WHOM TO BEFRIEND TO INFLUENCE PEOPLE

Manuel Lafond (University of Montreal)
Lata Narayanan (Concordia University)
Kangkang Wu (Concordia University)

The plan

Introduction

- MinLinks problem: diffusion by giving links

Hardness of MinLinks

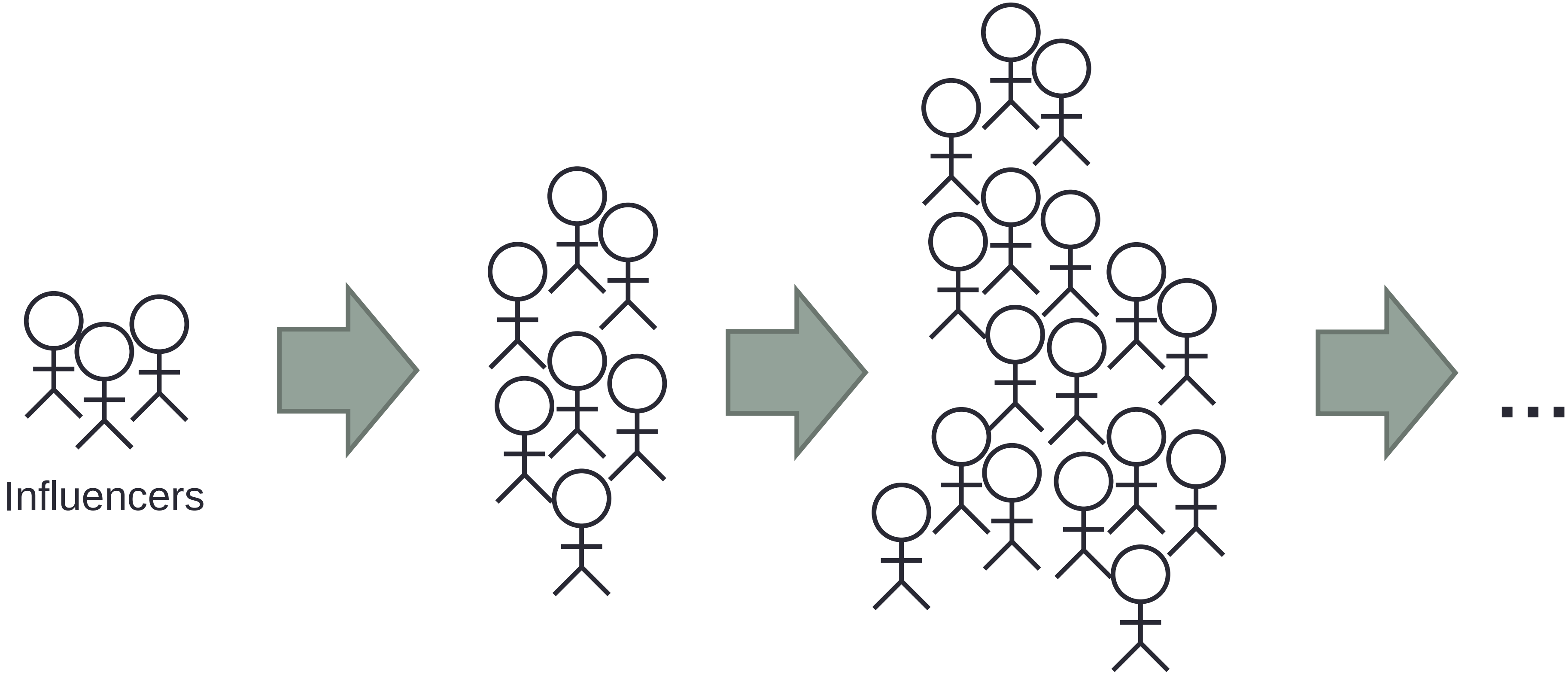
- Basic idea behind the reduction

Some cases we can handle

- Trees, cycles, cliques

Conclusion

Viral marketing



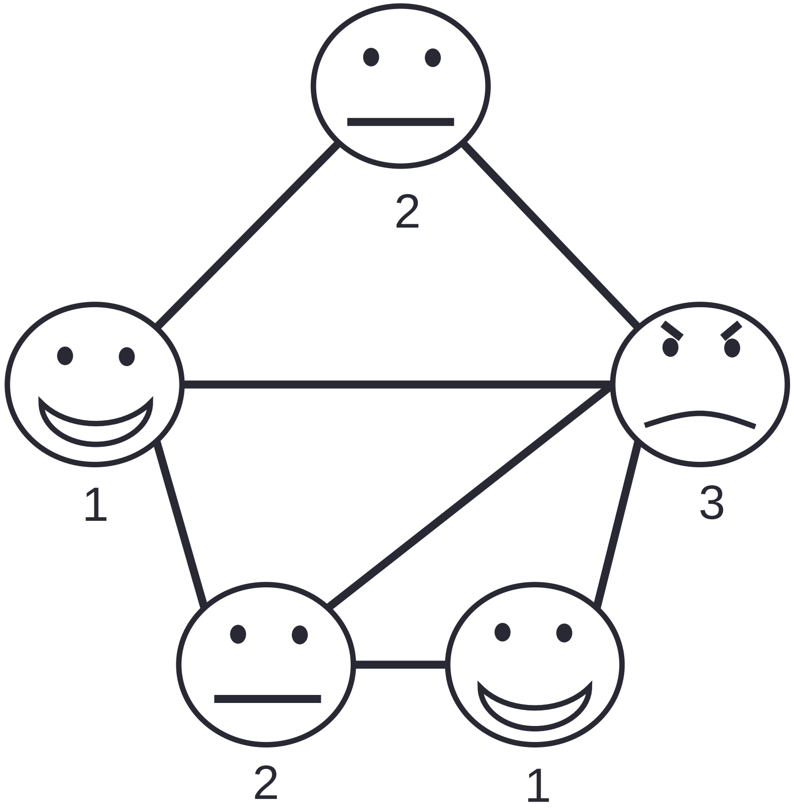
GOAL: make a network adopt a product/idea through the word-of-mouth effect, starting with a **small set of influencers**.

QUESTION: how do we find this small set of influencers?

Alice wants to be loved by all

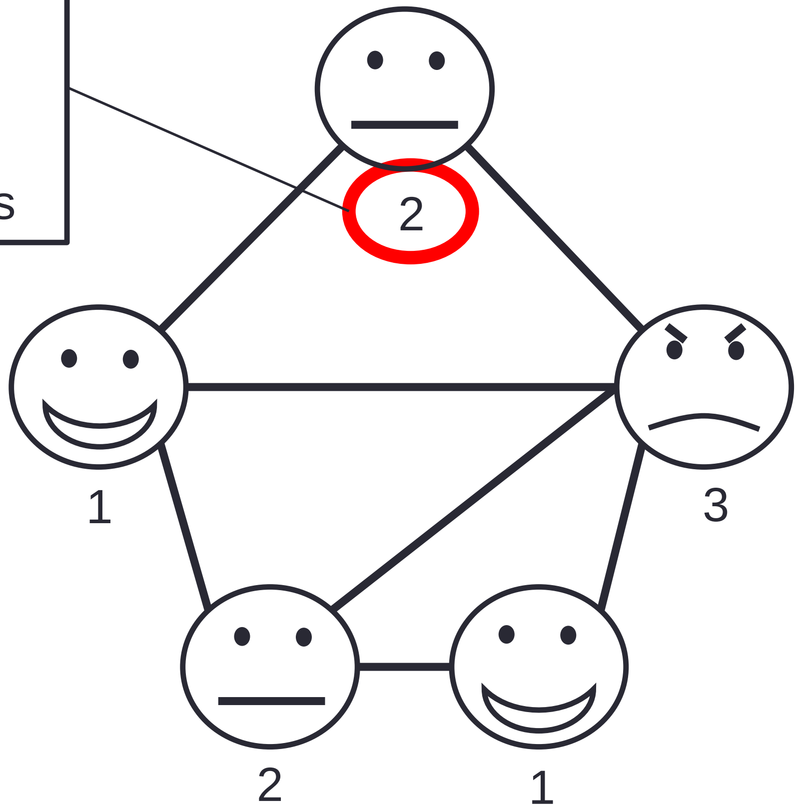
Nodes = people

Edges = friendship links



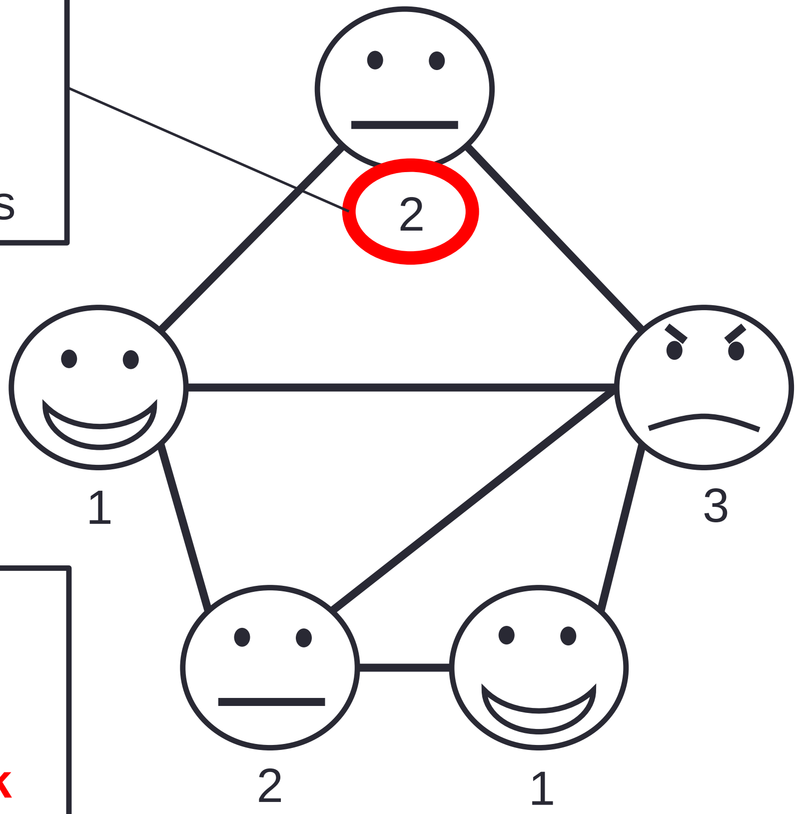
Alice wants to be loved by all

Influence threshold:
once **2 friends** tell this node how great Alice is, it will:
- become **"activated"**
- **tell its friends** how great Alice is



Alice wants to be loved by all

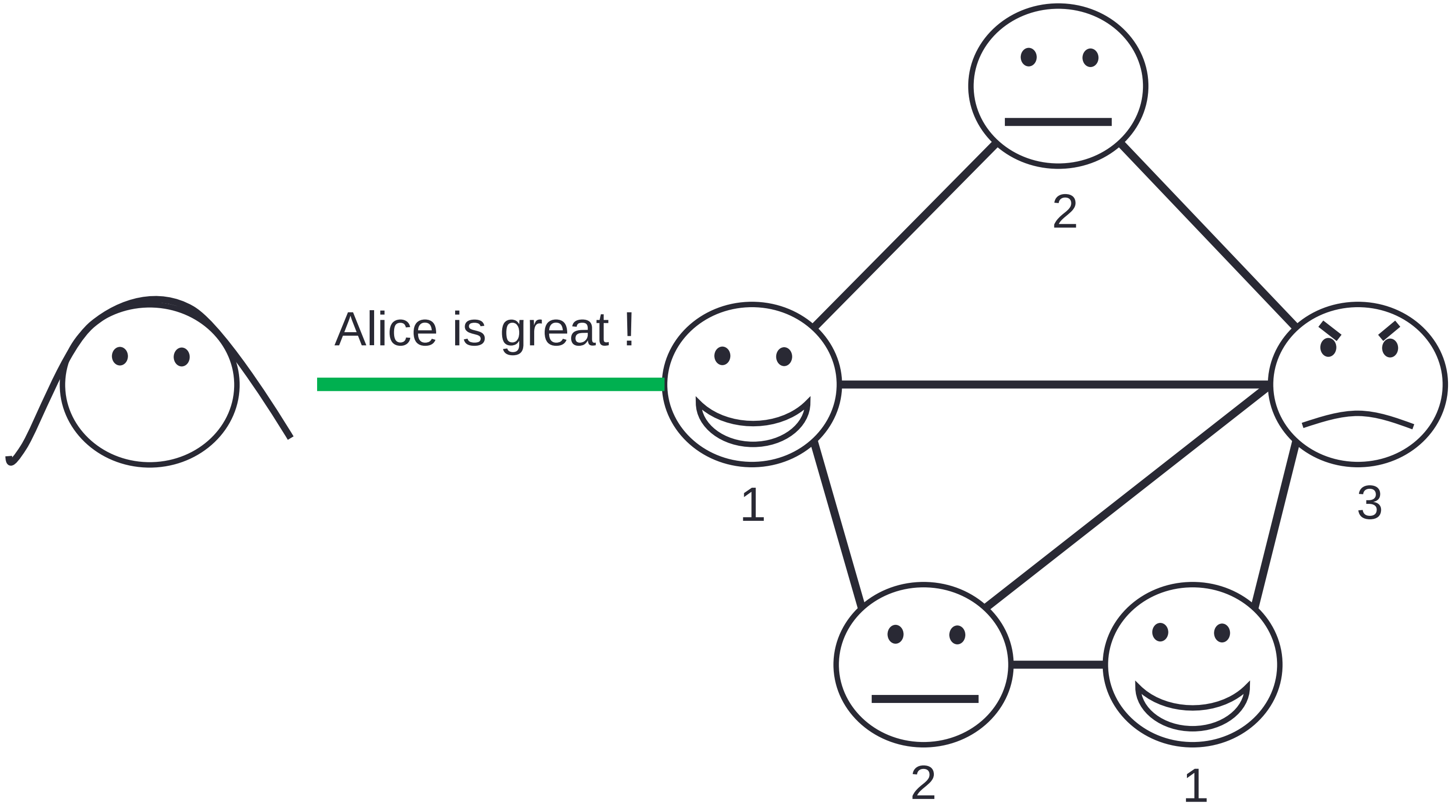
Influence threshold:
once **2 friends** tell this node how great Alice is, it will:
- become **"activated"**
- **tell its friends** how great Alice is



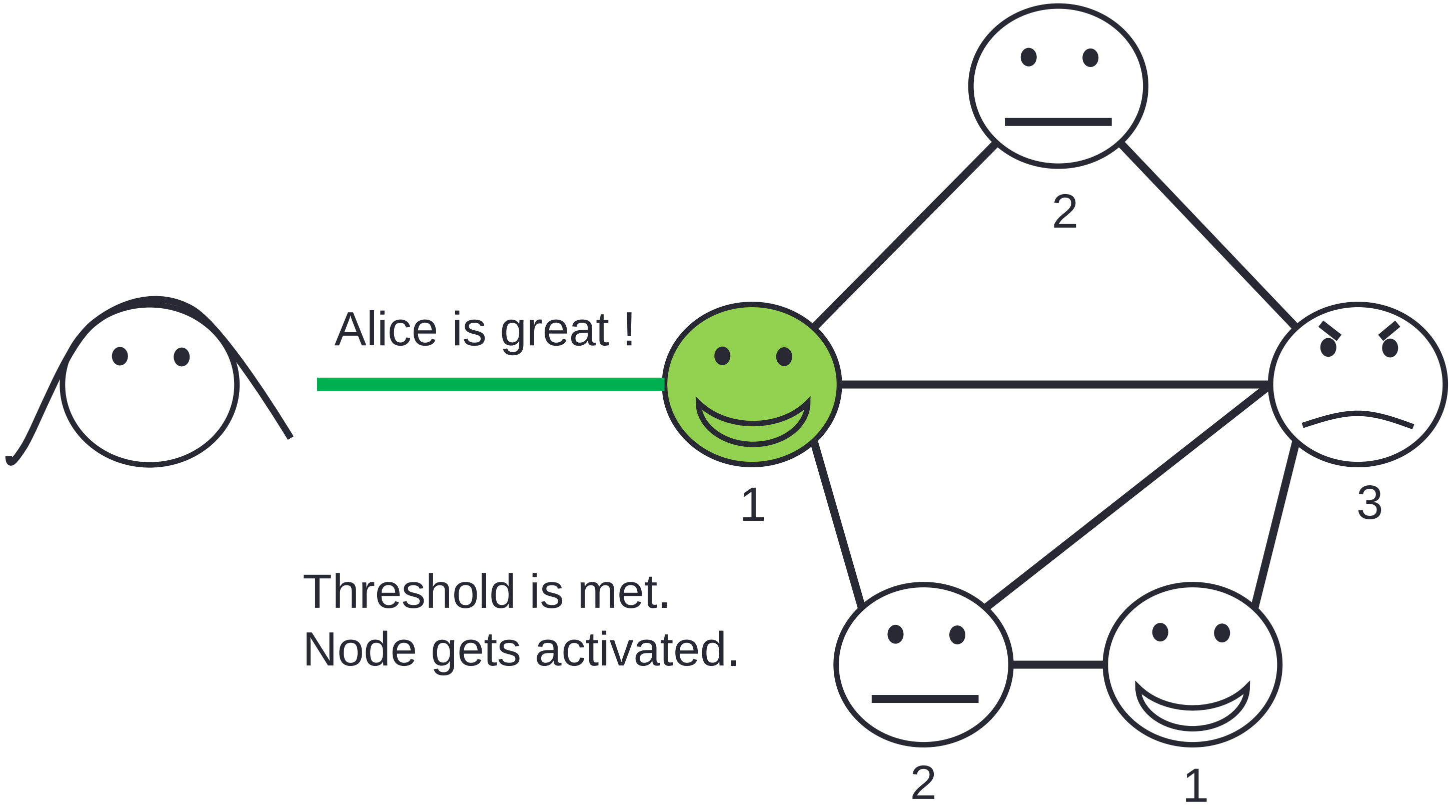
Alice can **buy links** to friends, and tell them how great she is.

GOAL: activate the **entire network** with a minimum number of links.

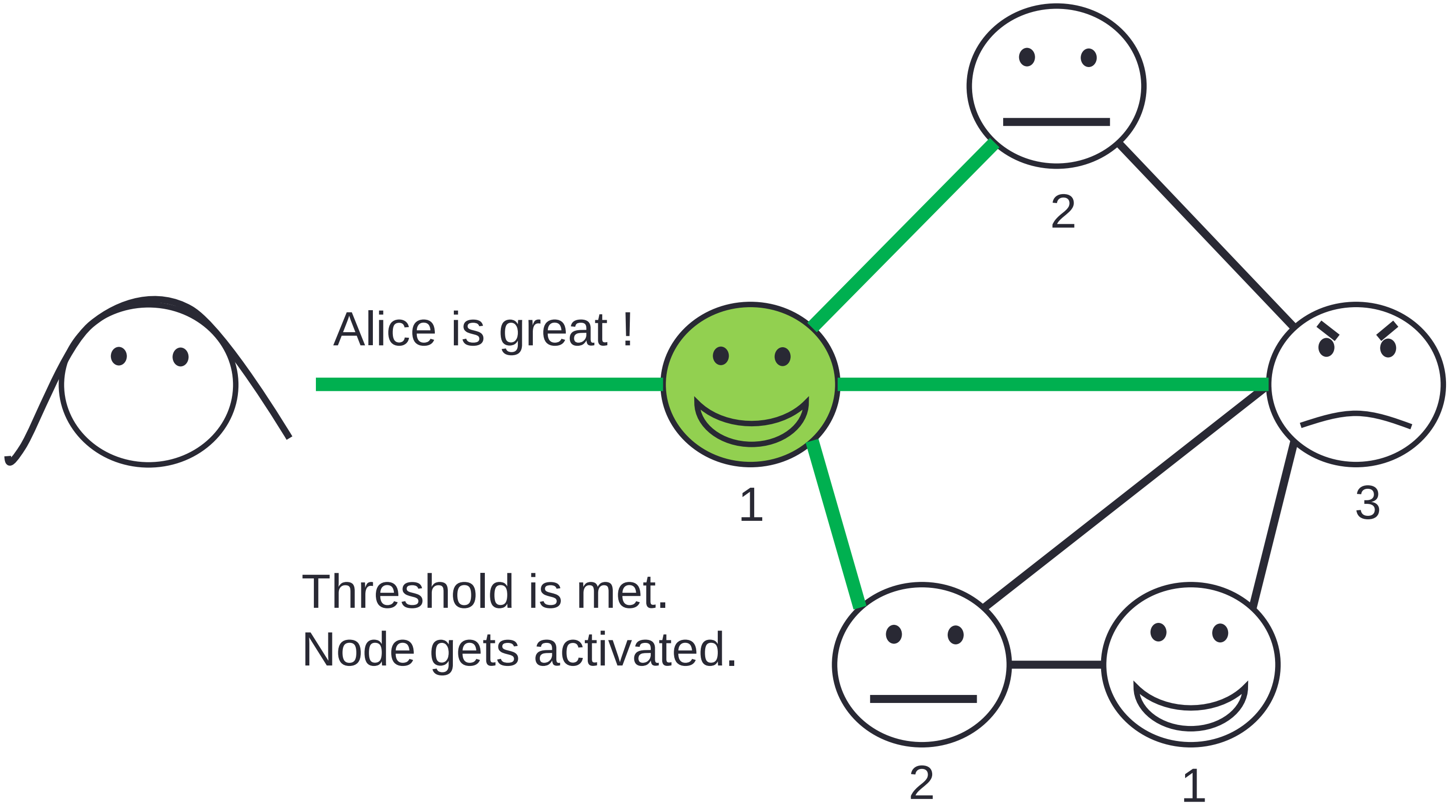
Alice wants to be loved by all



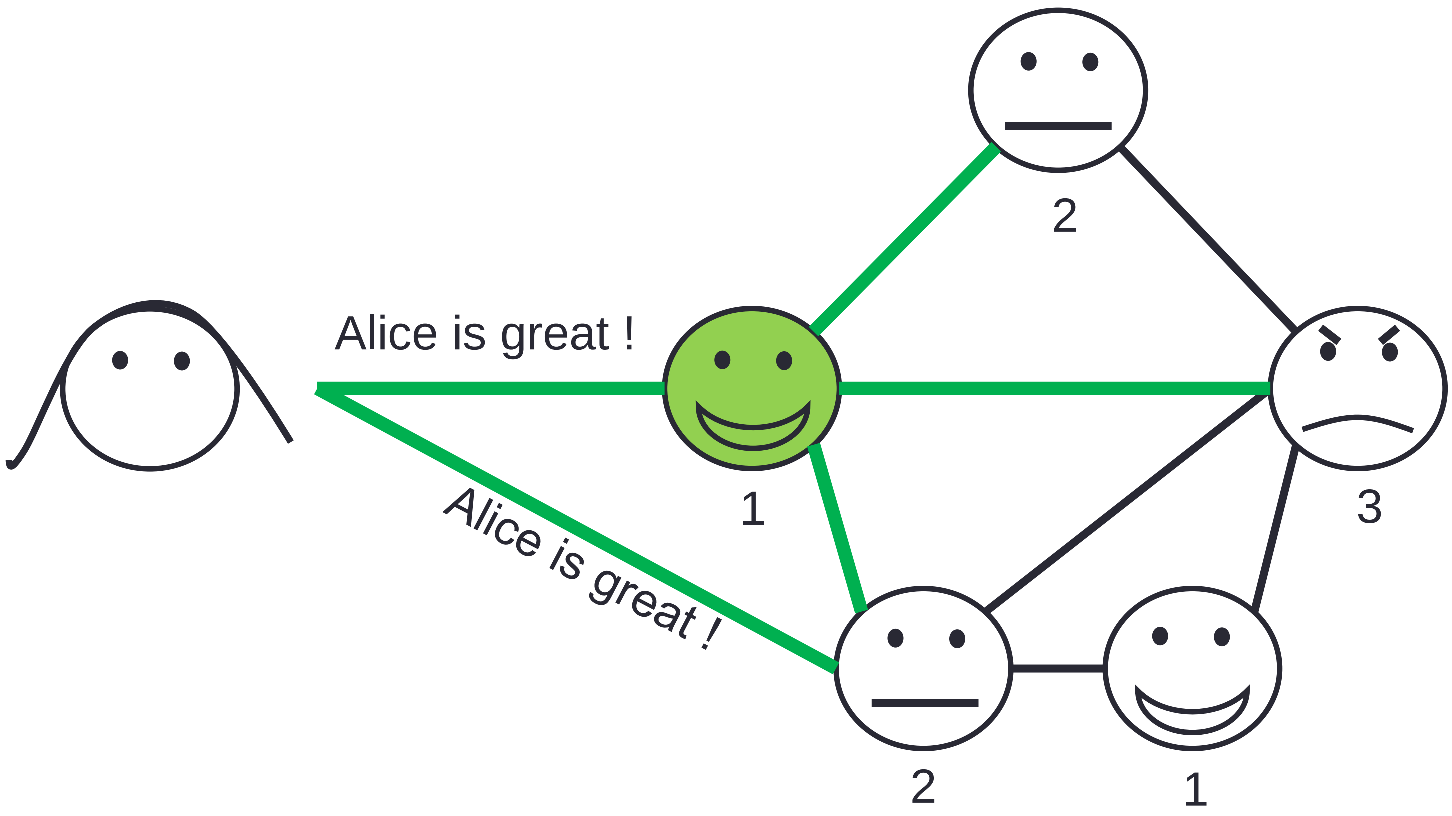
Alice wants to be loved by all



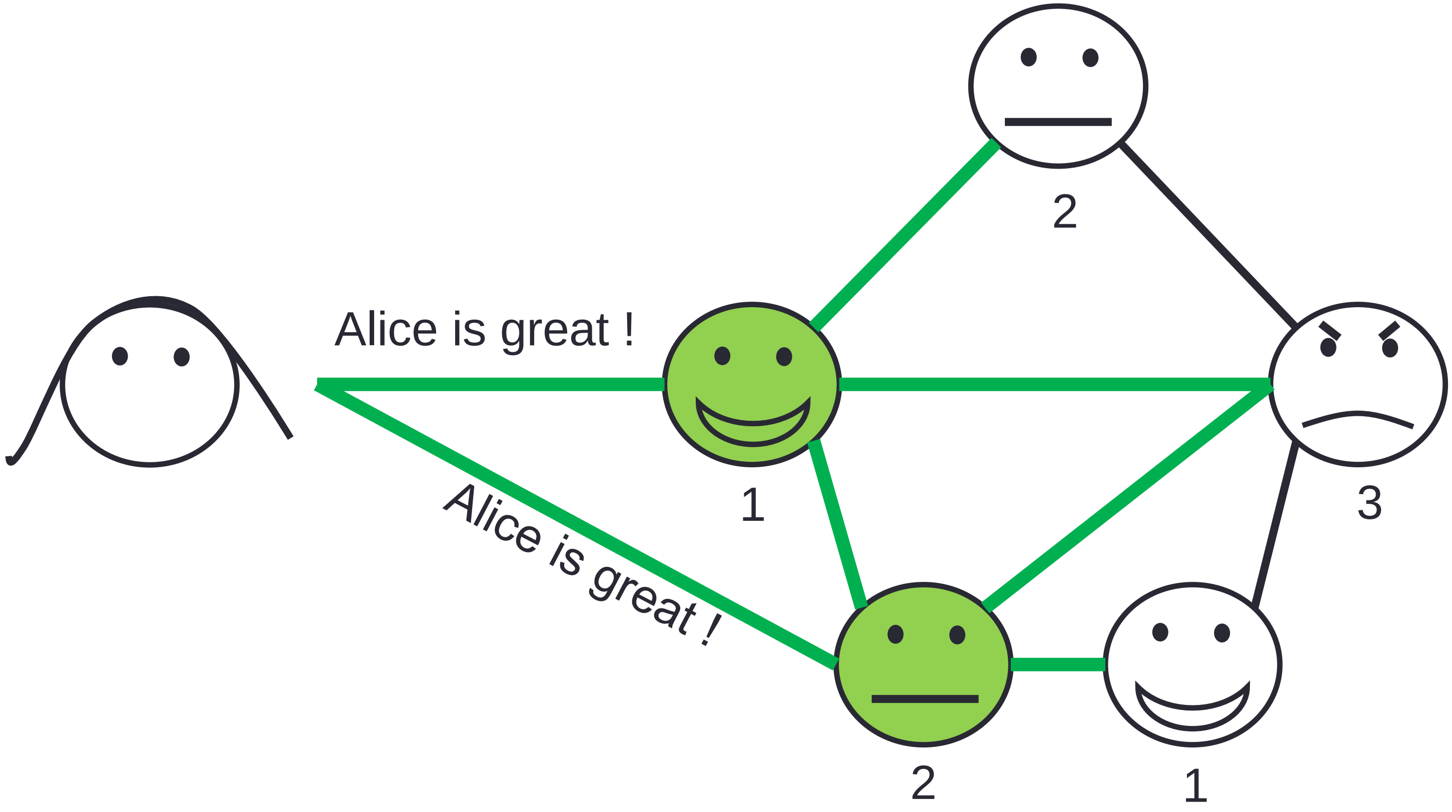
Alice wants to be loved by all



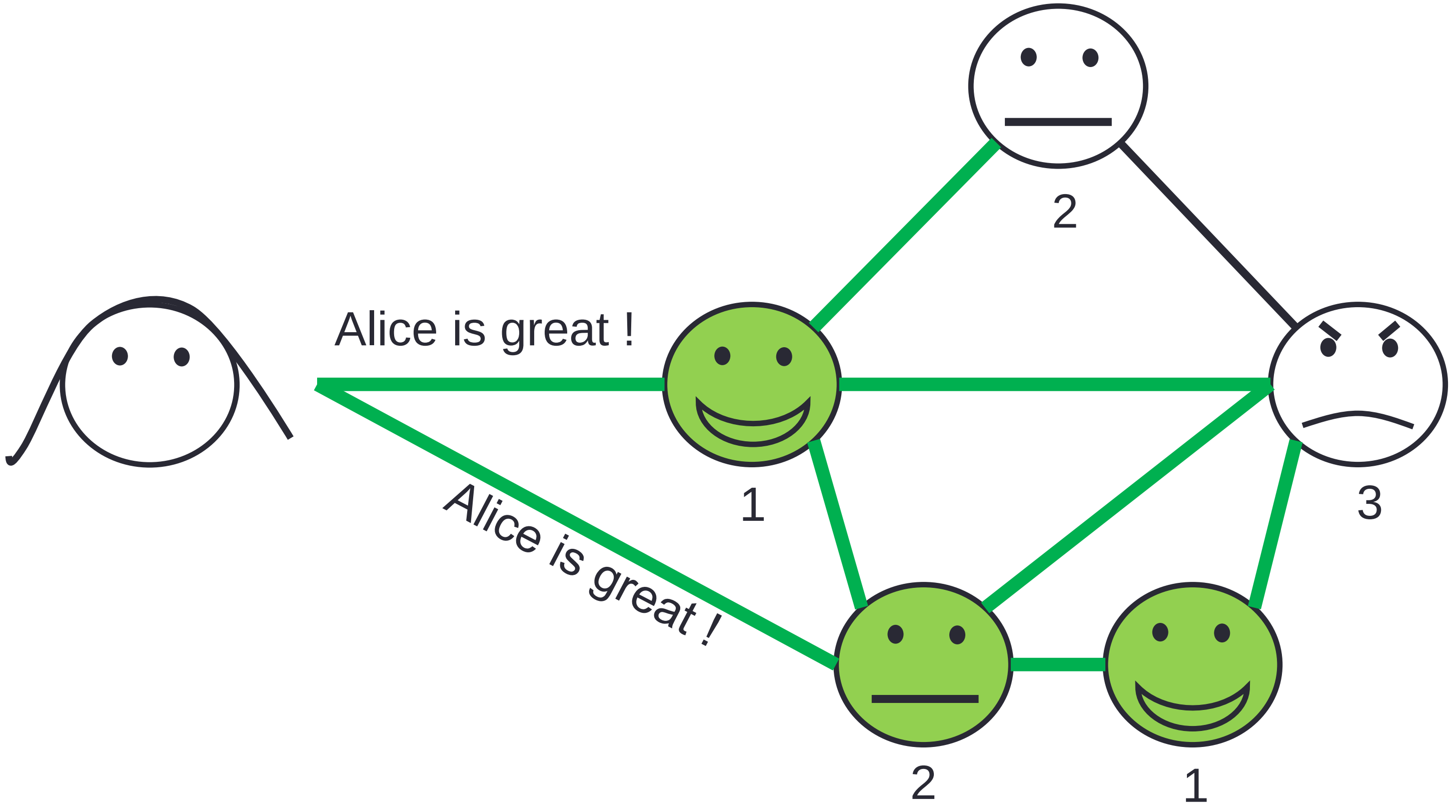
Alice wants to be loved by all



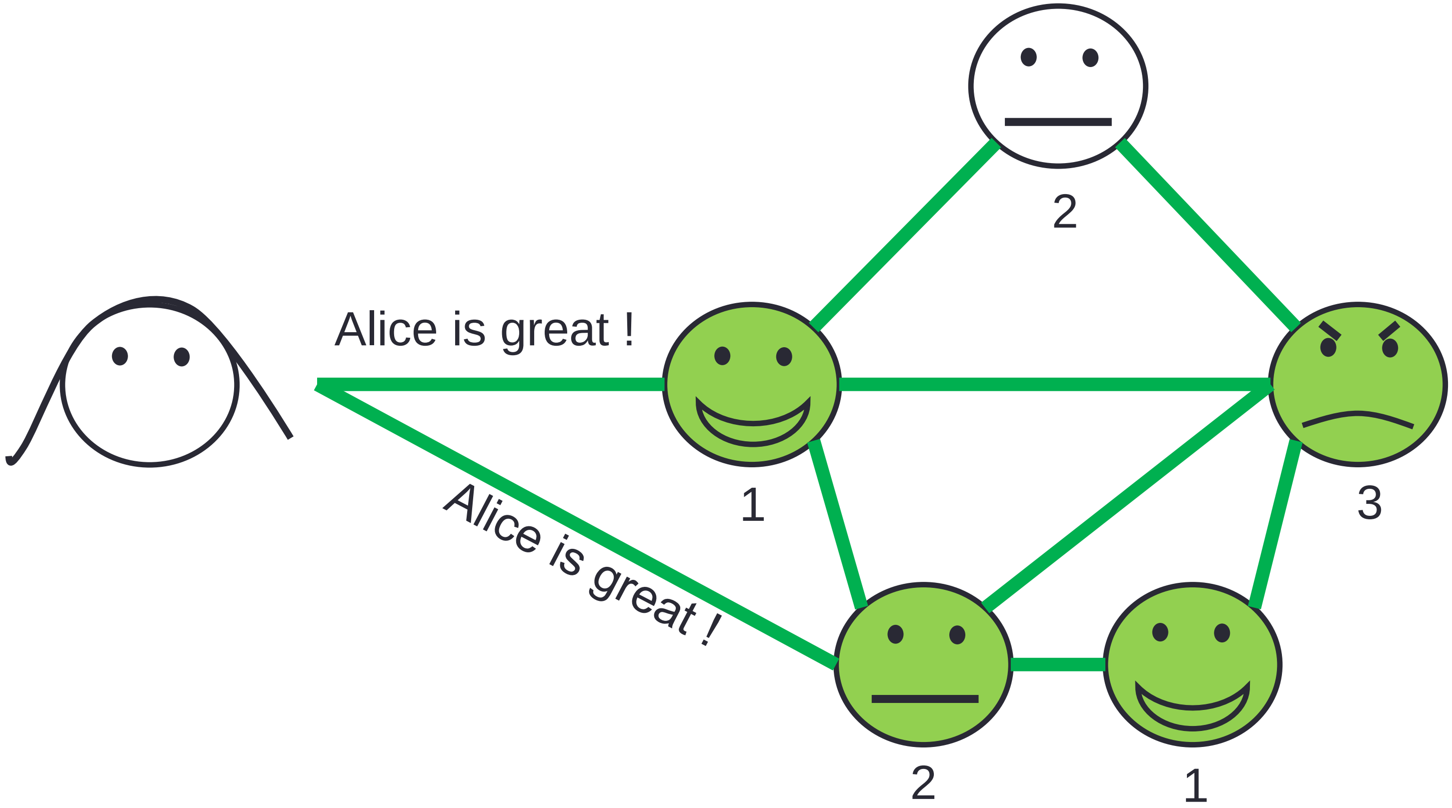
Alice wants to be loved by all



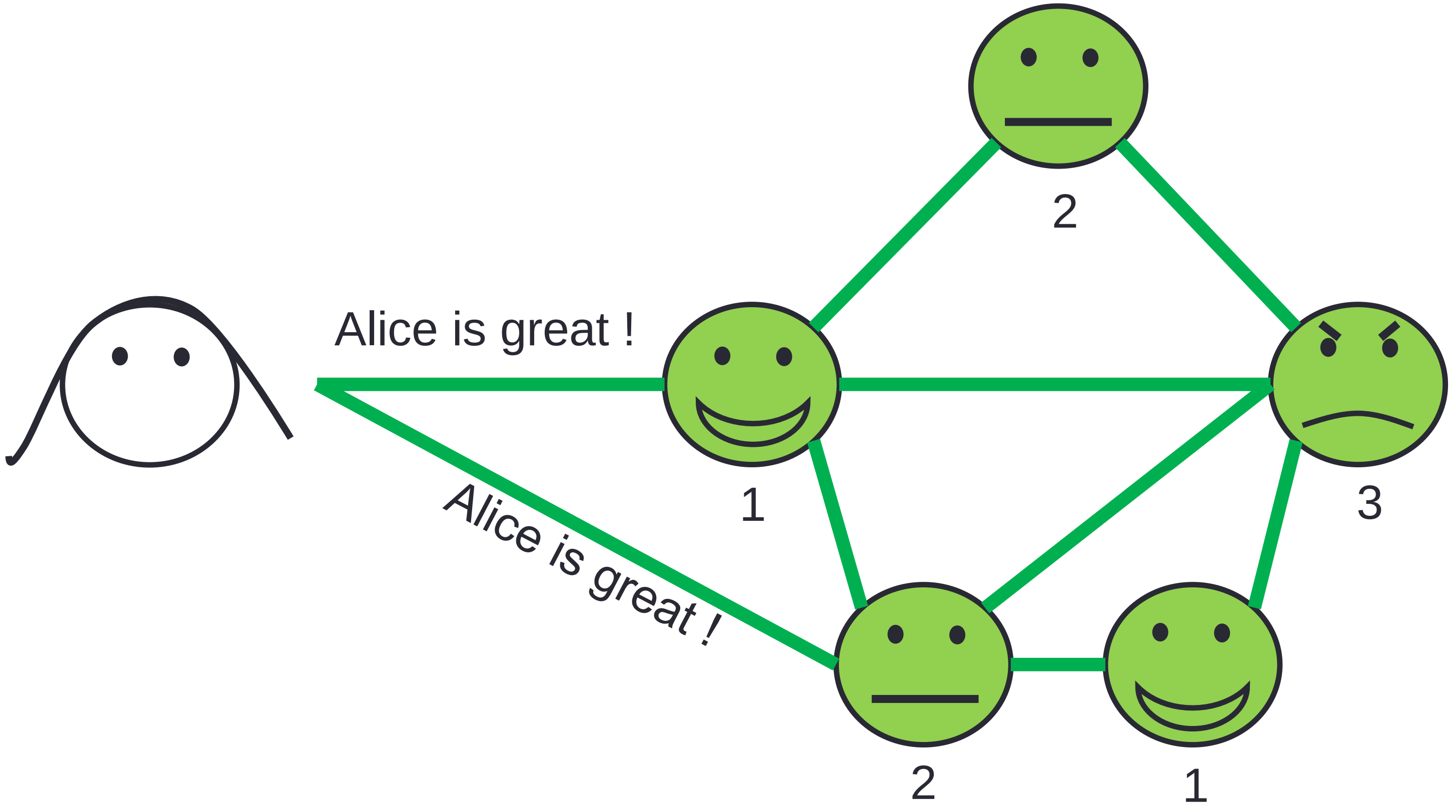
Alice wants to be loved by all



Alice wants to be loved by all



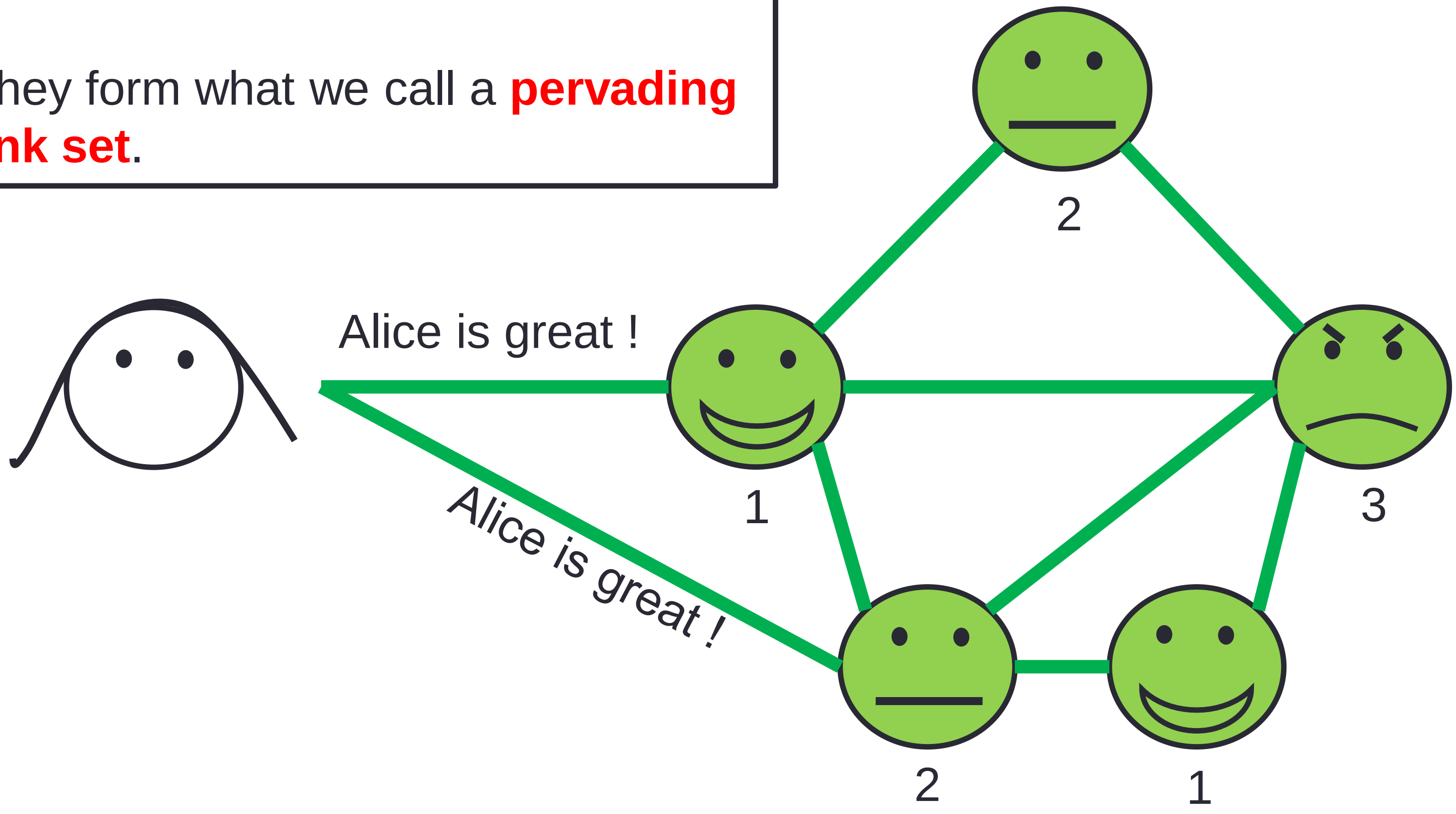
Alice wants to be loved by all



Alice wants to be loved by all

Adding a link to these two individuals activates the whole network.

They form what we call a **pervading link set**.



The model

Each individual has a different given threshold for activation – some people are easily influenced, while others are more resilient to new ideas.

Unit weights assumption: each friend of a given individual has the same influence.

Limited external incentives: the external influencer (i.e. Alice) has the same weight as every individual – Alice can only reduce a person's threshold by 1.

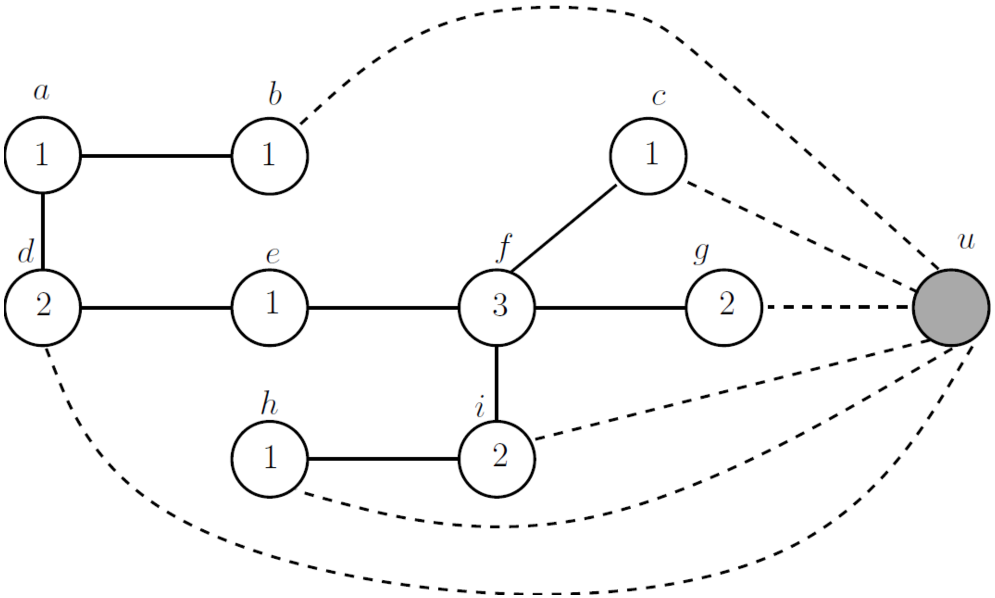
The MinLinks problem

Given: a network G = (V, E, t)

where t(v) is an integer defining the threshold node v

Find: a *pervading link set* S of minimum size

i.e. a subset S of V such that adding a link to each node in S activates the entire network.



Related work

k-target set problem: choose k nodes that influence a maximum number of nodes upon activation.

- Question posed by Domingos and Richardson (2001).
- Modeled as a discrete optimization problem by Kempe, Kleinberg, Tardos (2003).
  - Probabilistic model of thresholds/influence, the goal is to maximize expected number of activated nodes.
  - NP-hard, but admits good constant factor approximation.

Related work

Minimum target set problem: activate a minimum number of nodes to activate the whole network.

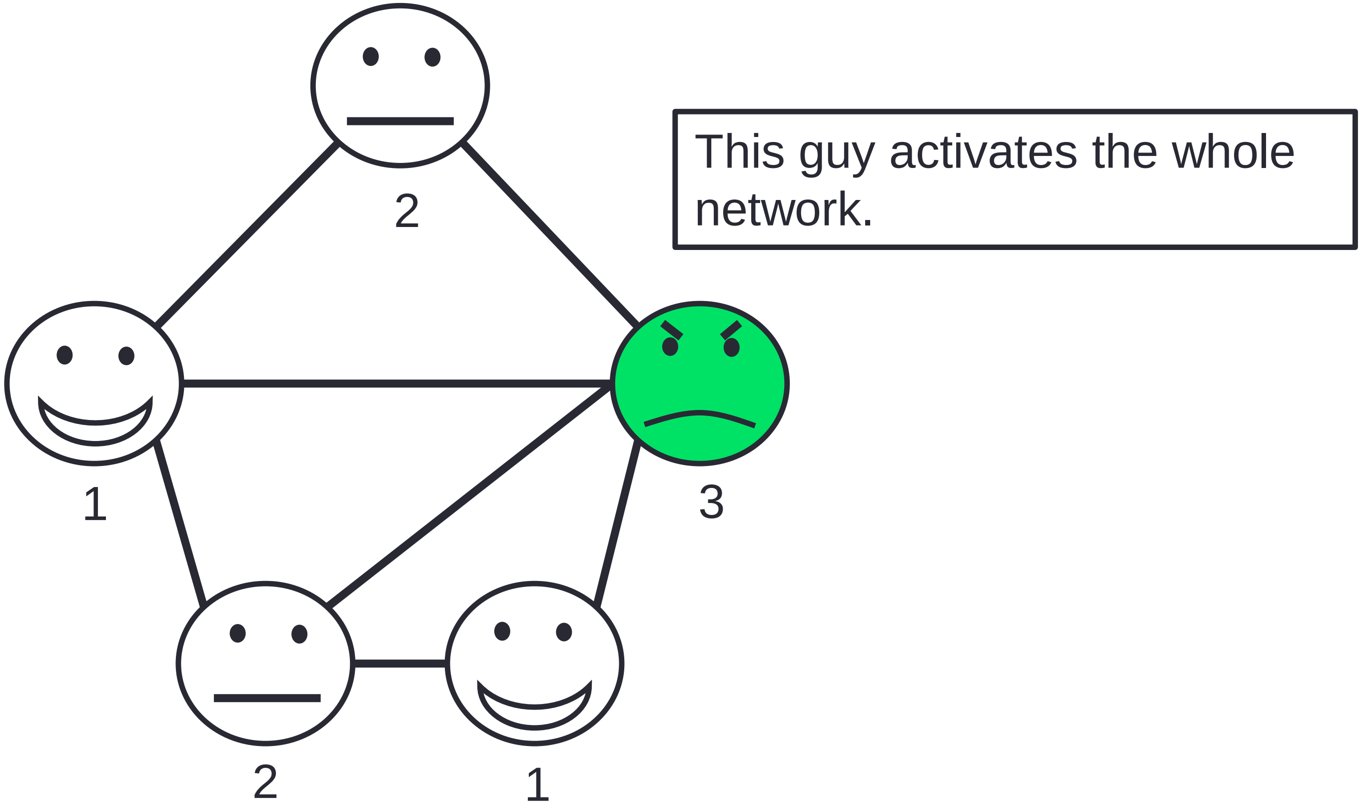
- Hard to approximate within a polylog factor (Chen, 2004).
- Admits FPT algorithm for graphs of bounded treewidth (Ben-Zwi & al., 2014).

Some issues with the minimum target set:

- A node can be activated at cost 1, **regardless of its threshold**. But some people cannot be influenced by external incentives only.
- **No partial incentive** can be given: we either activate a node, or we don't. (In our work, we allow nodes to be activated by a mix of internal and external influences.)

Related work

Minimum target set problem



Related work

Demaine & al. (2015) introduce a model allowing partial incentives.

- Maximization of influence using a fixed budget.
- Thresholds are chosen uniformly at random.
- Any amount of external influence can be applied to a node.

Feasibility of MinLinks

Given a network G, can Alice activate the whole network?

This can easily be checked in polynomial time (actually O(|E(G)|) time):

- give a link to everyone
- propagate the influence
- check if the network is activated

Hardness of MinLinks

Reduction from set cover.

Set cover

Given: a ground set U of size m, and a collection $S = \{S_1, \ldots, S_n\}$ of subsets of U;

Find: a subcollection S' of S such that each element of U is in some set of S'.

Hardness of MinLinks

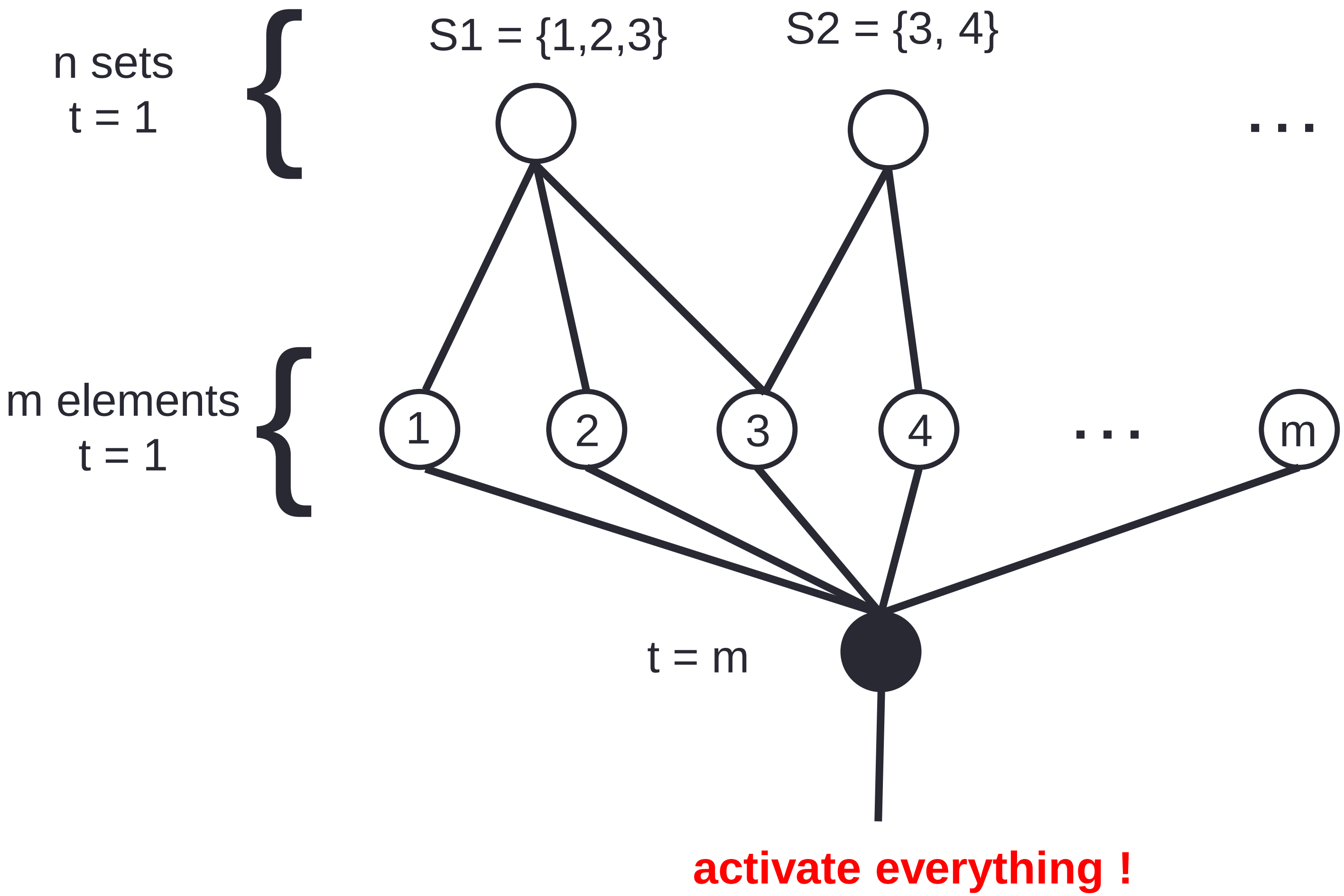
Constructing a MinLinks instance from S:

1) To each element of U corresponds a node of G.
2) To each set S_i corresponds a node of G which activates the elements of S_i.
3) If each element of U gets activated, the whole network gets activated.

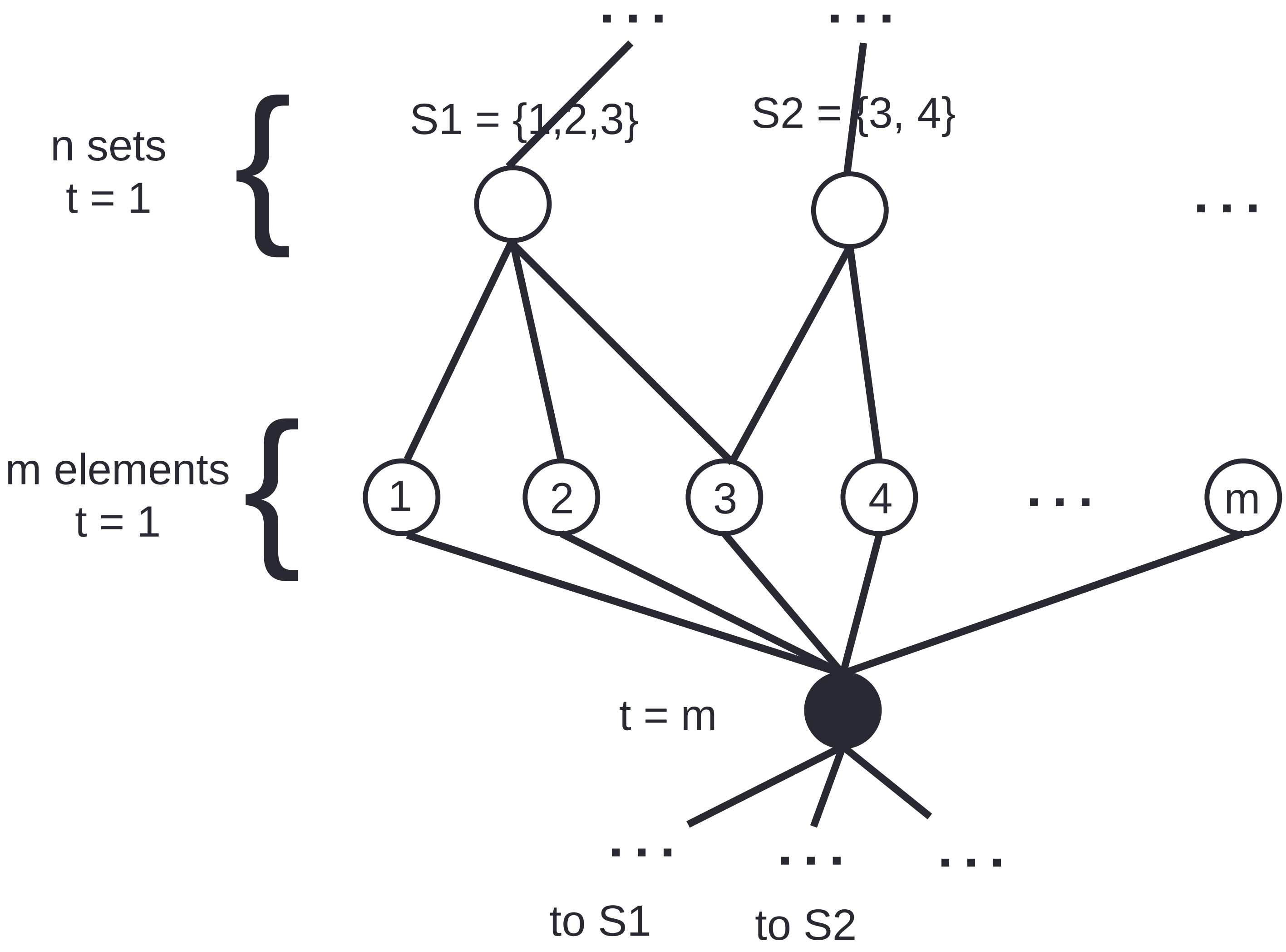
Objective: S has a set cover of size k

iff k links are enough the corresponding MinLinks instance.

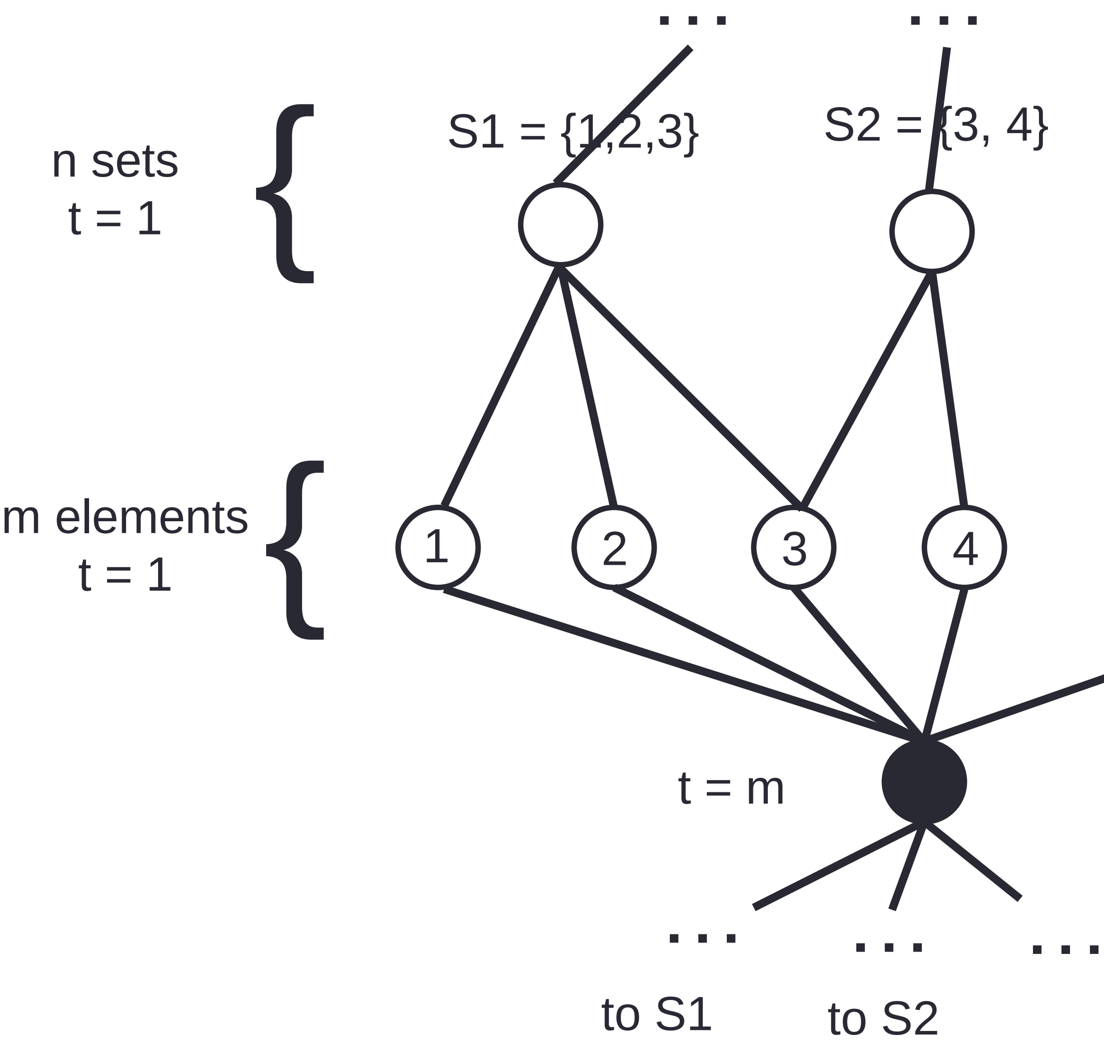
Hardness of MinLinks



Hardness of MinLinks

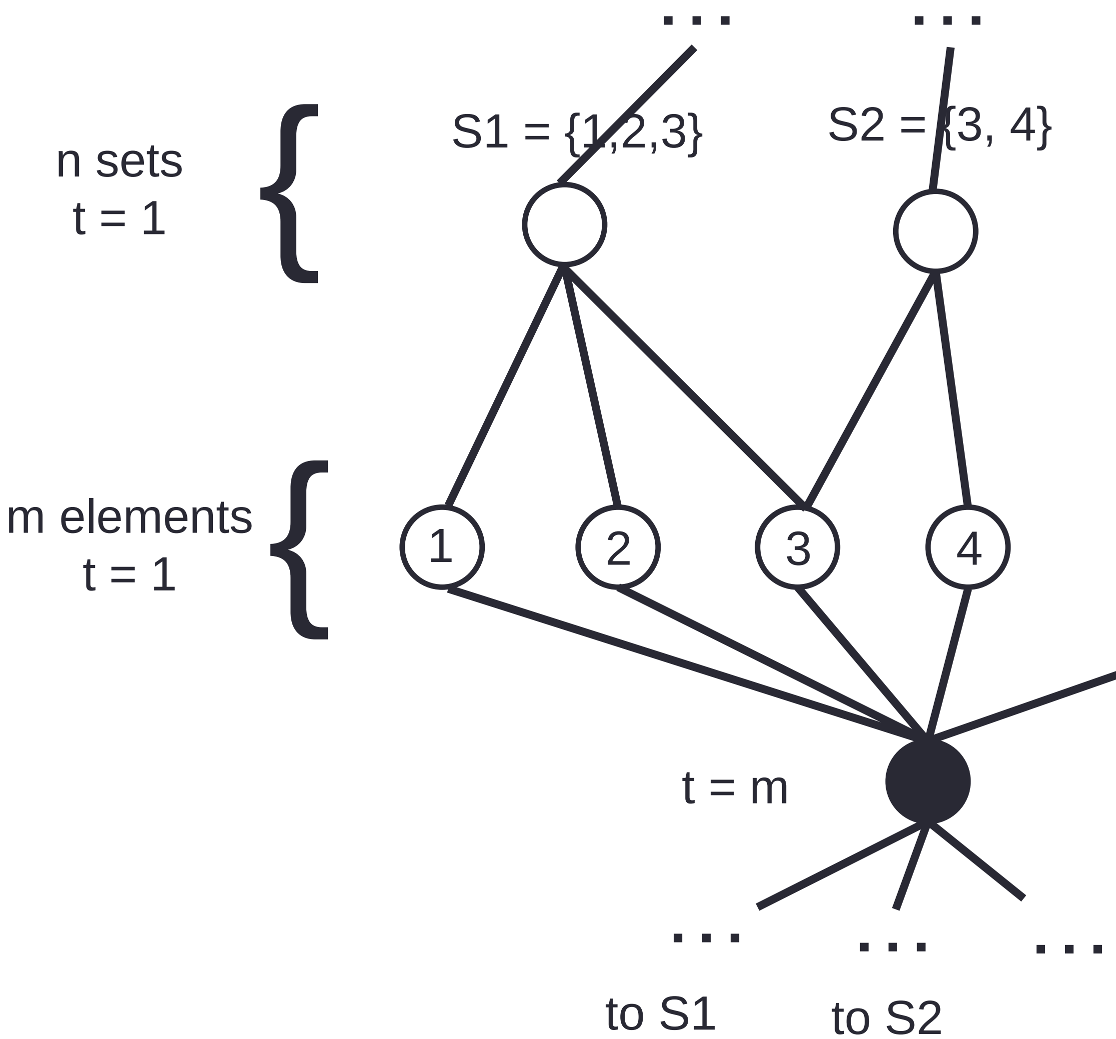


Hardness of MinLinks



If there is a set cover $S_1, \ldots, S_k$ activating the nodes corresponding to $S_1, \ldots, S_k$ activate the m elements, then the supernode, then everything!

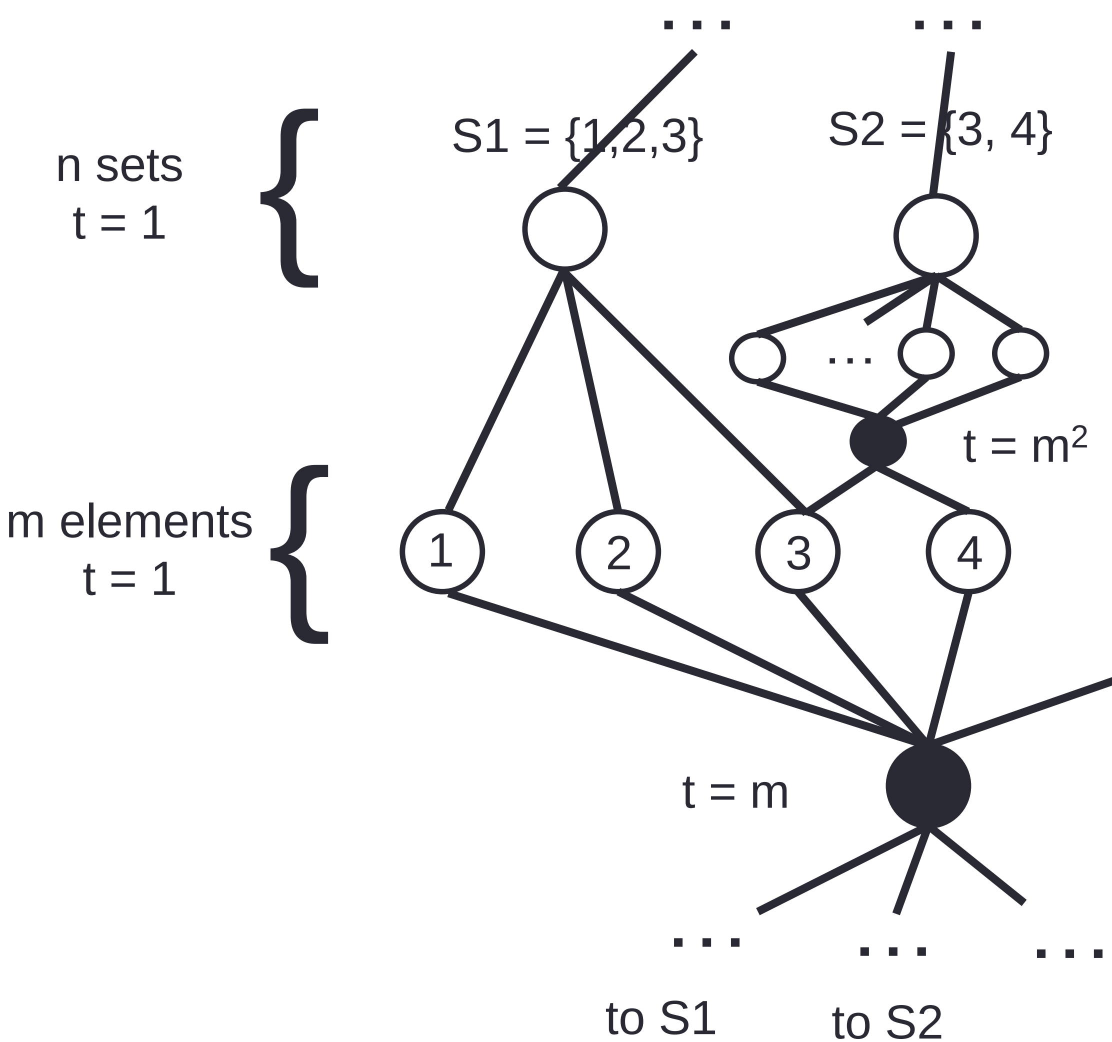
Hardness of MinLinks



If there is a set cover $S_1, \ldots, S_k$ activating the nodes corresponding to $S_1, \ldots, S_k$ activate the m elements, then the supernode, then everything!

Problem: S1 activates 3, then 3 activates S2

Hardness of MinLinks

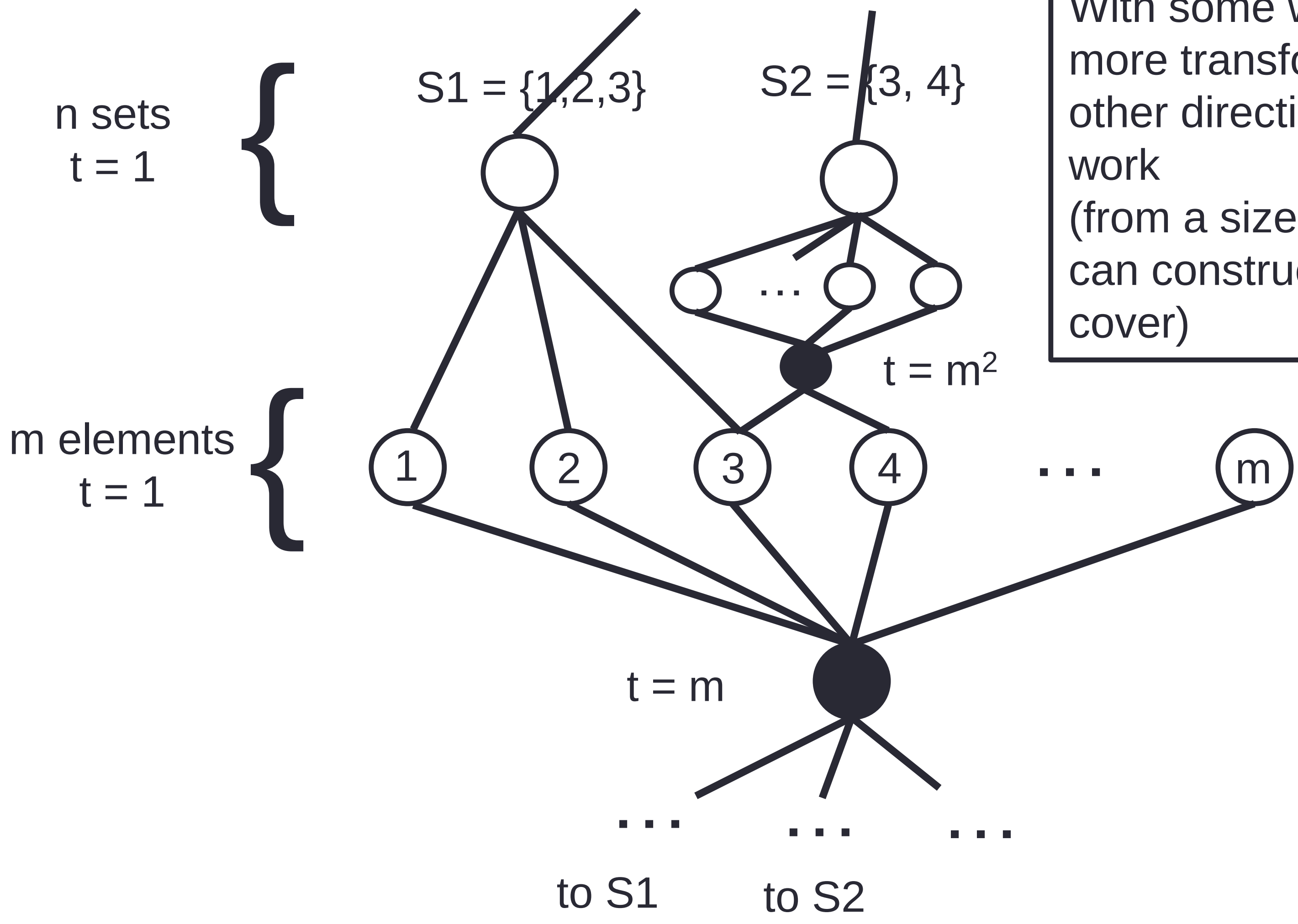


If there is a set cover
$S_1, \ldots, S_k$
activating the nodes corresponding to $S_1, \ldots, S_k$ activate the m elements, then the supernode, then everything!

Problem: S1 activates 3, then 3 activates S2.

Add a blocking node of high threshold (say m^2). Now 3 cannot activate S2 by itself. Repeat for every set element.

Hardness of MinLinks



With some work, and some more transformations, the other direction can be made to work
(from a size k MinLinks we can construct a size k set cover)

Hardness of MinLinks

This reduction can be extended to show the hardness of this restricted version of MinLinks:

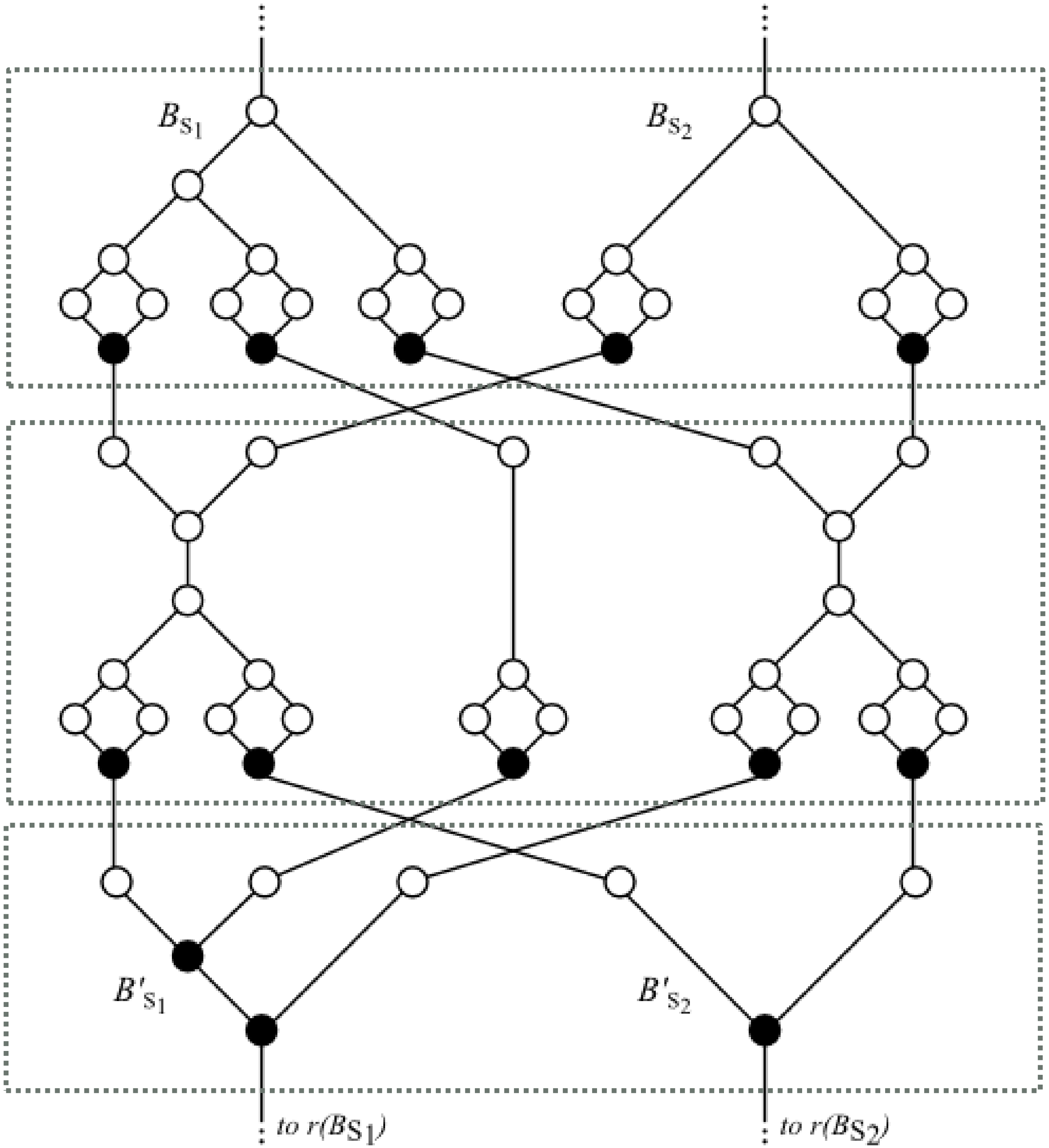
- all nodes have threshold 1 or 2
- each node has at most 3 neighbors

Since SetCover is O(log n)-hard to approximate, and we preserve the instance size and optimality value, MinLinks is also O(log n)-hard to approximate.

Set nodes => binary trees

Element nodes => binary trees

Super activator => binary trees



Trees

Theorem: if T has a pervading link set, then

for any node v, there is an optimal solution in which **v gets a link**.

Trees

Theorem: if T has a pervading link set, then

for any node v, there is an optimal solution in which **v gets a link**.

Leads to the following O(n) algorithm:

- Give a link to any leaf v of T
- If v gets activated, propagate its influence
- Remove v and all activated nodes
- Repeat on each component

Trees

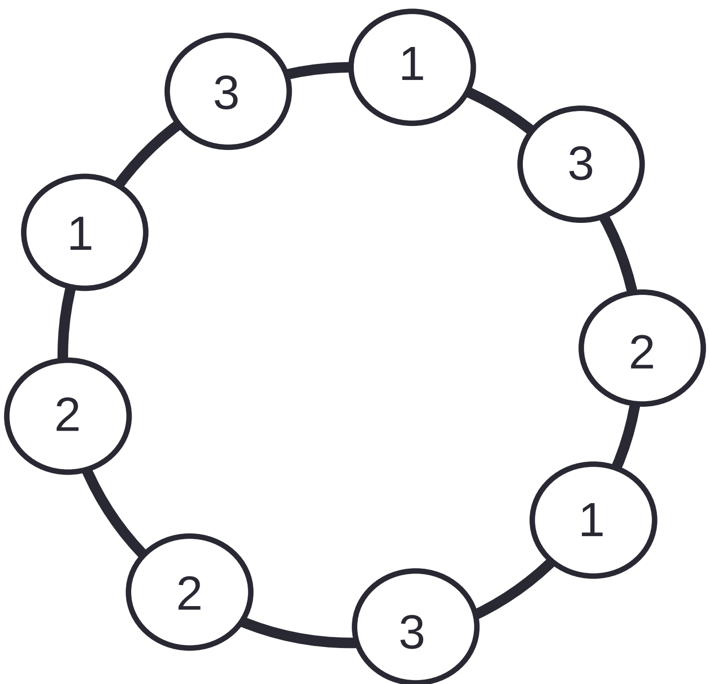
Theorem: The minimum number of links required to activate a tree T is

$$ML(T) = 1 + \sum_{v \text{ in } V(T)} (t(v) - 1)$$

Cycles

Lemma: a cycle has a pervading link set iff

- $t(v) \leq 3$ for every node
- there is at least one node of threshold 1
- between any two consecutive nodes of threshold 3, there is at least one node of threshold 1



Cycles

Lemma: a cycle has a pervading link set iff

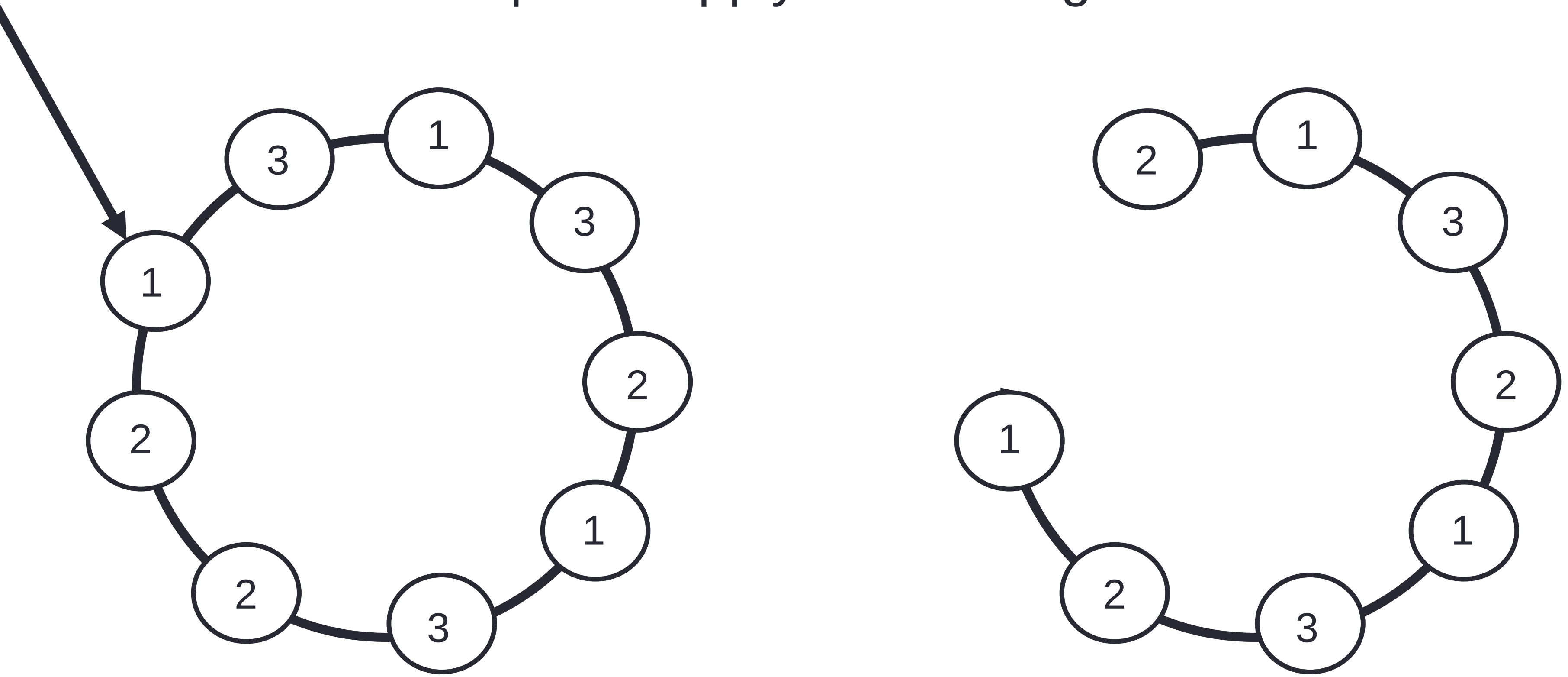
- $t(v) \leq 3$ for every node
- there is at least one node of threshold 1
- between any two consecutive nodes of threshold 3, there is at least one node of threshold 1

Theorem: if a cycle C has a pervading link set, then for any node v of threshold 1, there is an optimal solution in which v gets a link.

Cycles

Leads to the following O(n) algorithm:

- Give a link to a threshold 1 node
- Propagate the influence and remove activated nodes
- The result is a path. Apply the tree algorithm on it.



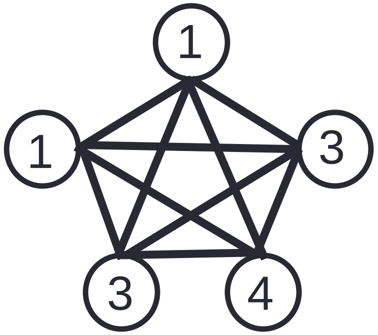
Cycles

Theorem: The minimum number of links required to activate a cycle C is

$$ML(C) = \max(1, \sum_{v \text{ in } V(T)} (t(v) - 1))$$

Cliques

Theorem: suppose that G is a clique. Order the vertices (v_1, v_2, …, v_n) by threshold in increasing order. Then G has a pervading link set iff $t(v_i) \leq i$ for each i.

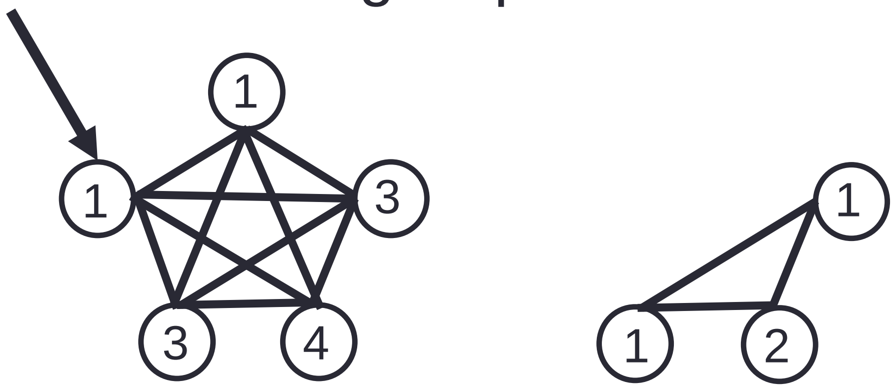


Cliques

Theorem: suppose that G is a clique. Order the vertices (v_1, v_2, …, v_n) by threshold in increasing order. Then G has a pervading link set iff $t(v_i) \leq i$ for each i.

Leads to the following O(n) algorithm:

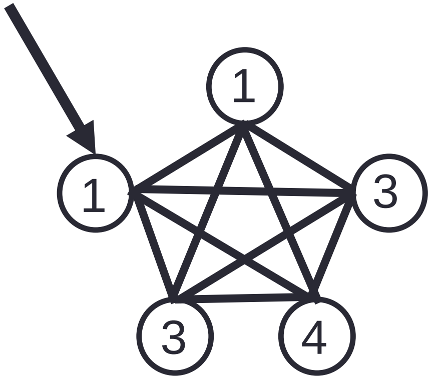
- Give any threshold 1 node a link.
- Propagate the influence, remove activated nodes.
- Repeat on the resulting clique until nothing remains.

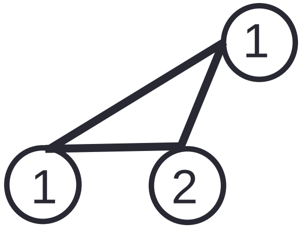


Cliques

Theorem: suppose that G is a clique. Order the vertices (v_1, v_2, …, v_n) by threshold in increasing order.

Then the minimum number of links required to activate G is equal to the number of v_i such that $t(v_i) = i$.





Conclusion

What if we allow multiple influencers?

- Allows buying multiple links to a single node
- Given k influencers, how many links are needed?

What if the "link budget" is limited?

- Maximize the number of activated nodes by using k links.
- NP-Hardness follows from our results, but approximability/FPT is unknown.