, let F be a field equipped with addition and multiplication

We define binary operations “+” on V, denoted $V \times V \to V$, and “.” on $F \times V$ denoted $F \times V \to V$

What is a vectorspace?

Let V be a set, let F be a field equipped with addition and multiplication

We define binary operations “+” on V, denoted $V \times V \to V$, and “.” on $F \times V$ denoted $F \times V \to V$

A **vectorspace** satisfies $\forall c, d \in F, \forall u, v, w \in V$

Closure under addition: $u + v \in V$

Closure under multiplication: $c \cdot v \in V$

What is a vectorspace?

Let V be a set, let F be a field equipped with addition and multiplication

We define binary operations "+" on V, denoted $V \times V \rightarrow V$, and "." on $F \times V$ denoted $F \times V \rightarrow V$

A **<u>vectorspace</u>** satisfies $\forall c, d \in F, \forall u, v, w \in V$

Closure under addition: $u + v \in V$

Closure under multiplication: $c \cdot v \in V$

Addition (+):

1. Commutative: $u + v = v + u$
2. Associative: $(u + v) + w = u + (v + w)$
3. Identity: $u + 0 = 0 + u = u$
4. Inverse: There exists an element (-1)
   such that: $u + (-1)u = 0$

What is a vectorspace?

Let V be a set, let F be a field equipped with addition and multiplication

We define binary operations "+" on V, denoted $V \times V \to V$, and "." on $F \times V$ denoted $F \times V \to V$

A **vectorspace** satisfies $\forall c, d \in F, \forall u, v, w \in V$

Closure under addition: $u + v \in V$

Closure under multiplication: $c \cdot v \in V$

Addition (+):

1. Commutative: $u + v = v + u$
2. Associative: $(u + v) + w = u + (v + w)$
3. Identity: $u + 0 = 0 + u = u$
4. Inverse: There exists an element (-1) such that: $u + (-1)u = 0$

Multiplication (.):

1. Compatibility: $(cd)u = c(du)$
2. Distributivity: $c(u + v) = cu + cv$
3. Distributivity: $(c + d)u = cu + du$
4. Identity: $1 \cdot u = u$

What is a metric?

For $\forall x, y, z$:

1. Non-negativity: $d(x, y) \geq 0$
2. Identity of indiscernibles: e$d(x, y) = 0$ if and only if $x = y$.
3. Symmetry: $d(x, y) = d(y, x)$
4. Triangle inequality: d(x, y) + d(y, z) ≥ d(x, z)

What we have been doing so far

What we have been working with so far, is technically called:

- An n-dimensional real vectorspace, equipped with a metric.
- We have been denoting the vectorspace as $\mathbb{R}^d$
- The euclidian vectorspace is one specific case where we use the euclidian metric.

What we have been doing so far

- The metric has often been implicit, common choices are:
  - Euclidian distance $L_2: \; d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$
  - Manhattan distance $L_1: \; d(x, y) = \sum_{n=1}^{n} |\, x_i - y_i \,|$

What we have been doing so far

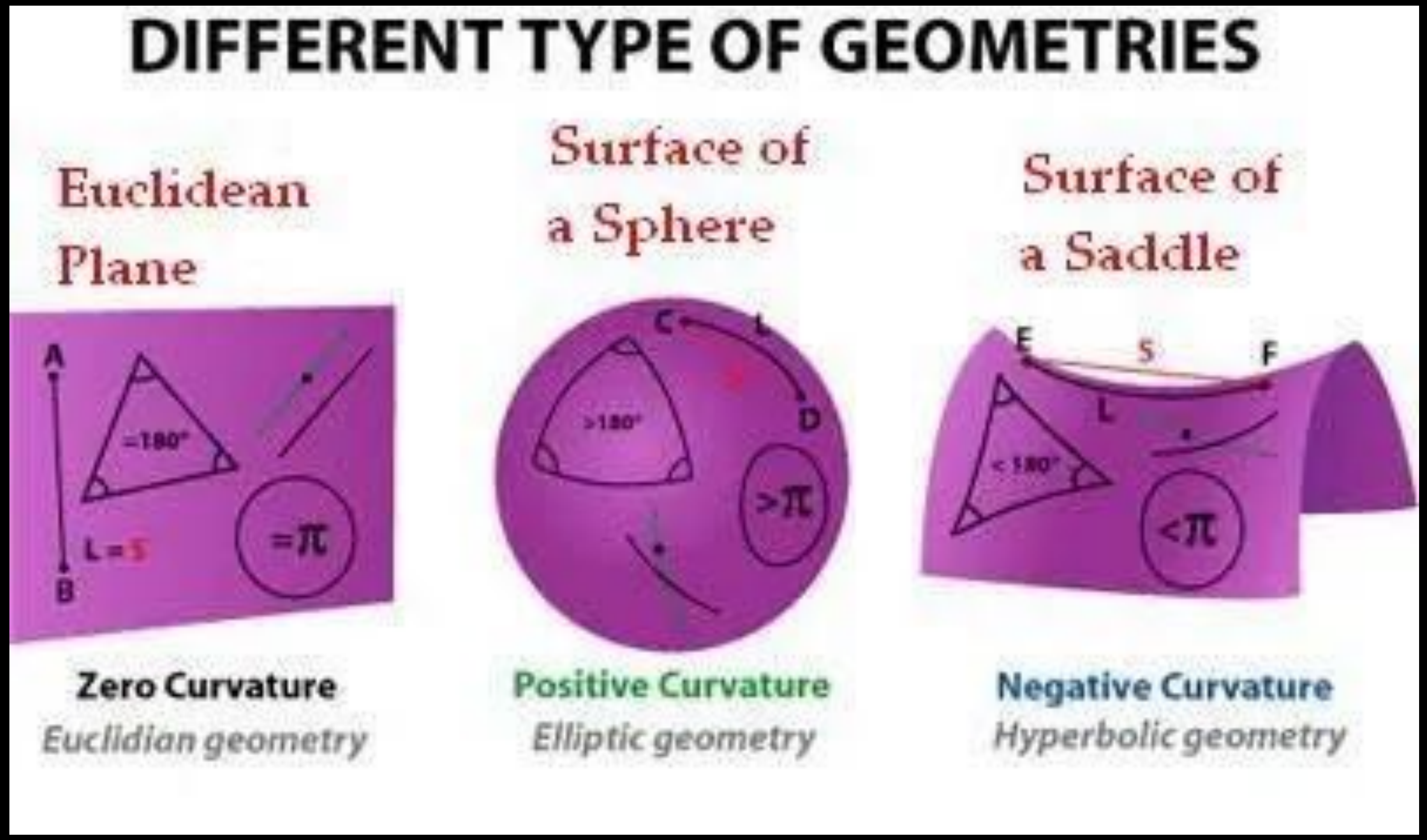
- Due to our training, we turned the vectorspace into a *semantic* vectorspace

What we have been doing so far

- Due to our training, we turned the vectorspace into a *semantic* vectorspace
- where the geometric relationships in the vectorspace correspond to semantic relationships in the original domain, like language.

What is Non-Euclidian geometry?

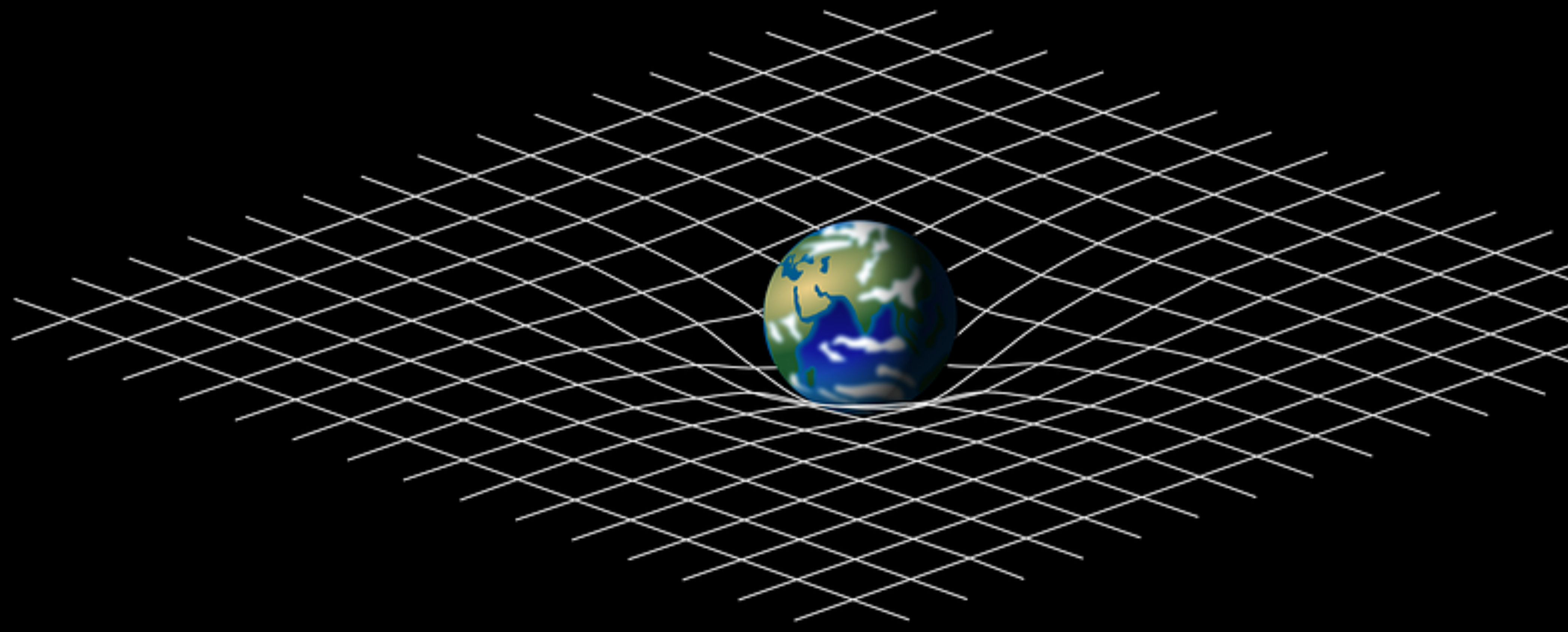
Rejection of the parallel postulate



First attempts at challenging the parallel postulate are by Ibn al-Haytham in the 11th century.

Early 19th century, the parallel postulate was rejected as “apriori true”

- 1823/1832, Bolyai’s (Hungarian) father writes in 1820 ““You must not attempt this approach to parallels. […] I have traversed this bottomless night, which extinguished all light and joy in my life”, but in 1823 Bolyai writes back “I have created a strange new universe”. It is published in 1932 by his father.
- 1829 “*A Concise Outline of the Foundations of Geometry*” by Lobachevsky (Russian)
- 1848 Bolyai learns that Lobachevsky has published a similar piece. Their work is the basis for “hyperbolic geometry”
- 1905 Poincare describes his disk model of hyperbolic space and suggests that space might be hyperbolic.
- 1915 Einstein publishes “The field equations of gravitation”, describing space as non-euclidian.

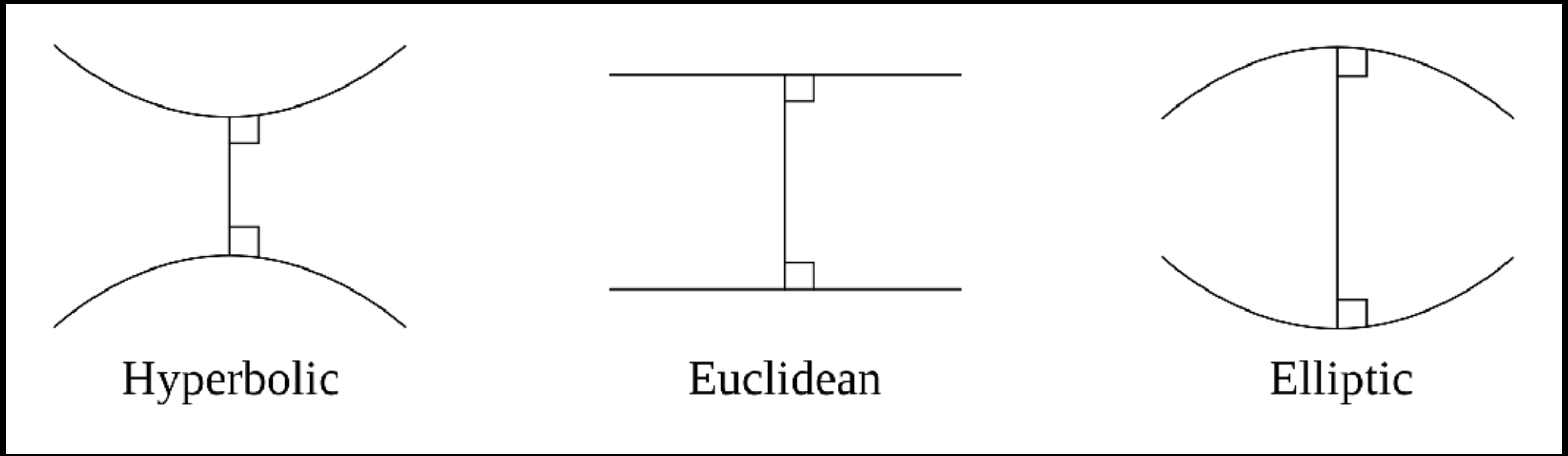


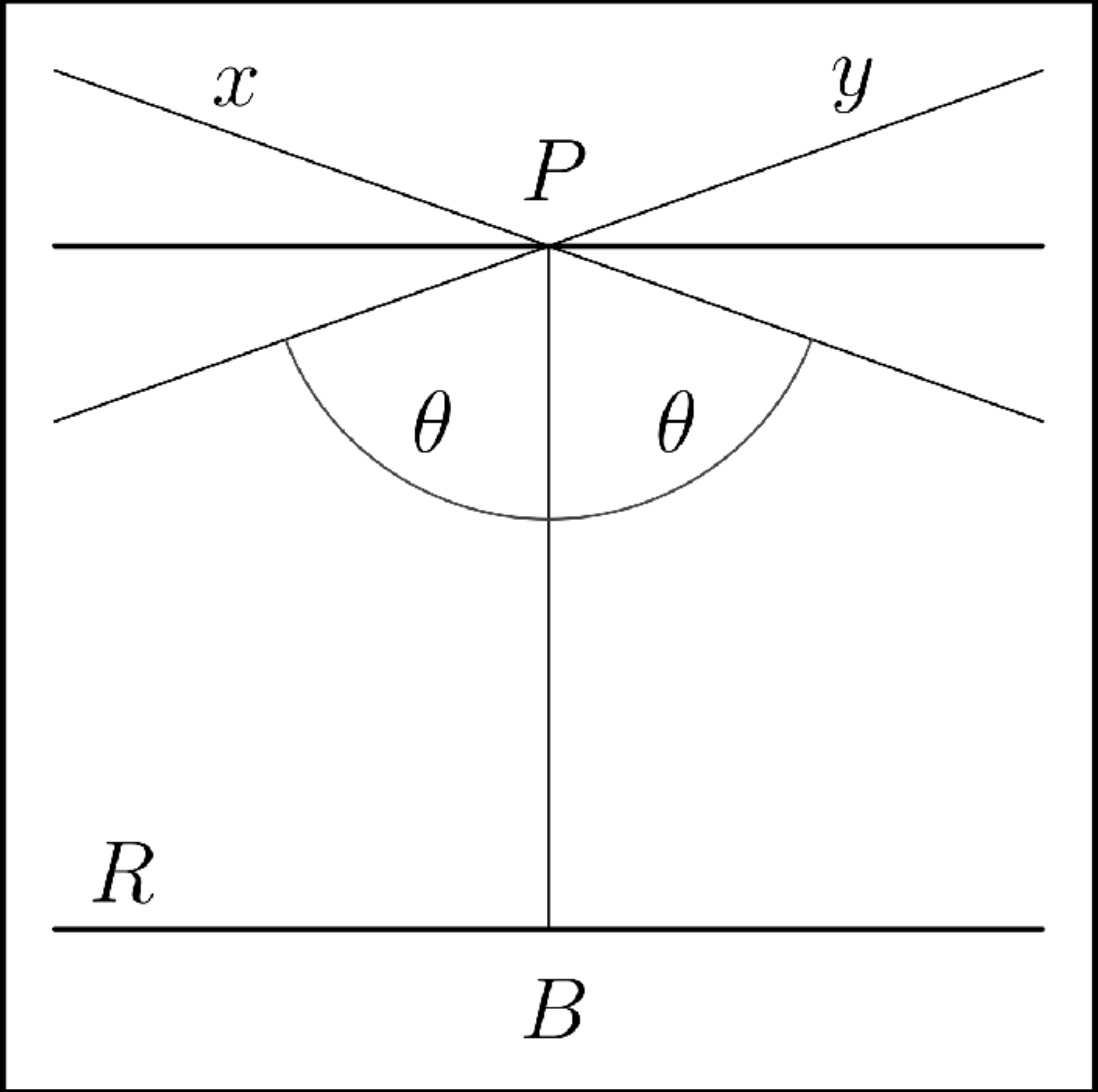
What is Non-Euclidian geometry?

Rejection of the parallel postulate

Bolyai-Lobachevskian geometry:

For any given line *R* and point *P* not on *R*, in the plane containing both line *R* and point *P* there are at least **two distinct lines** through *P* that do not intersect *R*.



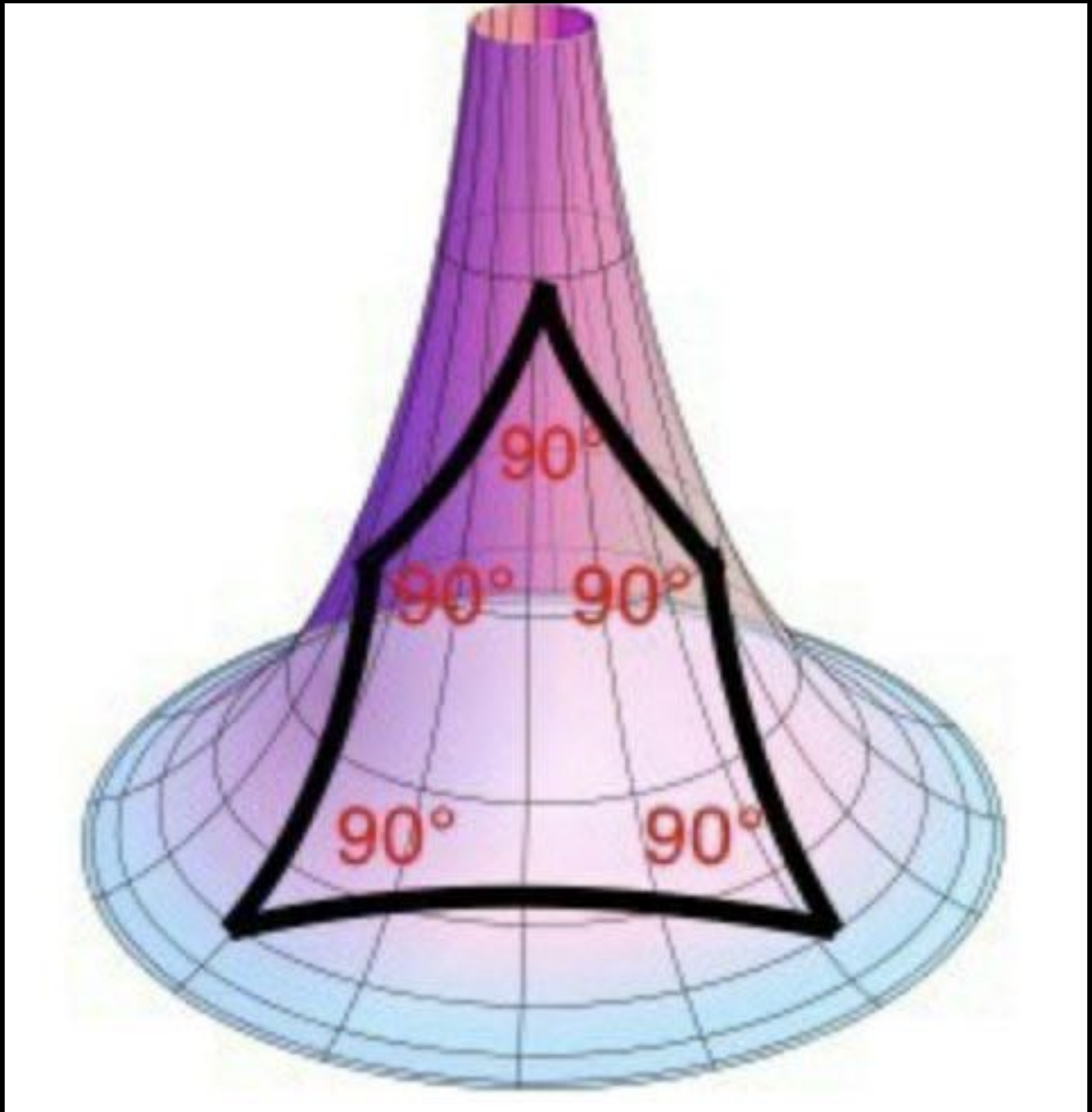


What is Non-Euclidian geometry?

Among other things, this gives us five sides squares.

Actually, our space *is* hyperbolic due to the gravity of the Sun.

This is a very nice explanation https://youtu.be/n7GYYerlQWs

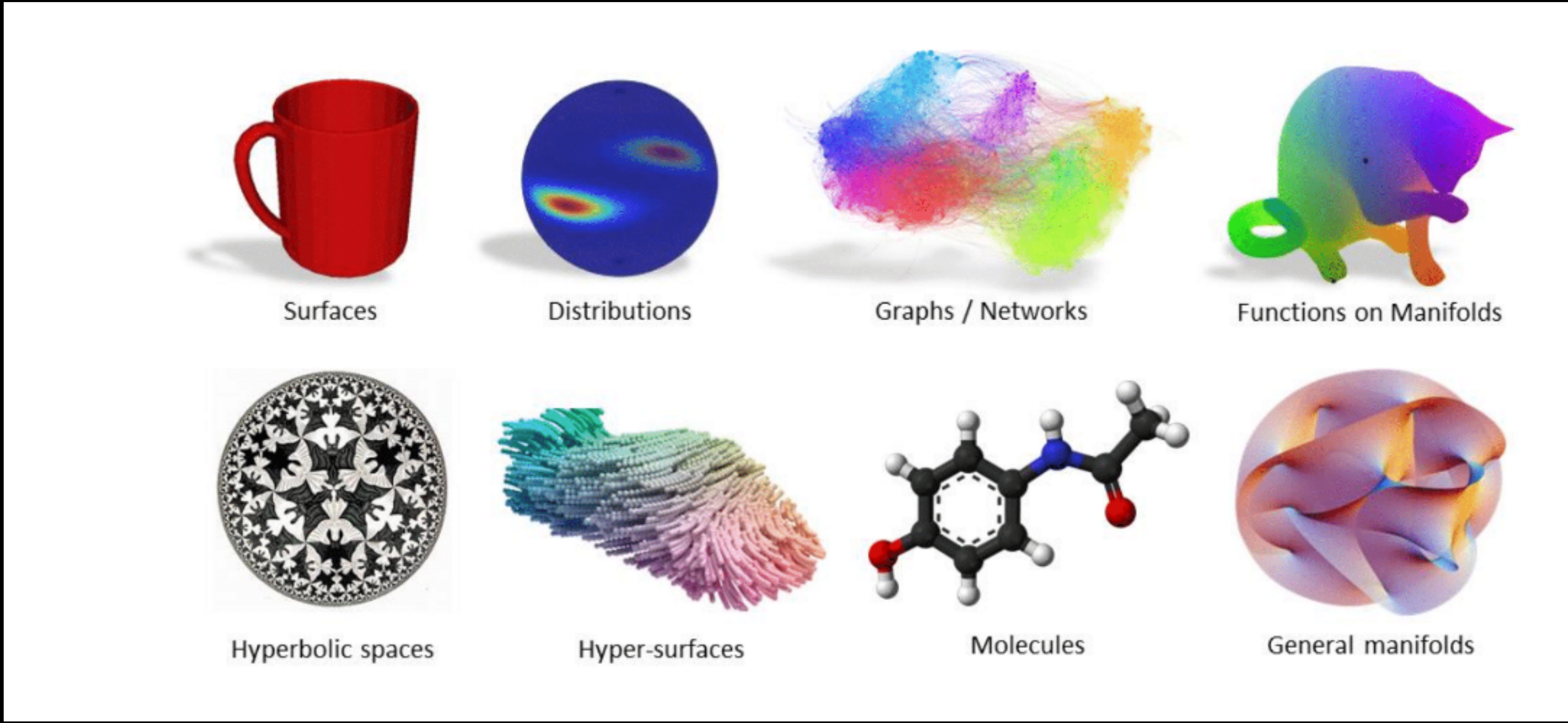


And why is this relevant for us?

What is Non-Euclidian data?

- Vectorspaces like $\mathbb{R}^d$ *can* have a metric (distance measure).
- But not all data is best modeled with the principles of a vectorspace
- or they do, but an euclidian metric is not the best way to model the data

What is Non—Euclidian data?



What is Non-Euclidian data?

- For example:
  - Some data doesn’t even has a good notion of distance, or is irregular (eg with holes or curves). These are no longer vectorspaces, but can still be topologies.
  - Topologies just need to define what it means to be “near” without a specific distance.

What is Non-Euclidian data?

- For example:
  - For some data, their natural shape is not “flat”, for example hierarchies
  - There are hyperbolic spaces $\mathbb{H}^d$ where the parallel postulate does not hold. They can still have a metric. Eg the space we live in, which is arguably relevant.

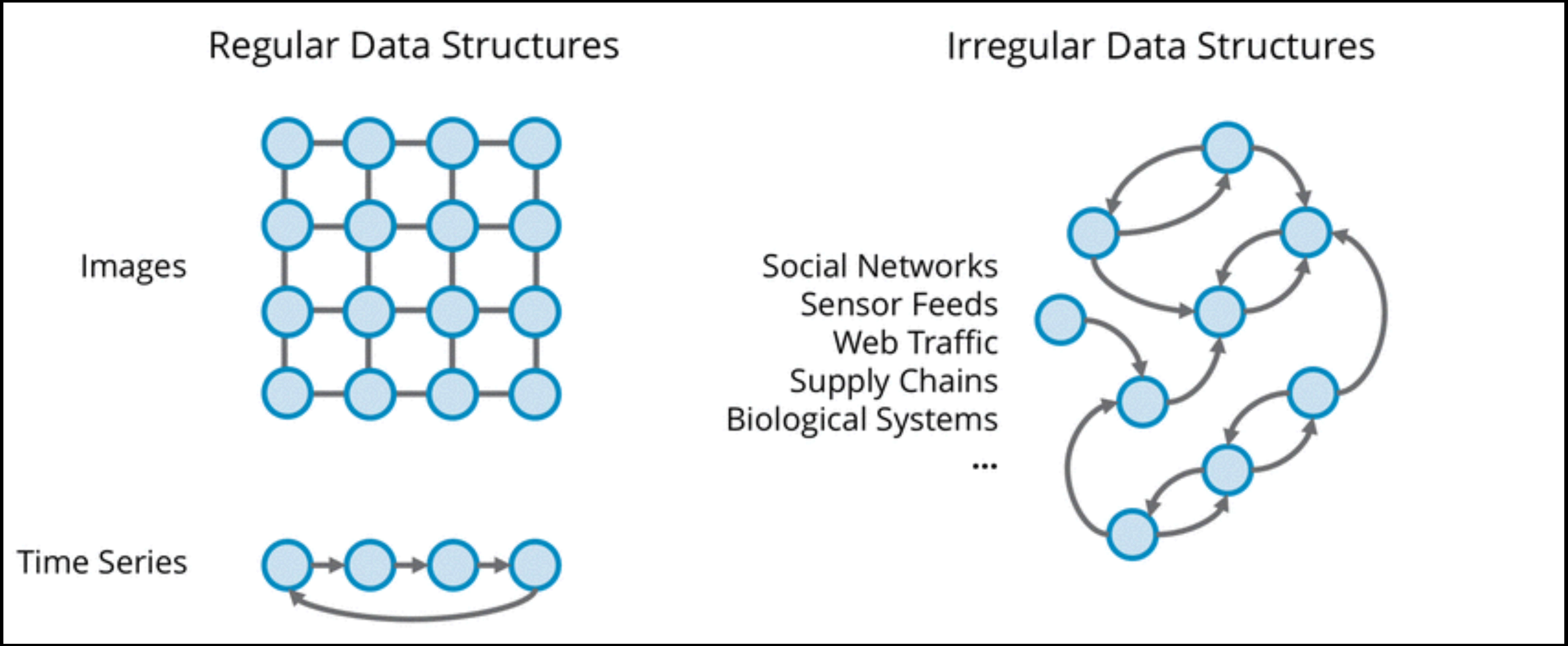
What is Non-Euclidian data?

- Try to imagine a 2D space with all possible normal distributions.
- We put μ and σ on an axis

What is Non-Euclidian data?

- Try to imagine a 2D space with all possible normal distributions.
- We put μ and σ on an axis

1. What is a point? Can we make a point more “fuzzy”?
2. What happens when we “wiggle” μ? And σ?
3. What does it mean for a distibution to be close? Or “different”?
4. Is $\mathcal{N}(0,0)$ closer to $\mathcal{N}(1,0)$ or $\mathcal{N}(0,1)$?
5. Is “close” or “different” everywhere the same in the graph? Eg $\mathcal{N}(100,100)$ and $\mathcal{N}(100,101)$?How is our ruler “impacted”?


Regular Data Structures
Irregular Data Structures
Images
Time Series
Social Networks
Sensor Feeds
Web Traffic
Supply Chains
Biological Systems
...

What is Hyperbolic space?

- In a vectorspace, one of the key axioms is that you can scale vectors (closure under multiplication)
- hyperbolic space is just an example of non-euclidian spaces
- In hyperbolic space scaling doesn’t work the same

What is Hyperbolic space?

- There still is a notion of distance (metric), a notion of structure (you can move in any direction) and the structure is smooth
- But it doesn’t have:
  - The axiom of cluser under multiplication
  - The axiom of distributivity $c(u+v) = cu + cv$
  - Linear structure in general

Hyperbolic embeddings

- In Euclidian space ($\mathbb{R}^d$), the volume of a ball grows polynomially:

$$V_E(r) \propto r^d$$

- In Hyperbolic space ($\mathbb{H}^d$), the volume of a ball grows exponentially for larger r:

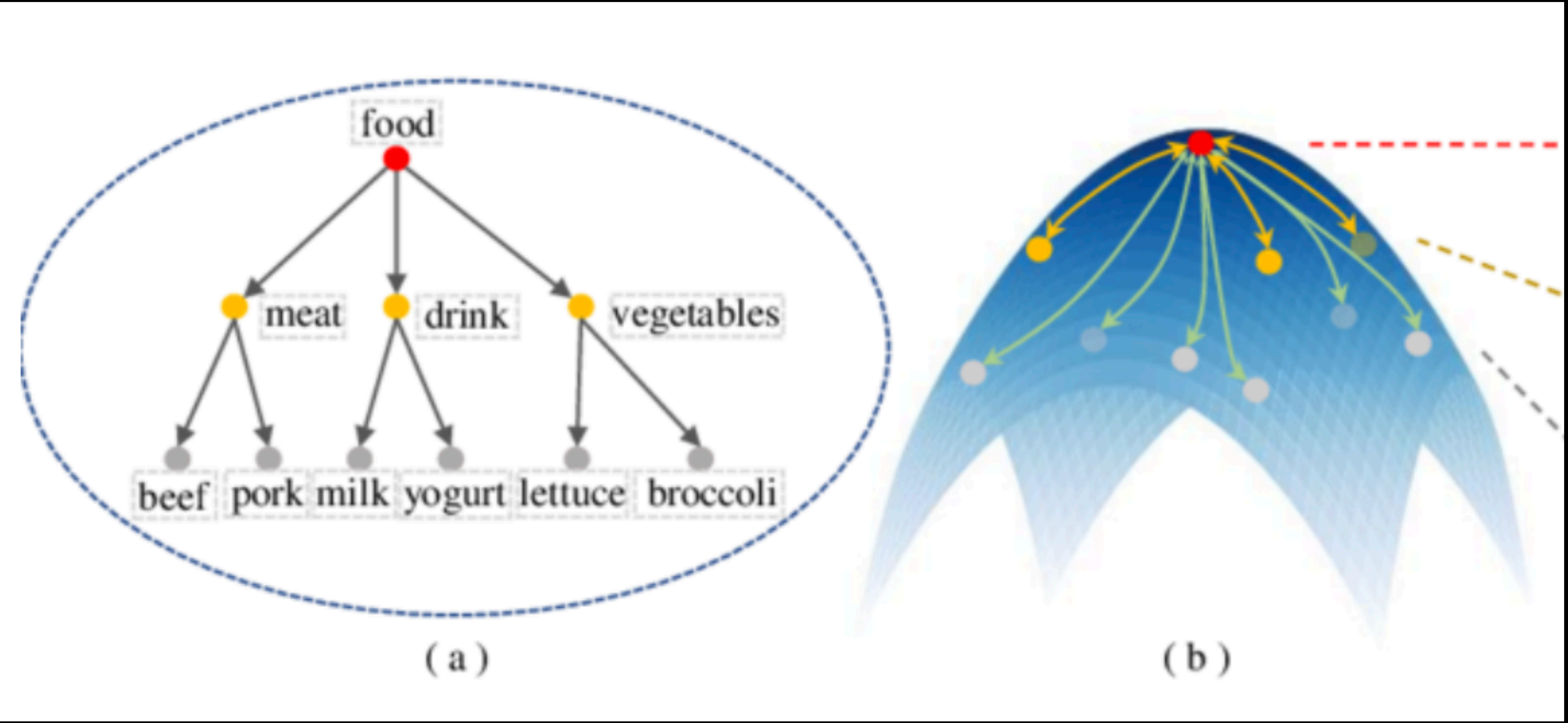
$$V_H(r) \propto sinh^{d-1}(r) \propto e^r$$

Okay, but why do I care?
What do I need to do different?

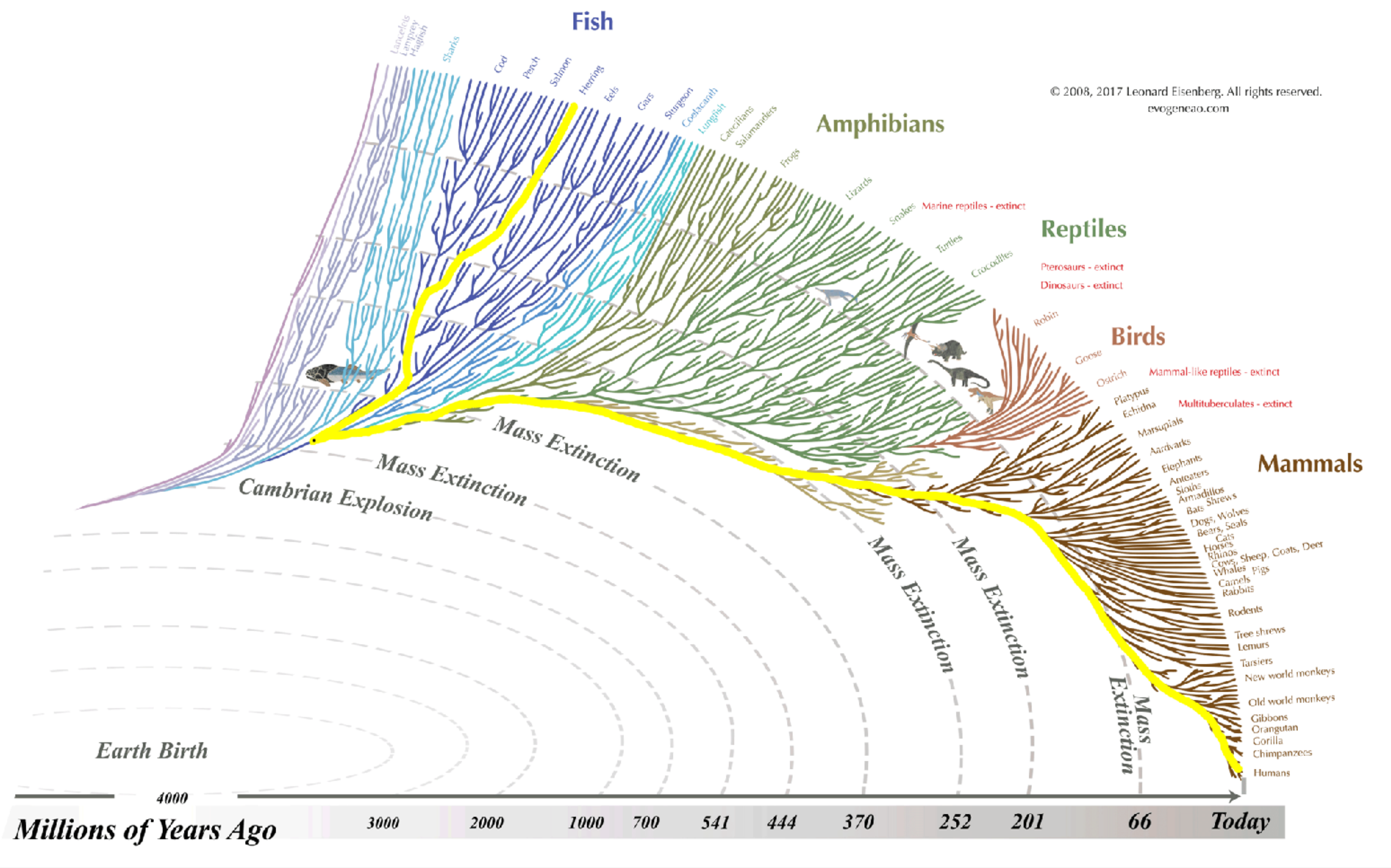
Hyperbolic embeddings

- In plain language, this means there is increasingly “more space” when you go towards the edge
- This is the reason hyperbolic embeddings are particularly well suited to use with tree-like structures that need more space for the “leaves” of the tree
- For example, hierarchical structures with main categories and subcategories.

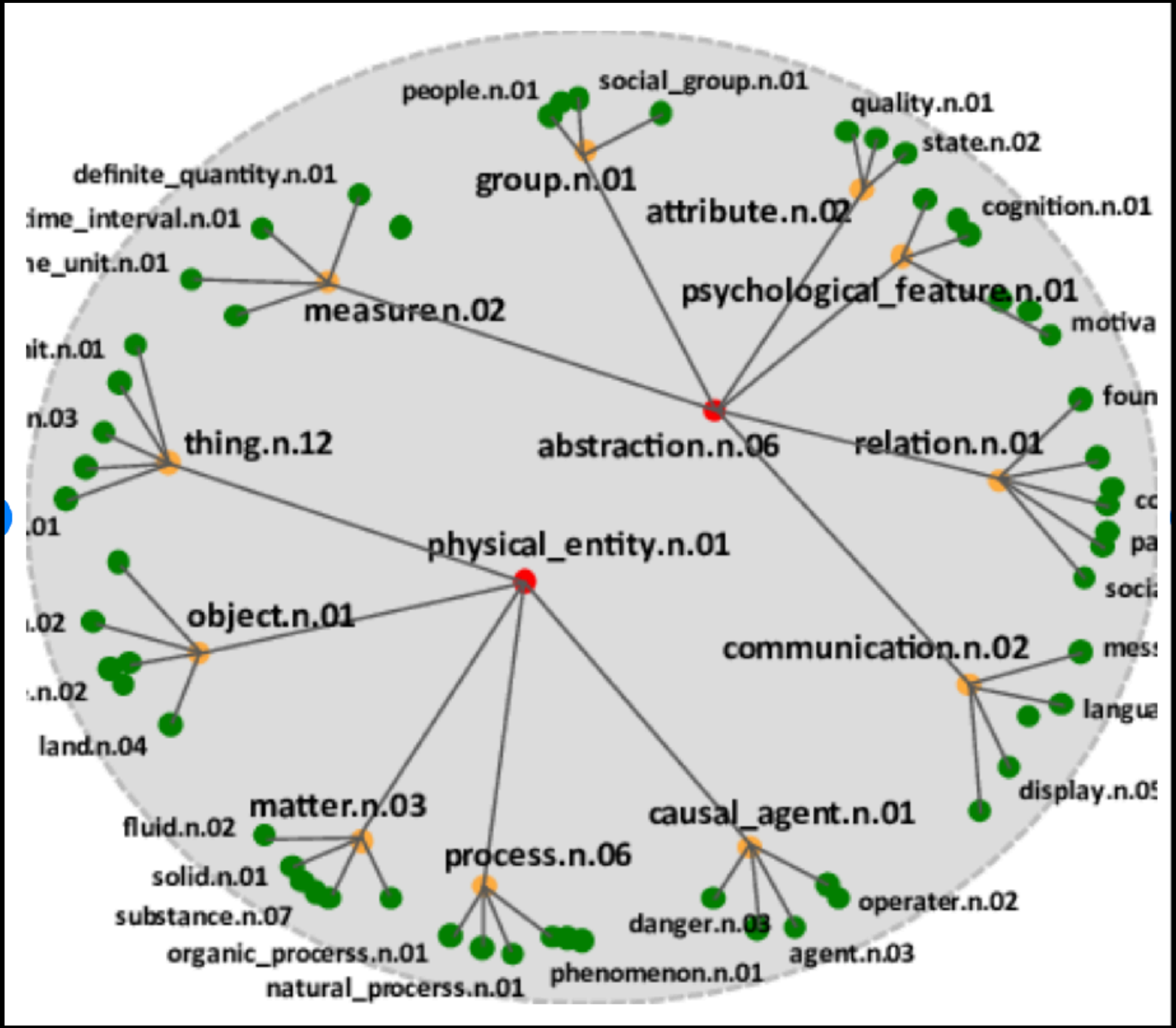
Hyperbolic embeddings



Hyperbolic embeddings



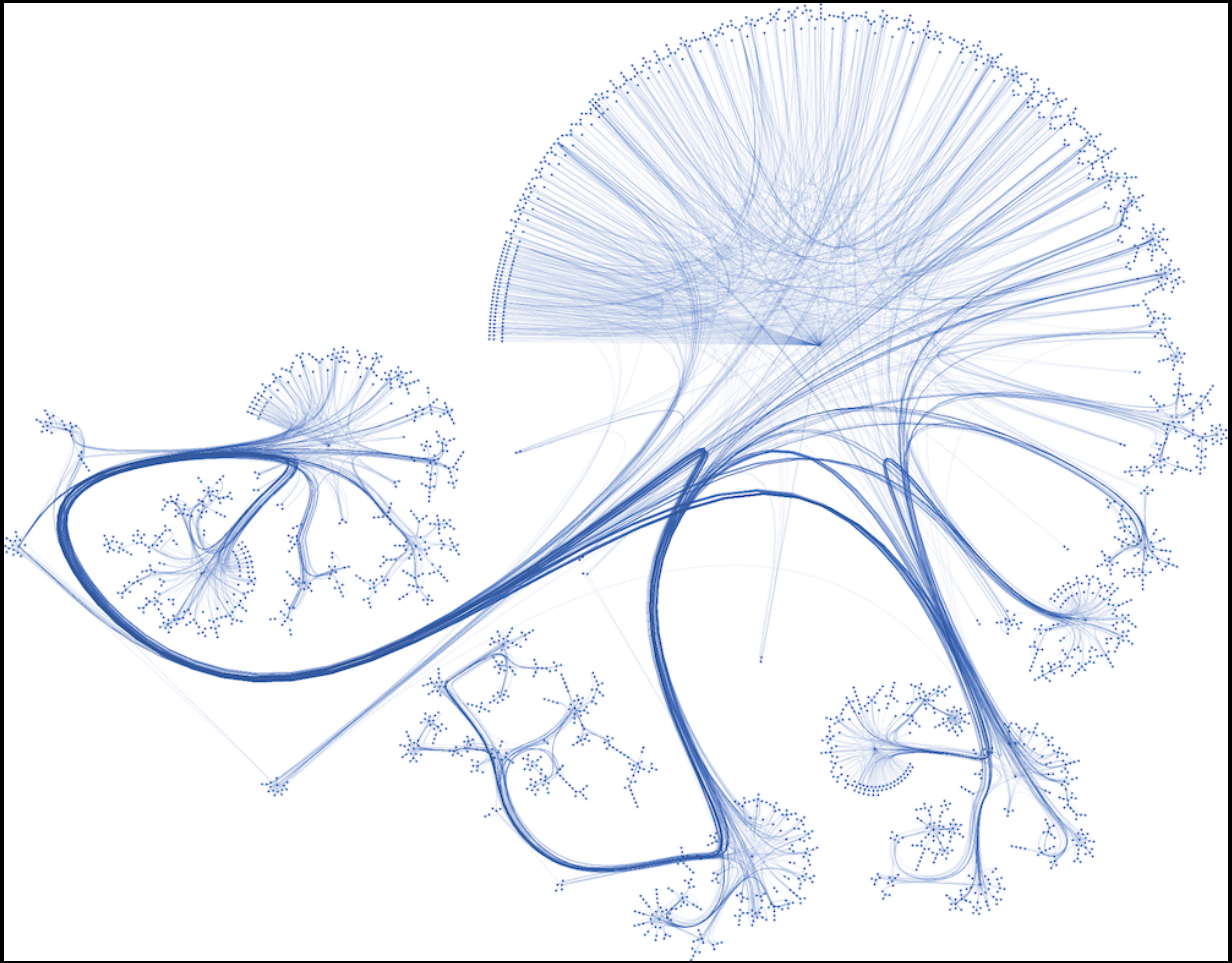
Hyperbolic embeddings



Graphs

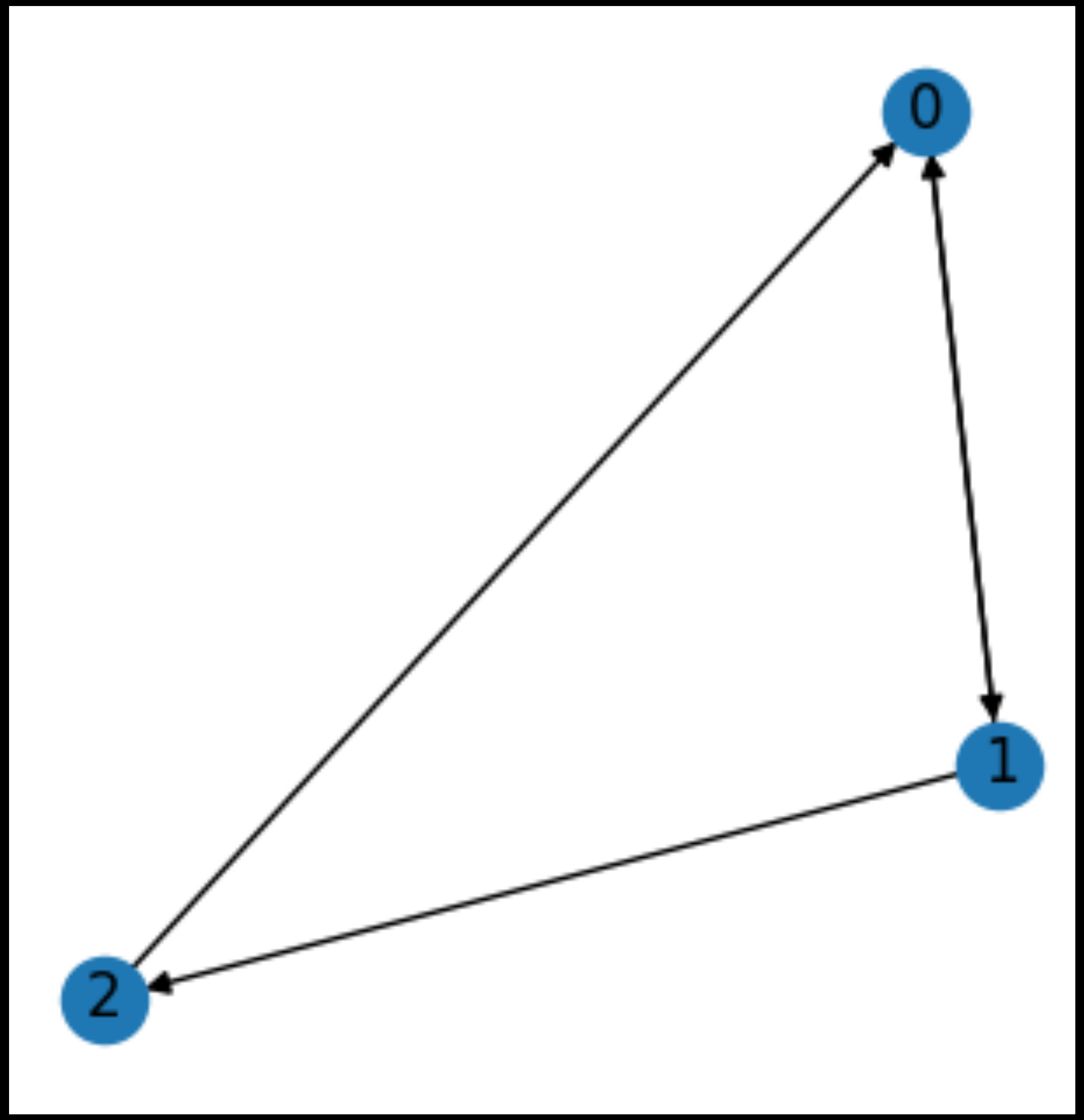
Cora dataset

- 2708 scientific publications
- classified into one of seven classes.
- The citation network consists of 5429 links.
- Each publication in the dataset is described by a 0/1-valued word vector indicating the absence/presence of the corresponding word from the dictionary. The dictionary consists of 1433 unique words



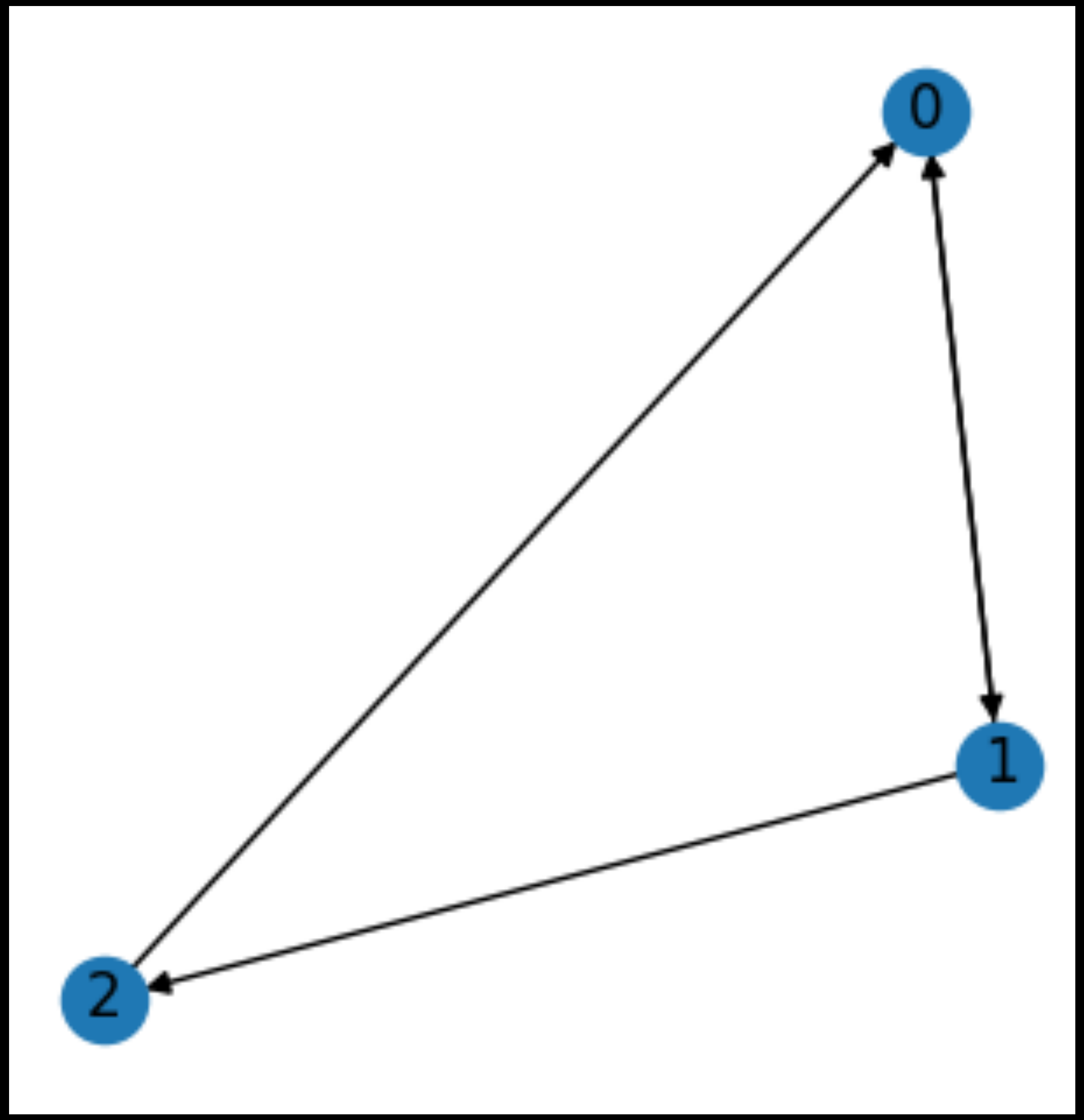
What is a Graph?

- A Graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ is defined by:
  - A set of nodes $\mathcal{V} = \{v_1, \ldots, v_n\}$
  - A set of edges between nodes $\mathcal{E} = \{(v_i, v_j) \mid v_i, v_j \in \mathcal{V}\}$



What is a Graph?

- The adjecency matrix A has a 1 on every position where there is and edge: $A[i,j] = 1$ if $e_{i,j} \in \mathscr{E}$
- Excercise: let's draw A for this graph!



Some statistics on graphs

Degree

- Degree: the number of edges for a node $d_u = \Sigma_{v \in V} A[u, v]$
- The degree matrix D has on each diagonal element the degree of the note: $D[i, i] = d_i$

Some statistics on graphs
Laplacian

The Laplacian matrix is defined as

$$L = D - A$$

Among other things, it can tell you if there are groups of nodes who are all connected to each other but not much to others in the network.

It's like spotting cliques in your group

Some statistics on graphs
Betweenness

- Betweenness centrality is the sum of the fraction of all shortest paths through v:

$$cb(v) = \Sigma_{s,t \in V} \frac{\sigma(s, t \mid v)}{\sigma(s, t)}$$

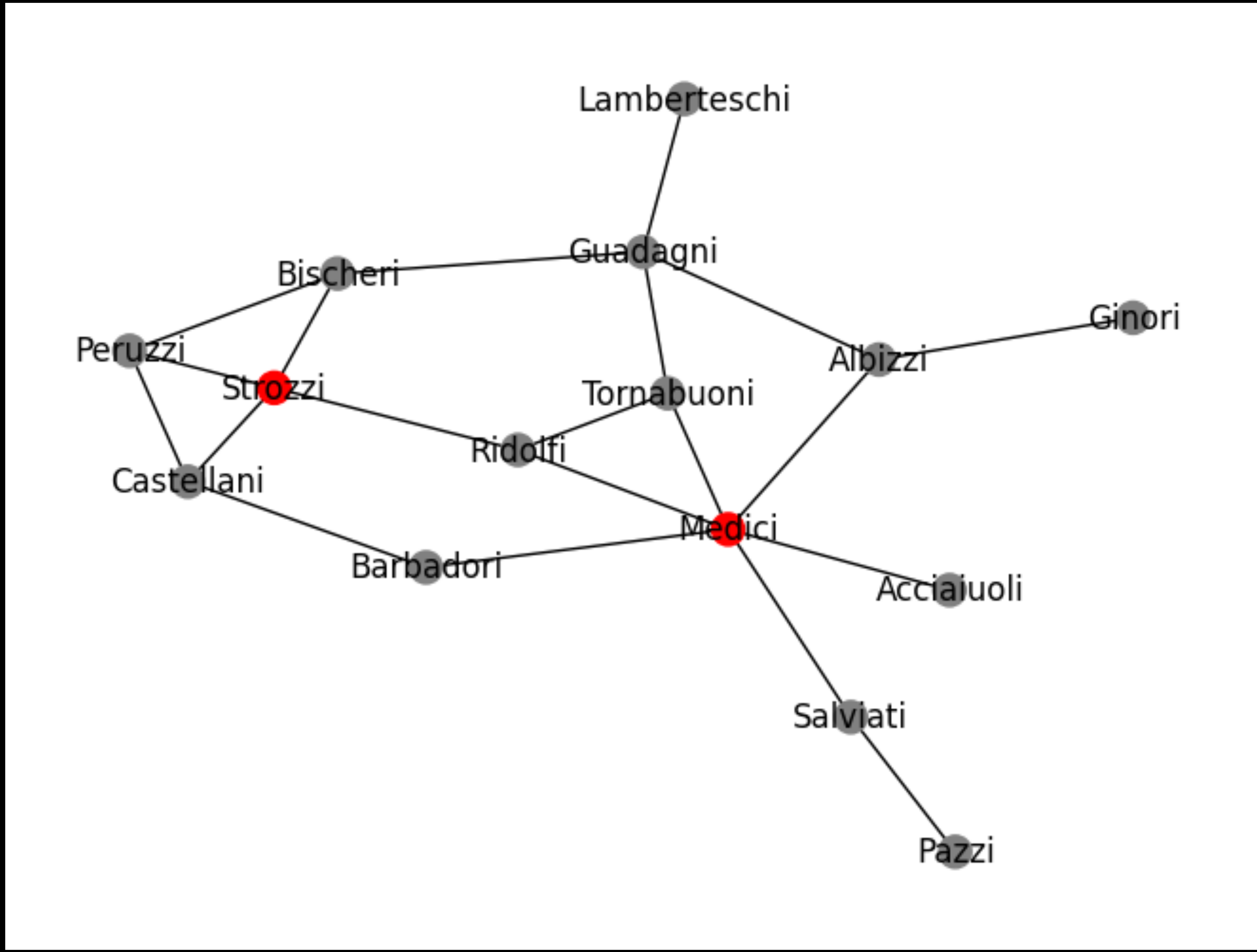
- with $\sigma(s, t)$ the number of shortest (s,t) paths and $\sigma(s, t \mid v)$ the paths through v.

Florentine Families

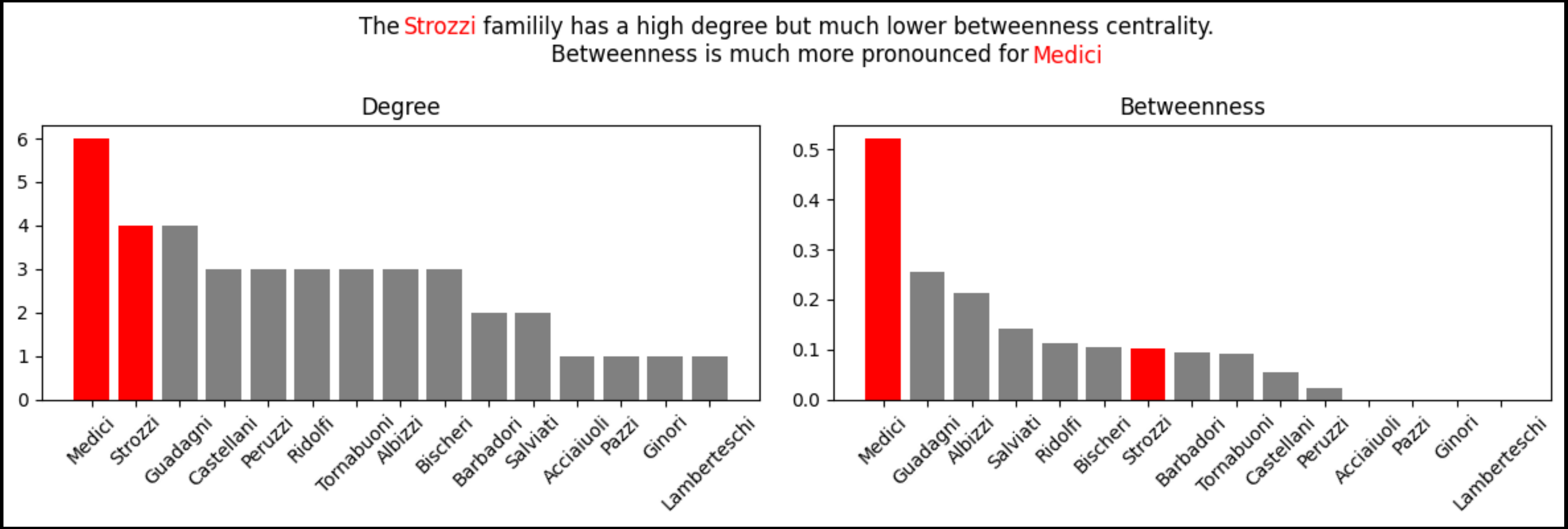
Renaissance Florentine families around 1430, collected by John Padgett from historical documents.

The graph shows marriage alliances.

The data include families who were locked in a struggle for political control. Two factions were dominant in this struggle: one revolved around the Medicis, the other around the Strozzis.



Florentine Families



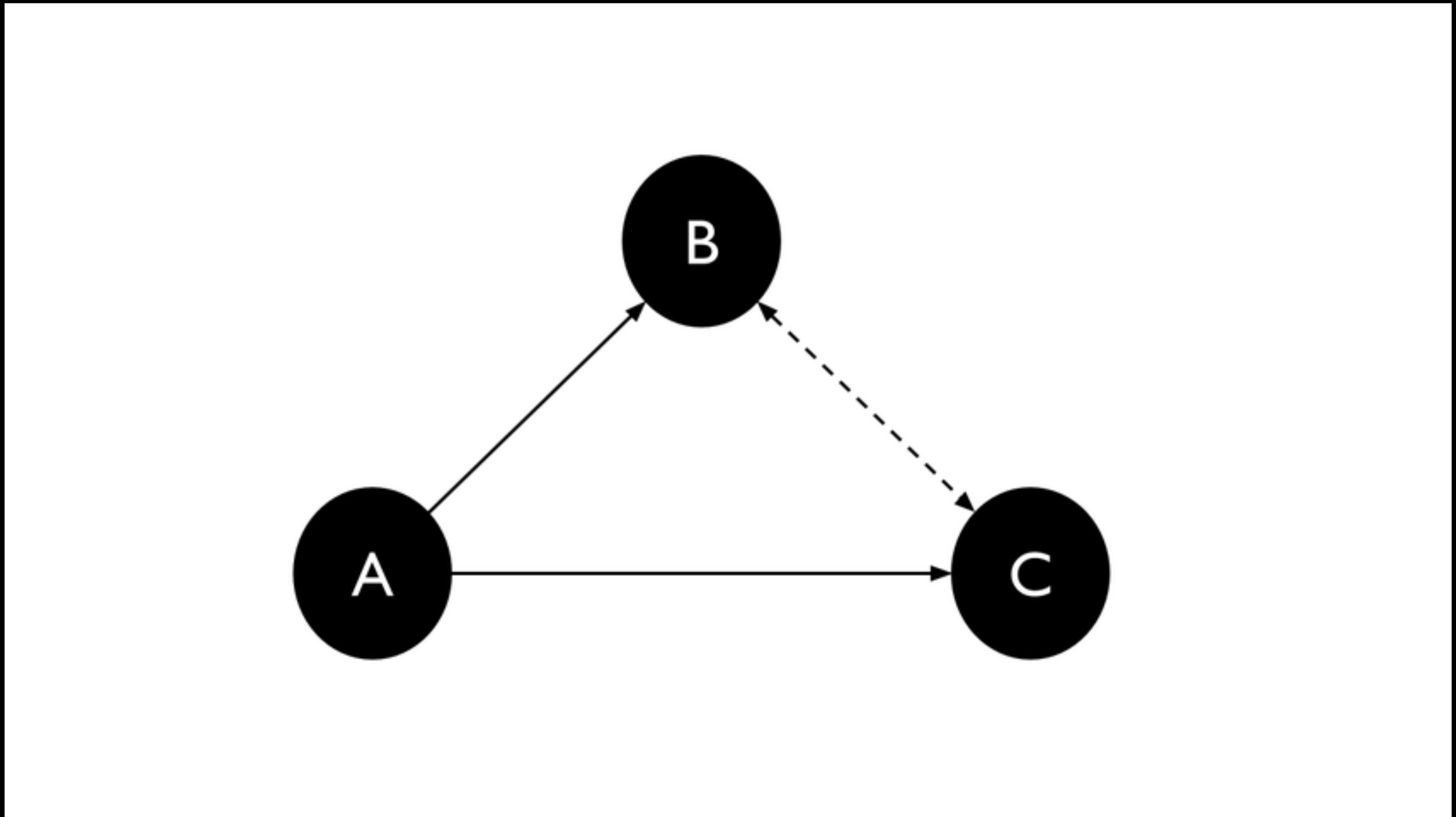
Some statistics on graphs

Triadic closure

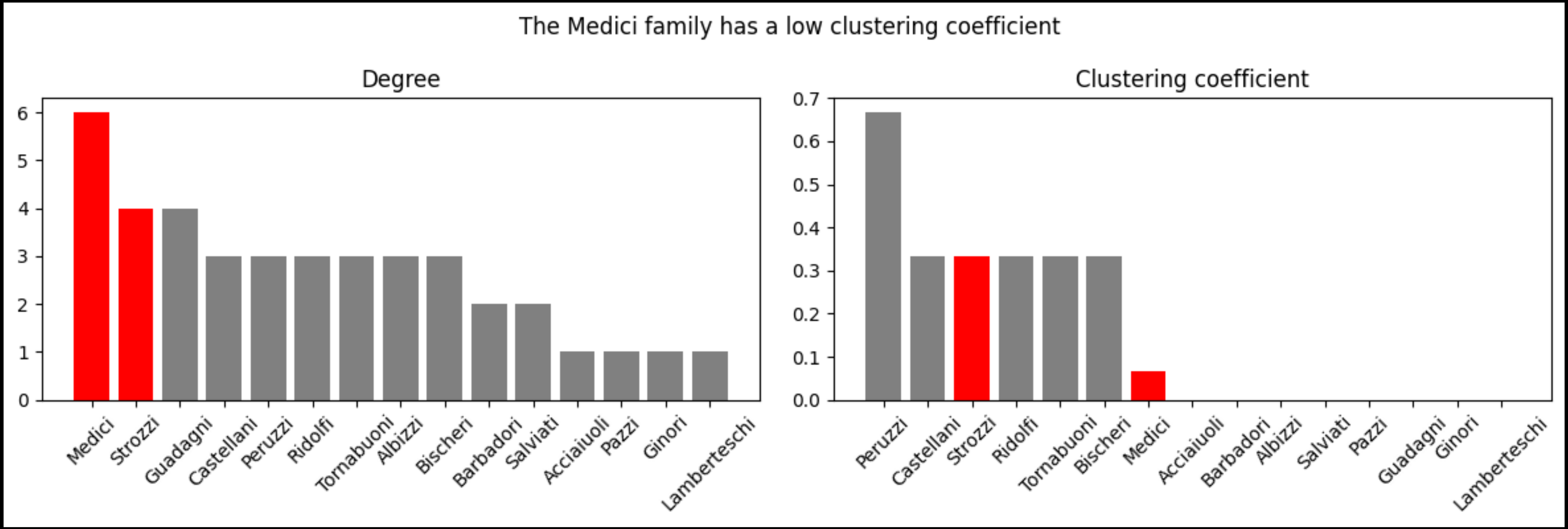
- Triangles: if your friend are also friends
- Clustering coefficient: the fraction of possible triangles.
- This can be a very relevant metric: e.g. there is a correlation between (lack of) triadic closure for teenagers and depression!

Some statistics on graphs

Triadic closure

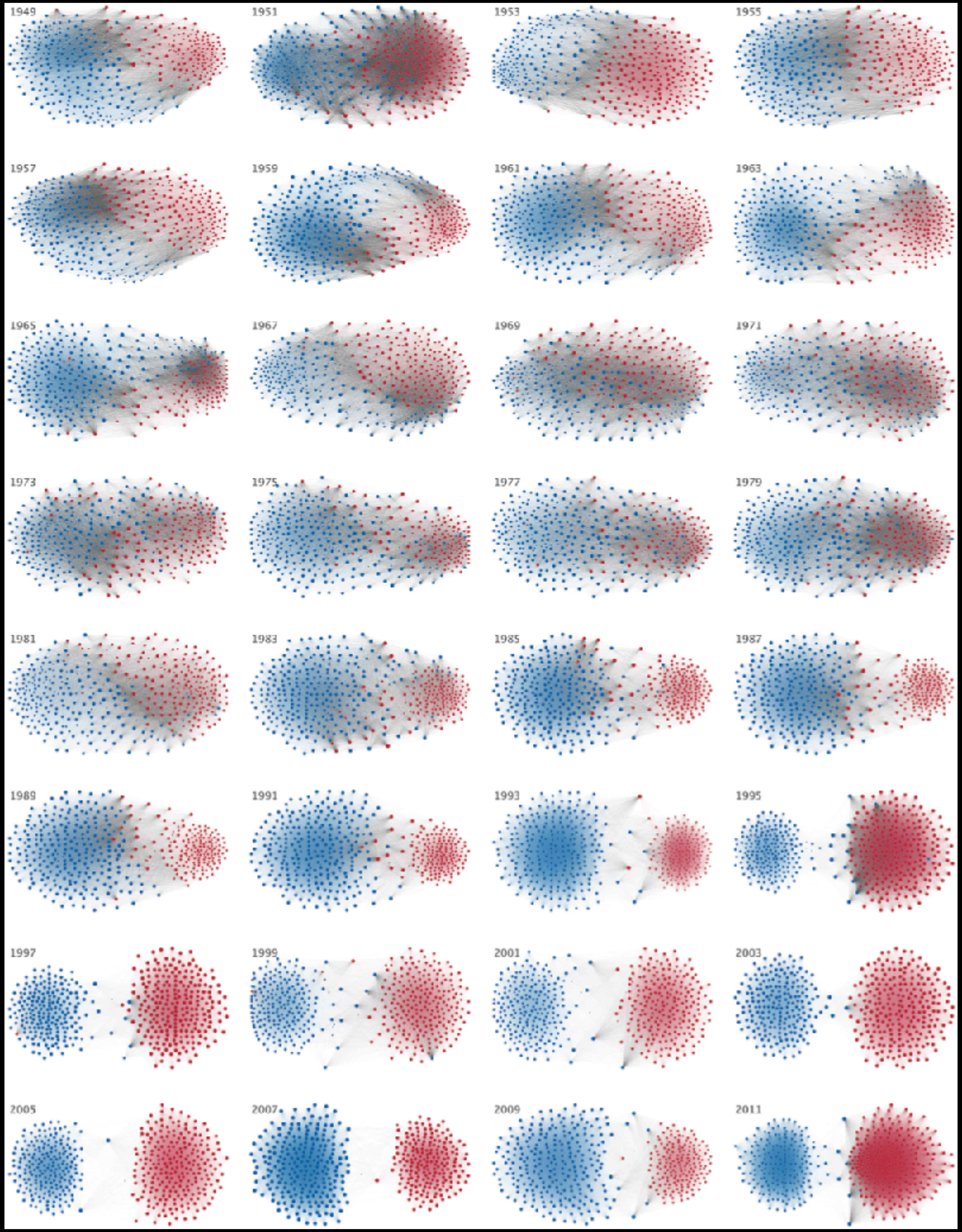


Florentine Families



See US Congress polarize over the past 60 years

See how likely the House of Representatives' Democrats (in blue) and Republicans (in red) are to vote with their own party, or to cross party lines.

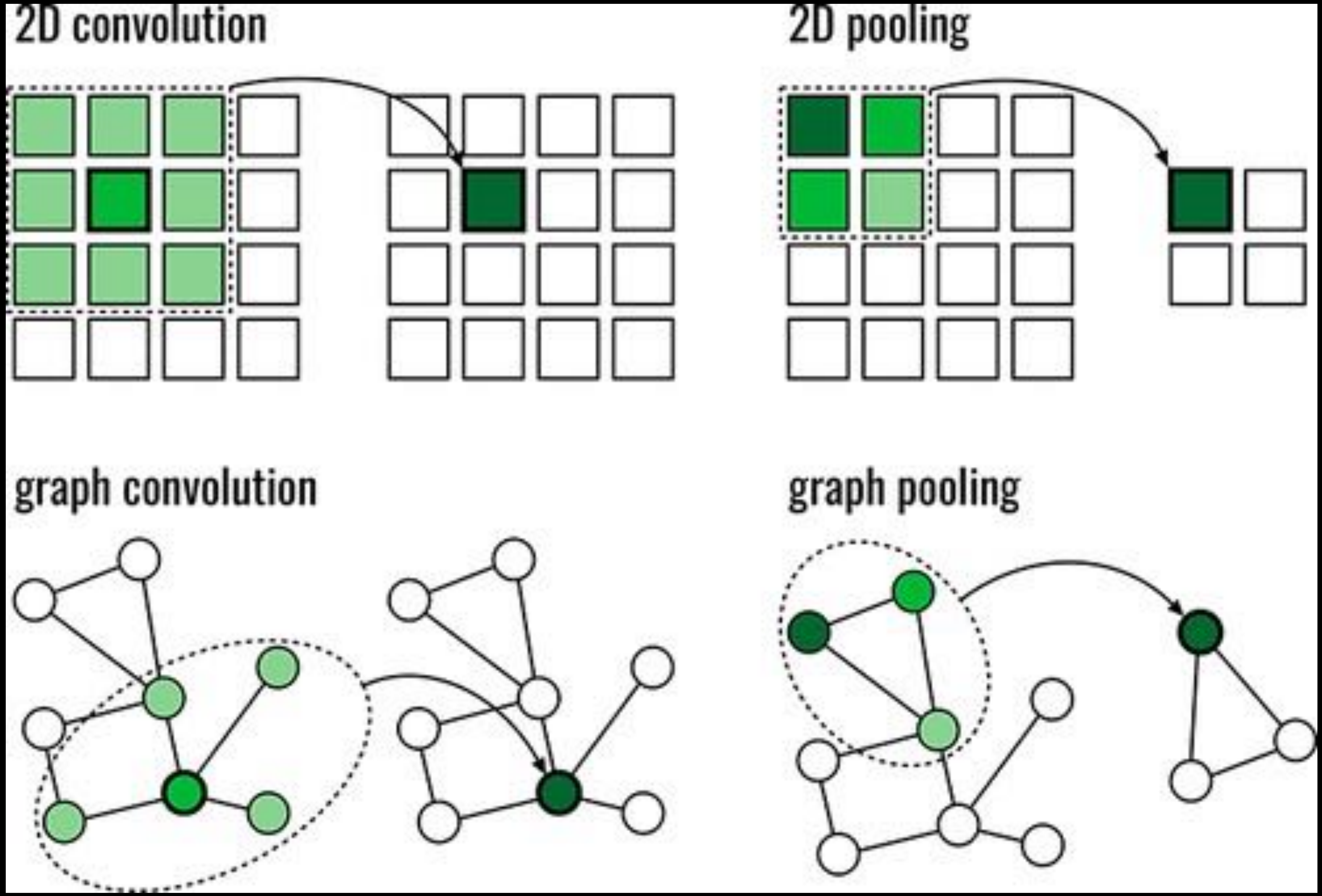


Can we do deep learning with graphs?

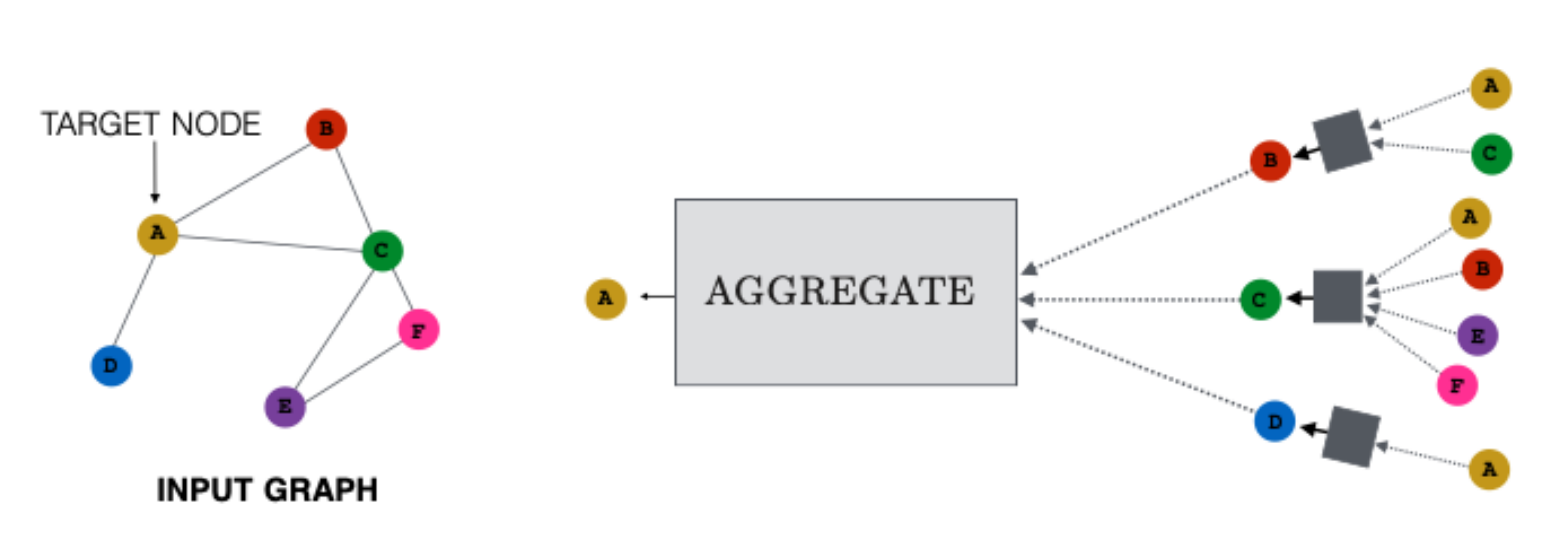
Issues with machine learning on graphs

- Size and shape: graphs can have wildly different shapes. Unlike images, we can just resize.
- Isomorphism: We can make multiple adjecency matrices for the same graph, by arbitrary reordering the nodes.
- However, these *different* matrices represent the *same* graph. We don't want arbitrary reordering to matter.

Solution: Graph convolutions



Graph convolutions



Graph convolutions

- At each iteration, every node aggregates information from its local neighborhood
- After k convolutions, each node contains information from its k-hop neighborhood

Graph convolutions

Basic message passing

- $h_u^{(k)}$ is the embedding of node u after k convolutions.
- $W_{self}, W_{neigh} \in \mathbb{R}^{d_k \times d_{k-1}}$ are trainable weights

$$\mathbf{h}_u^{(k)} = \sigma\left(\mathbf{W}_{\text{self}}^{(k)}\mathbf{h}_u^{(k-1)} + \mathbf{W}_{\text{neigh}}^{(k)} \sum_{v\in\mathcal{N}(u)} \mathbf{h}_v^{(k-1)} + \mathbf{b}^{(k)}\right)$$