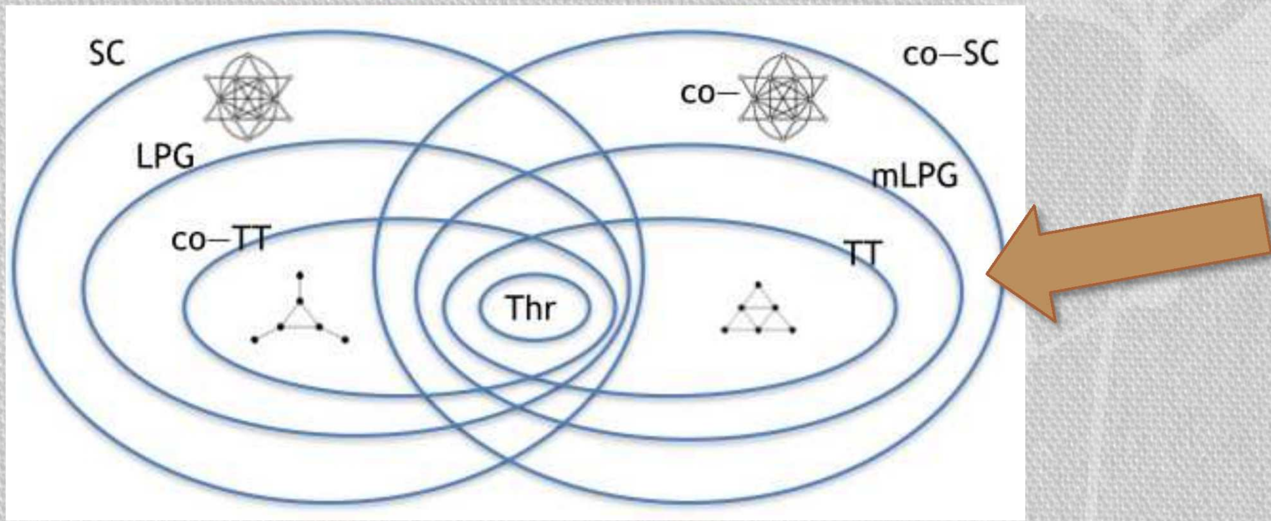
RELATING THRESHOLD TOLERANCE GRAPHS TO OTHER CLASSES OF GRAPHS

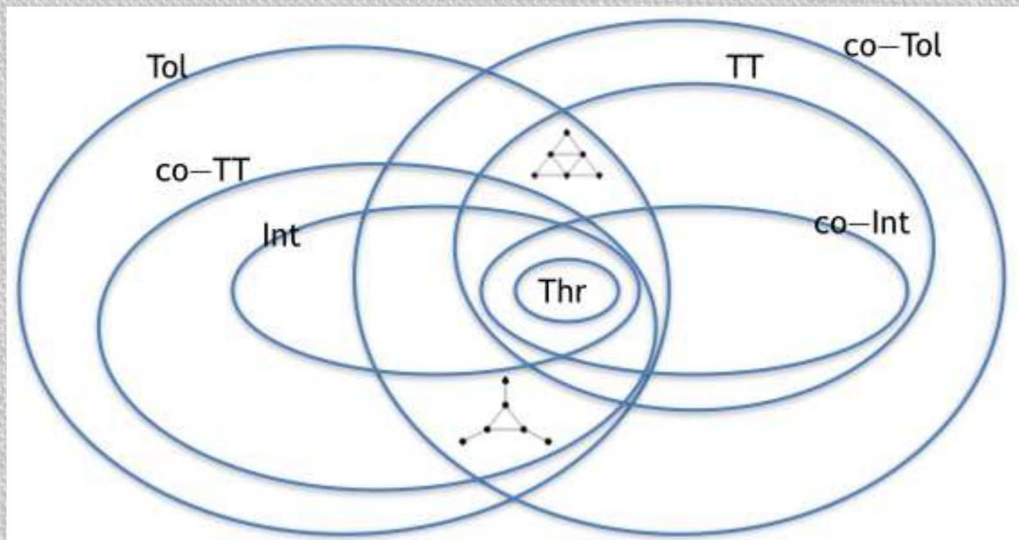
Tiziana Calamoneri & Blerina Sinaimeri

ICTCS 2014

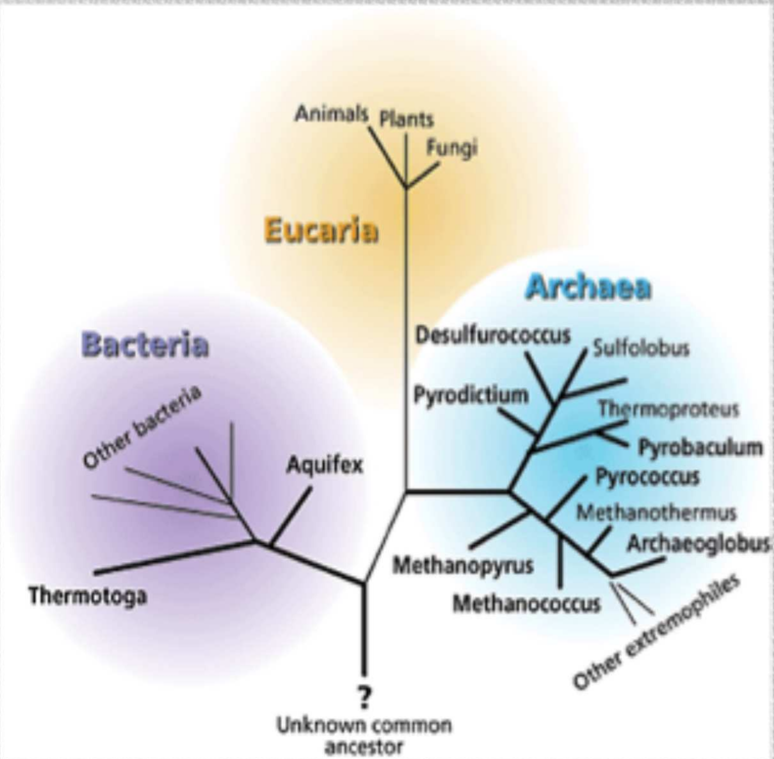
Perugia, 17th September 2014

TTGs w.r.t. other graph classes





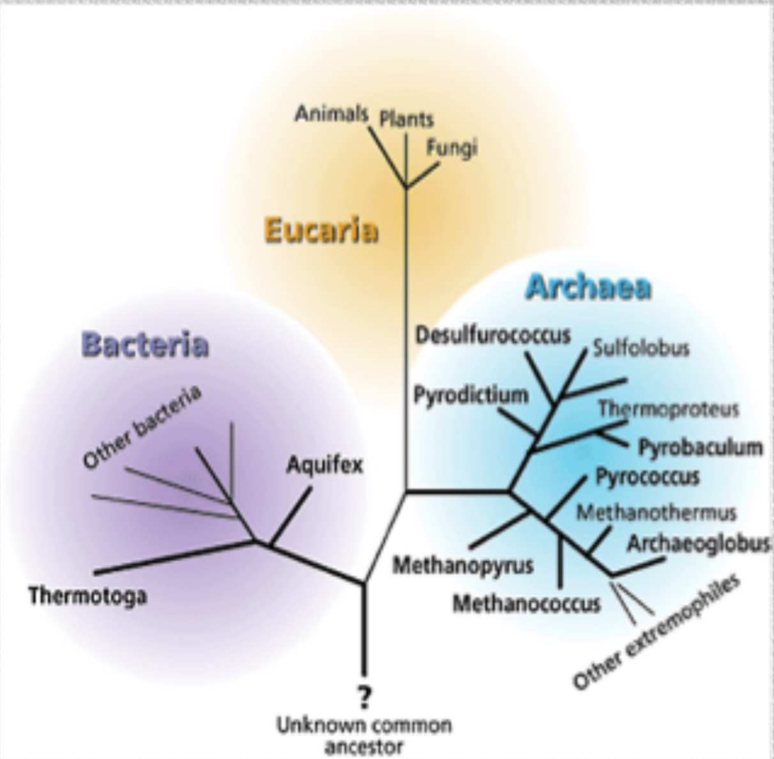
A phylogenetic problem (1)



The aim is to reconstruct a tree expressing evolutionary relations among organisms, on the basis of biological data [Jones & Pevzner '04].

- ♣ Leaves ⇔ known species (taxa)
- ♣ Internal nodes ⇔ (hypotetical) ancestors
- ♣ Edges ⇔ evolutionary relations between nodes
- ♣ weight of the edges ⇔ node distance in terms of evolution

A phylogenetic problem (2)



- ♣ Somehow, internal nodes represent moments of speciation (i.e. the creation of new species from an old one).
- ♣ It is not possible to know which is the "true" tree of a certain set of taxa, nevertheless biologists are quite sure of certain speciations (e.g. on the basis of the study of fossils).
- ♣ For some sets of taxa, biologists have built some phylogenetic trees accepted as "true".

A phylogenetic problem (3)

- ♣The automatization of the creation process of a phylogenetic tree given the set of taxa involves solving an optimization problem (e.g. find the tree that minimizes the total number of evolutionary events has to be individuated)
- ♣Usually, these optimizations lead to NP-hard problems, so the reconstruction algorithms are in fact heuristics that need to be tested.
- ♣The results of these heuristics are compared with the trees considered as “true”.

A phylogenetic problem (4)

- ♣ In general, these “true” trees are huge and the reconstruction heuristics are slow…
- ♣ It is hence important to extract subtrees from the “true” trees in order to test the reconstruction heuristics on the subsets of taxa that are involved in these subtrees.

A phylogenetic problem (5)

- ♣ Many reconstruction heuristics fail in the reconstruction if the considered taxa have a very large evolutionary distance.
- ♣ Analogously, these heuristics fail when the considered taxa have a very small evolutionary distance [Felseinster '78].
- ♣ Hence:
  try to find a set of taxa that are neither too close nor too far in order to test the heuristics on the subtree induced by this set → sampling.

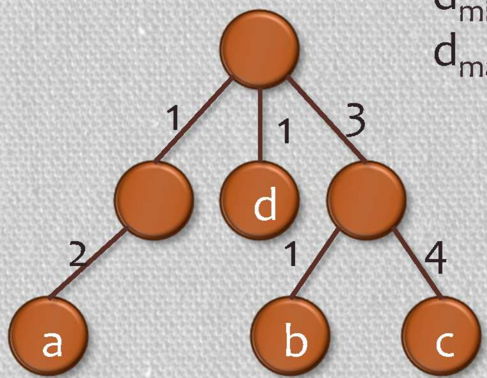
Pairwise Compatibility Graphs (1)

[Kearney, Munro & Phillips '03] formulated these constraints in graph theory, so introducing the **Pairwise Compatibility Graphs (PCGs)**.

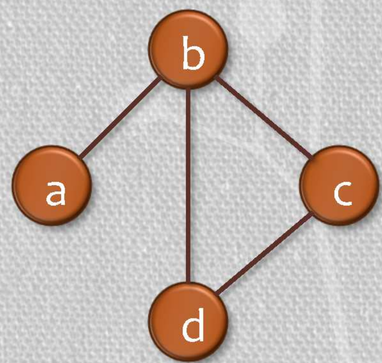
Given a phylogenetic tree (i.e. an edge-weighted tree) T and two positive values d_{min} and d_{max}, a pairwise compatibility graph $G=PCG(T,d_{min},d_{max})$ is defined as follows:

nodes of G ⇔ leaves of T

edges of G ⇔ paths in T having length between d_{min} and d_{max}



$d_{min}=5$
$d_{max}=8$



Pairwise Compatibility Graphs (2)

Given a tree T, to solve the sampling problem is equivalent to seek for a maximum cardinality clique in $G=PCG(T,d_{min},d_{max})$.

In [Kearney, Munro & Phillips '03] it is proved that, for this class of graphs, MAX CLIQUE can be solved in polynomial time.

Given T, d_{min} and d_{max} it is trivial to determine G.

Pairwise Compatibility Problem:

Given a graph G, there exists a tree T and two positive values d_{min} and d_{max} such that $G=PCG(T,d_{min},d_{max})$?

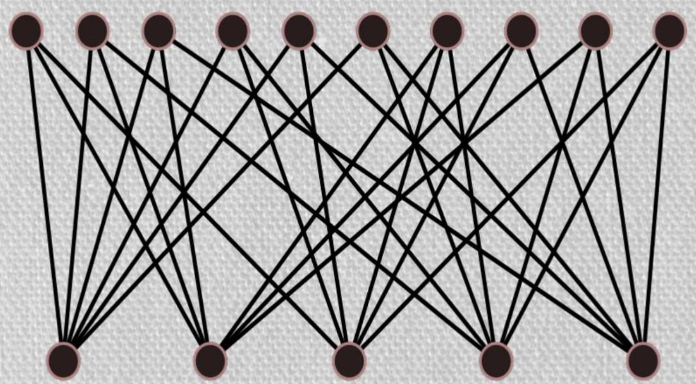
Pairwise Compatibility Graphs (3)

The problem is not trivial:

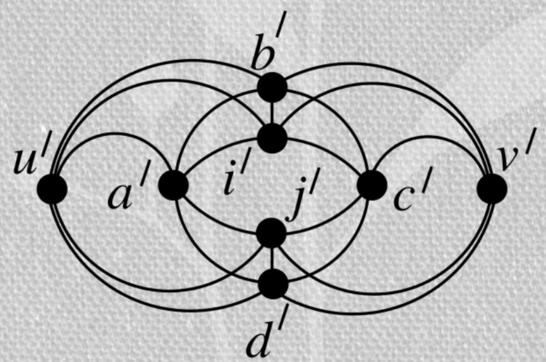
♣All graphs with at most 7 nodes are PCGs

[Philips '02; C., Frascaria & Sinaimeri '12]

♣Not all graphs are PCGs:



[Yanhaona, Bayzid & Rahman '10]



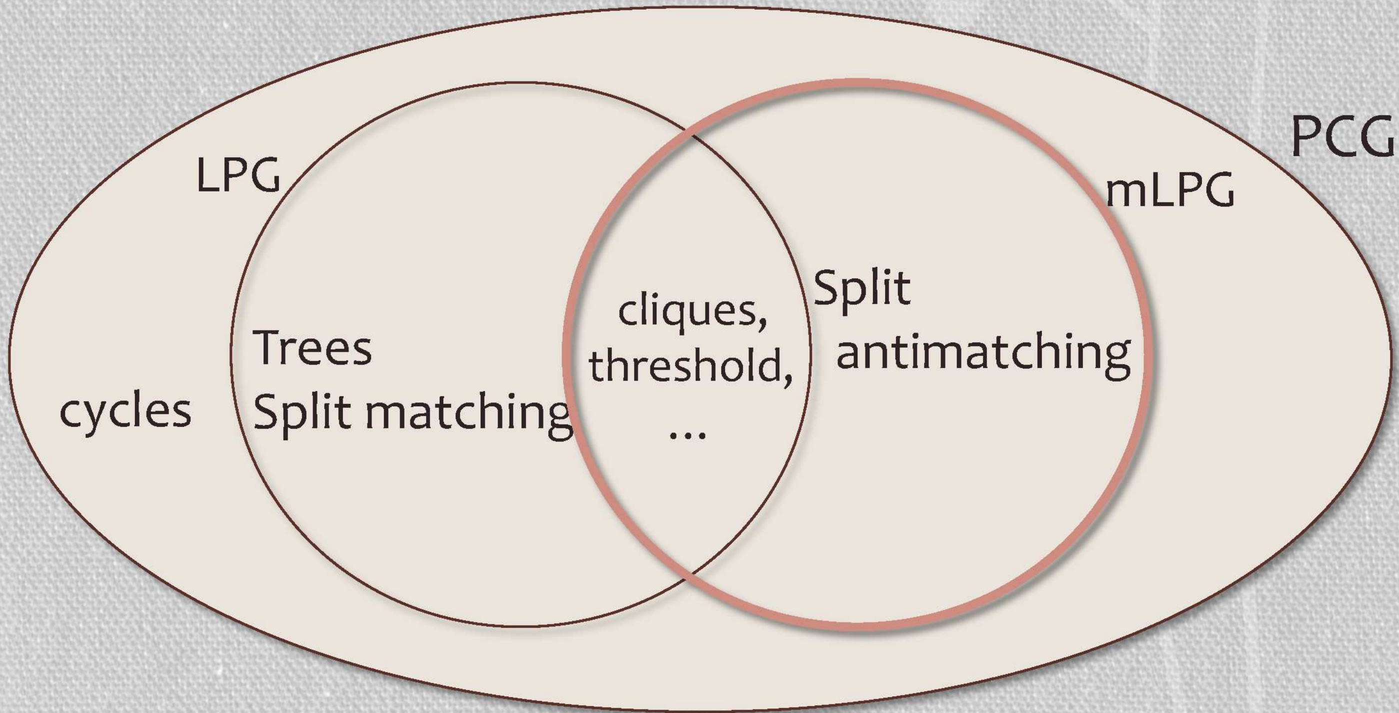
[Durocher, Mondala & Rahman '13]

Pairwise Compatibility Graphs (4)

It is possible to relax the requirements:

♣ Leaf Power Graphs [Nishimura, Ragde & Thilikos '02]: $d_{min}=0$

♣ minLeaf Power Graphs [C., Petreschi & Sinaimeri '12]: $d_{max}=+\infty$



Threshold tolerance graphs

- ♣ A graph $G=(V,E)$ is **threshold tolerance (TT)** if it is possible to associate weights g and tolerances t (in $\mathbb{R}^+$) with each node of G so that two nodes are adjacent exactly when the sum of their weights exceeds either one of their tolerances → $G=(V,E,g,t)$.
- ♣ It is not restrictive to assume that g and t are defined in $\mathbb{N}^+$.
- ♣ Threshold tolerance graphs generalize the class of **threshold graphs** which are also extensively studied in literature.

- ♣ Here we relate the threshold tolerance graphs with **min leaf power graphs (mLPGs)**.

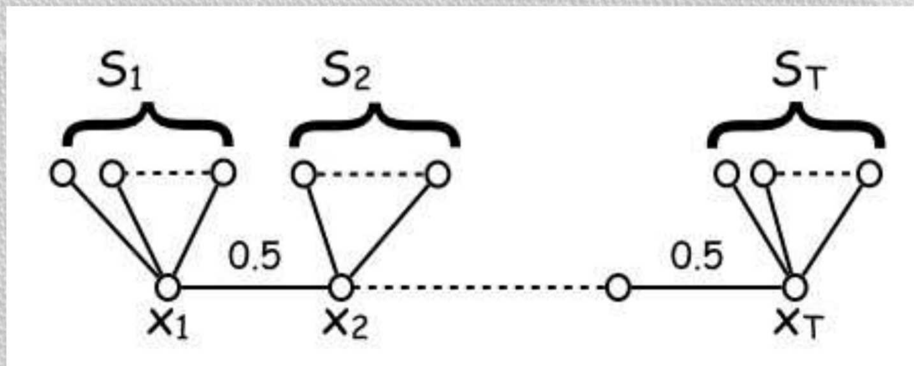
TTGs vs. mLPGs (1)

♣ **Theorem.** *Threshold tolerance graphs are mLPGs.*

♣ **Proof.** Let $G = (V, E, g, t)$ be a TTG. Let $T = \max_v t(v)$.
Split the nodes of G into groups $S_1, \ldots, S_T$ such that

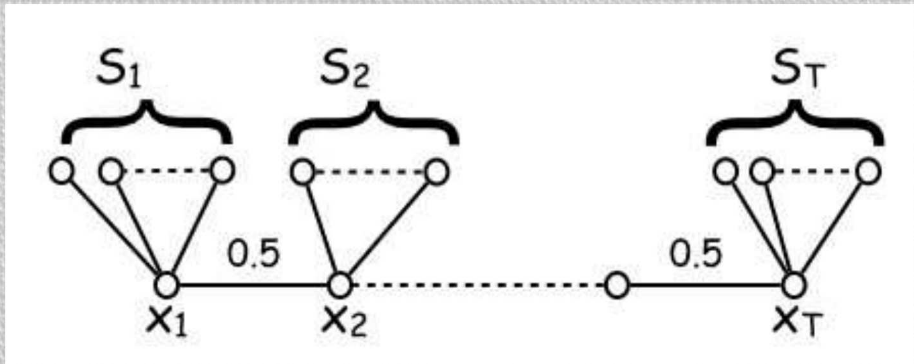
$$S_i = \{v \in V(G) : t(v) = i\}.$$

♣ Associate to G a caterpillar C (i.e. a tree in which all the nodes are within distance 1 of a central path, called *spine*):



TTGs vs. mLPGs (2)

Proof of Theorem. *Threshold tolerance graphs are mLPGs. (cnt.d)*

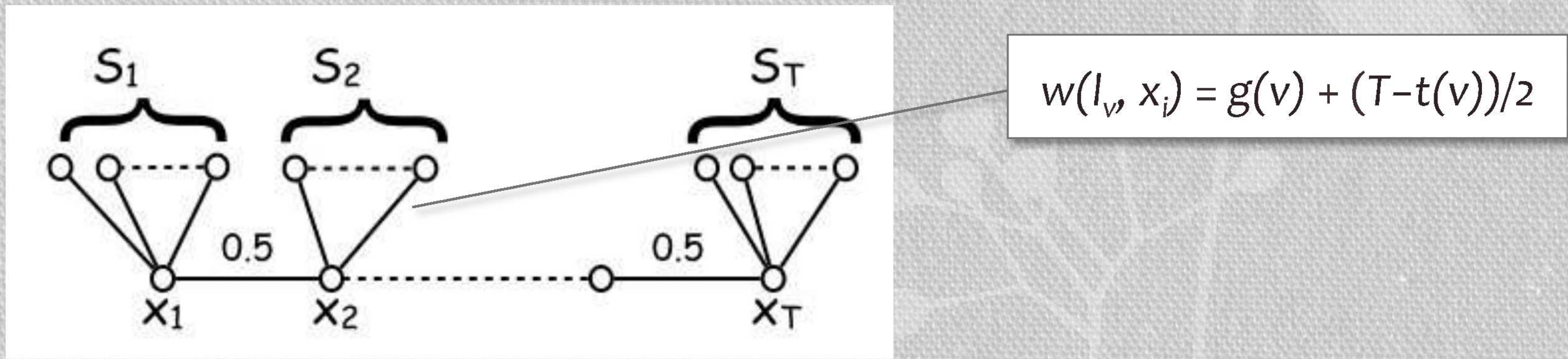


The weights *w* of the edges of C are defined as follows:

- For each edge of the spine $w(x_i,x_{i+1})=0.5$ for $0 \le i \le T-1$.
- For each leaf l_v connected to the spine through node x_i we assign a weight $w(l_v, x_i) = g(v) + (T-t(v))/2$.

TTGs vs. mLPGs (3)

Proof of Theorem. *Threshold tolerance graphs are mLPGs. (cnt.d)*



$G = mLPG(C, w, T)$ indeed:

for each two nodes u and v in G, in C we have that l_u is connected to $x_{t(u)}$ and l_v to $x_{t(v)}$, where $t(u)$ and $t(v)$ are not necessary distinct.

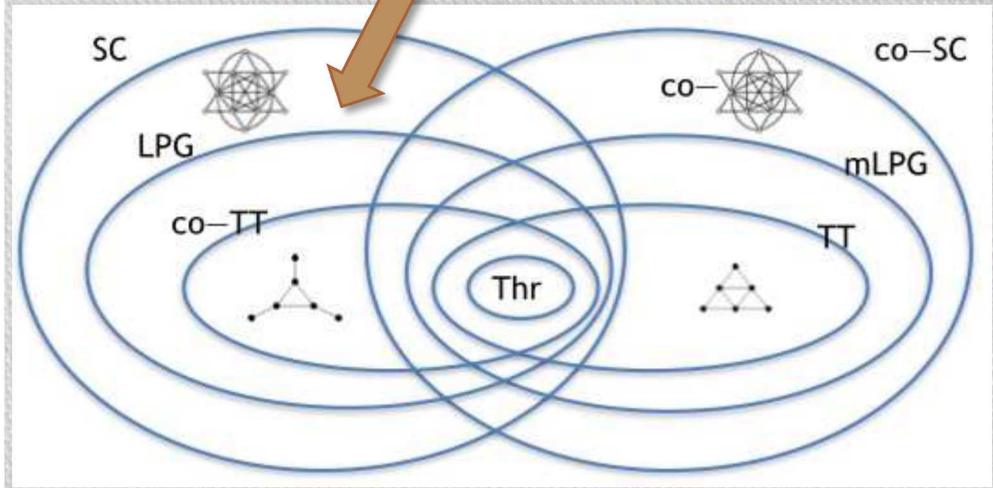
W.l.o.g. $t(v) \geq t(u)$, i.e. $t(u) = \min(t(u), t(v))$.

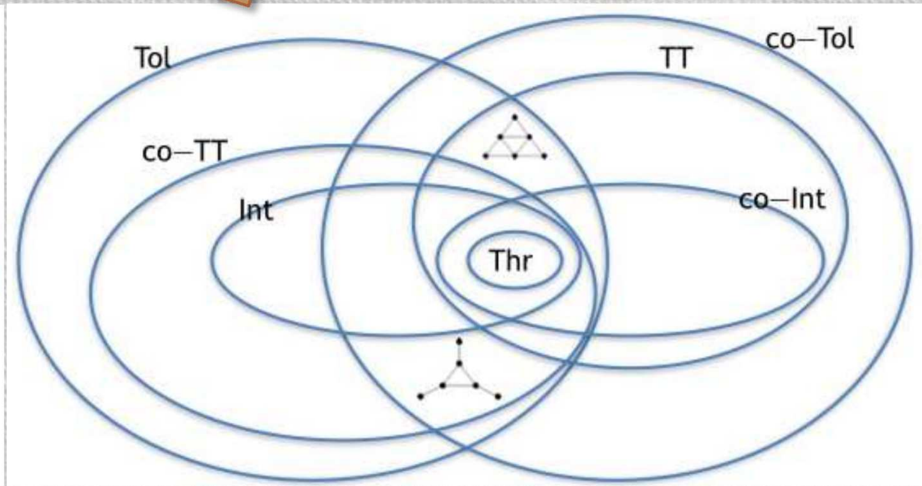
$d_T(l_u, l_v) = w(l_u, x_{t(u)}) + (t(v) - t(u))/2 + w(l_v, x_{t(v)}) = g(u)+g(v)+T-t(u)$.

Clearly, $d_T(l_u, l_v) \geq T$ if and only if $g(u) + g(v) \geq t(u) = \min(t(u), t(v))$

QED

Open problems





- ♣A graph is a **tolerance graph** if to every node *v* can be assigned a closed interval I_v on the real line and a tolerance t_v such that *x* and *y* are adjacent if and only if $|I_x \cap I_y| \geq \min\{t_x, t_y\}$.
- ♣How are related tolerance graphs and leaf power graphs (and, analogously, co-tolerance and min leaf power graphs)?

ANY QUESTION?

THANK YOU!