Open problems in complex networks analysis and modelling

Christophe Crespelle

christophe.crespelle@univ-cotedazur.fr

UNIVERSITÉ CÔTE D'AZUR

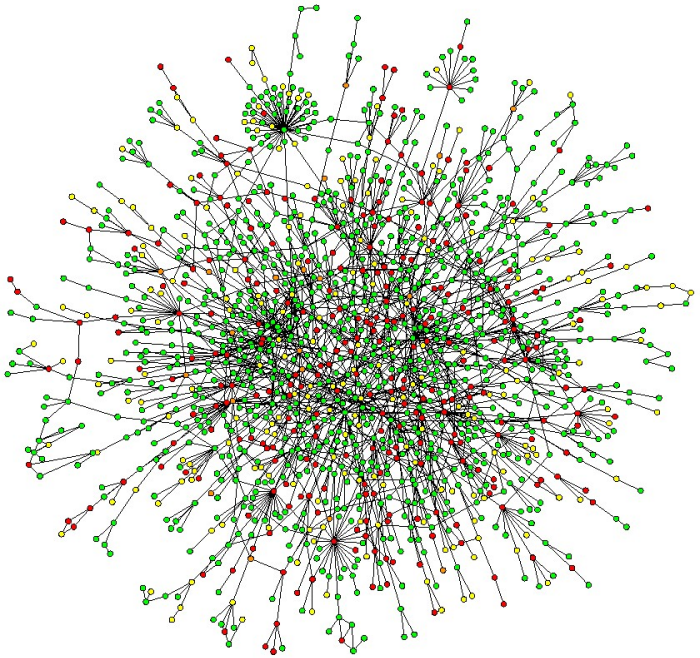
i3S sophia antipolis

DIGITAL SYSTEMS FOR HUMANS
GRADUATE SCHOOL AND RESEARCH

Complex networks

- Real-world data

  Ex of contexts :
  computer science,
  social sciences,
  biology, linguistics,
  medecine,
  transportation,
  communications,
  industry, economy, ...



complex

=

large

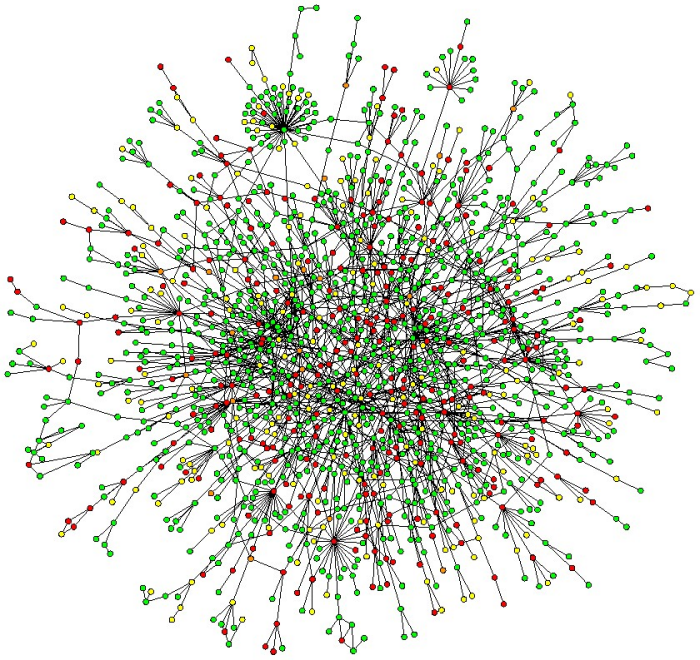
+

unordered

Complex networks

- Real-world data

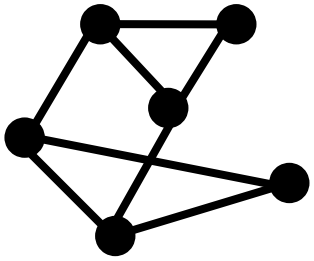
  Ex of contexts :
  computer science,
  social sciences,
  biology, linguistics,
  medecine,
  transportation,
  communications,
  industry, economy, ...



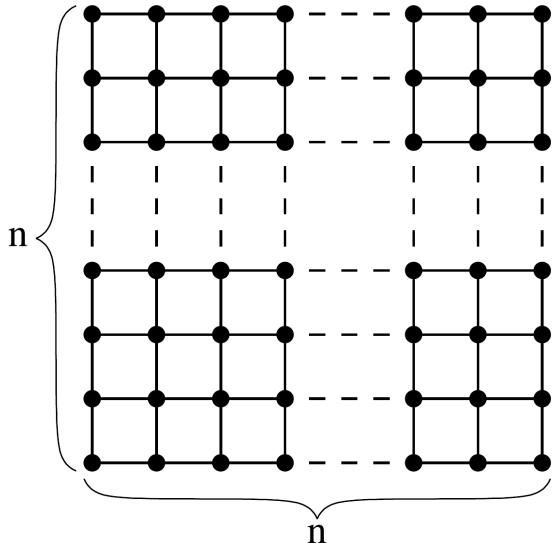
complex
=
large
+
unordered

- Not complex

small



ordered

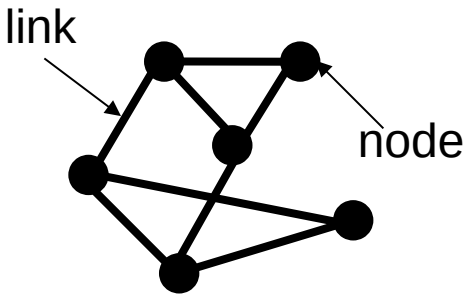


Complex networks

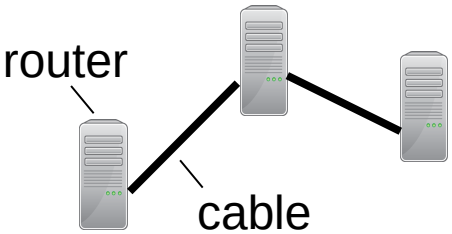
- Real-world data *(not formally defined)*

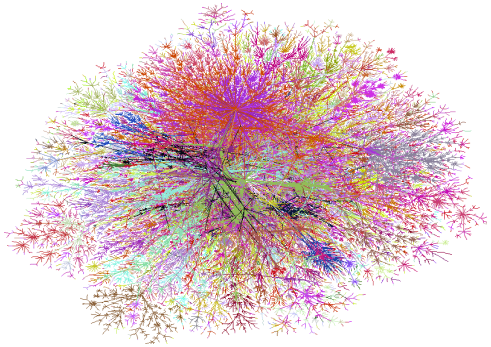
  Ex of contexts :

  computer science, social sciences, biology, linguistics, medecine, Transportation, communications, industry, economy, ...



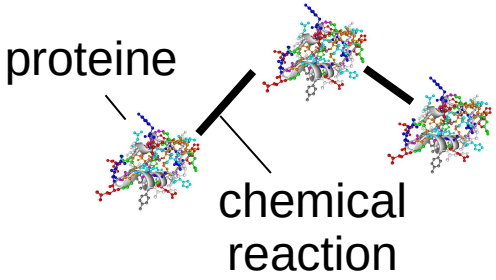
Internet

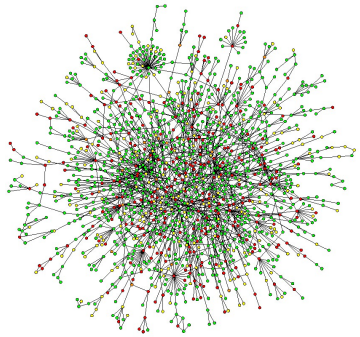




How to carry information across the Internet?

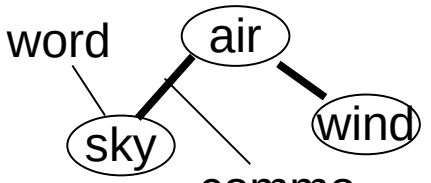
Proteine interactions

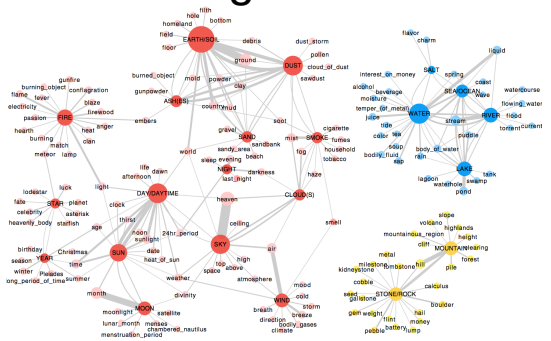




How does a living cell work?

Word networks





How does a language evolve?

Four big classes of problems

- Measurement
- Analysis
- Modelling
- Algorithms

Analysis

Communities in complex networks

What is a community ?

"Moral" definition

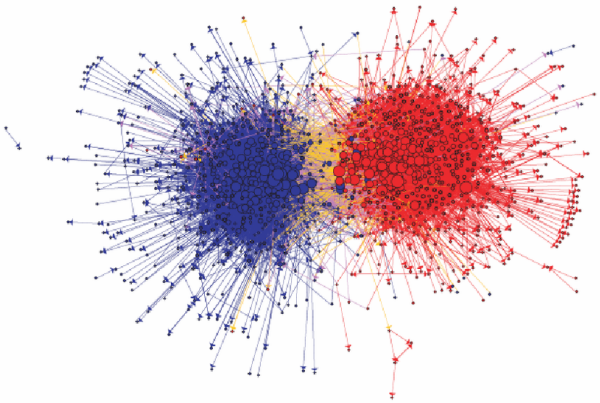
- A group of nodes that share something...
  - People with a common interest
  - Web pages with similar content
  - Proteins realising a common function

Communities in complex networks

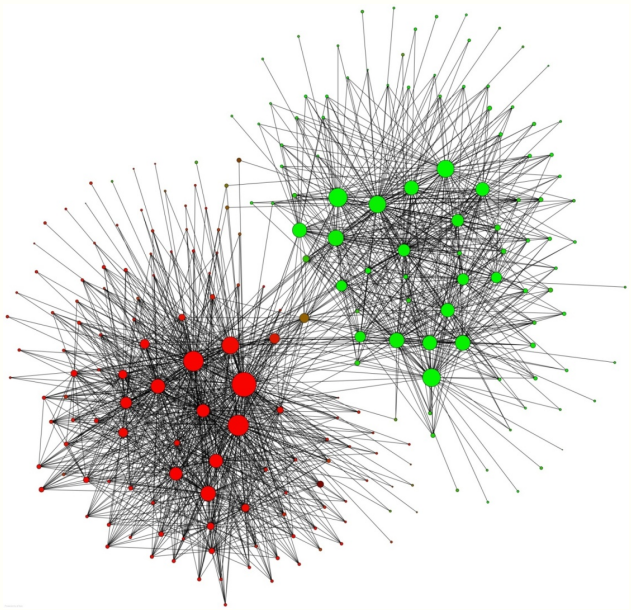
What is a community ?

"Moral" definition

- A group of nodes that share something...
  - People with a common interest
  - Web pages with similar content
  - Proteins realising a common function
- ... that makes them be in relationship in the network !



Political blogs in US



Languages in Belgium

Communities in complex networks

What is a community ?

Structural definition

- A highly connected group of nodes

Communities in complex networks

What is a community ?

Structural definition

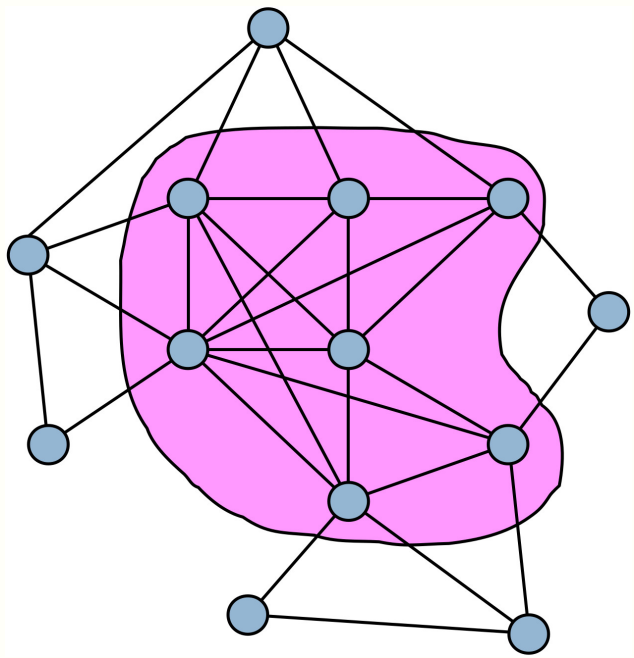
- A highly connected group of nodes
  - Density inside the community much higher than global density of the network

Communities in complex networks

What is a community ?

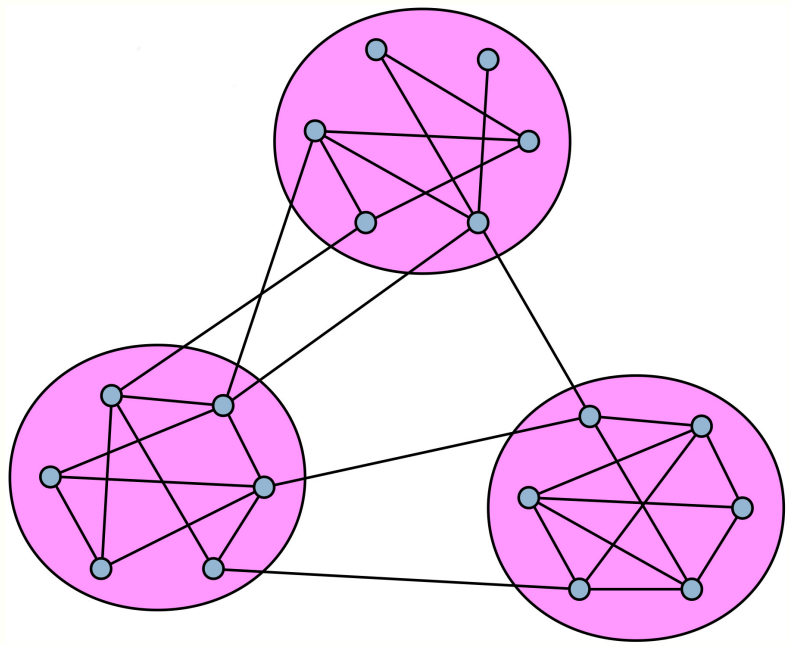
Structural definition

- A highly connected group of nodes
  - Density inside the community much higher than global density of the network
  - Only few edges toward the rest of the network



Types of structural communities

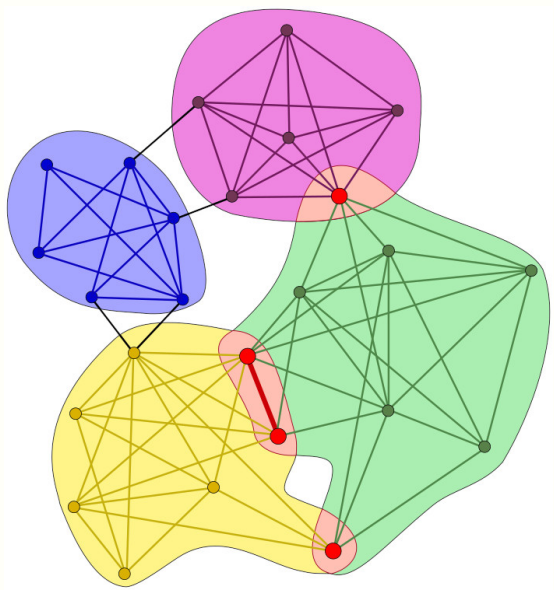
- Partition of the nodes into dense parts sparsely connected between them
  - High density inside communities
  - Few edges between communities



Types of structural communities

- Partition of the nodes into dense parts sparsely connected between them
  - High density inside communities
  - Few edges between communities
- Overlapping communities

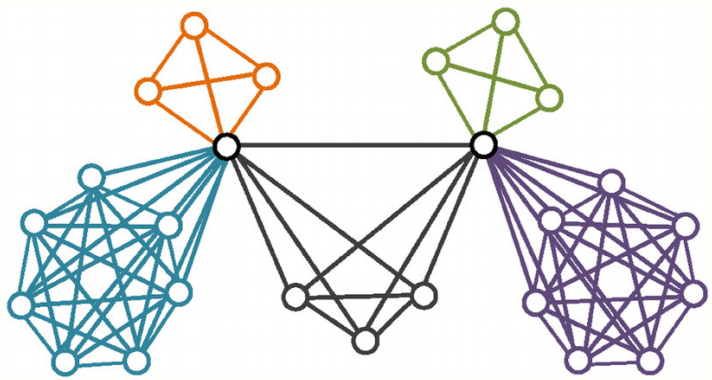
  A node can belong to several communities
  - more realistic
  - problem : how to separate communities ?



Types of structural communities

- Partition of the nodes into dense parts sparsely connected between them
  - High density inside communities
  - Few edges between communities
- Overlapping communities

  A node can belong to several communities
  - more realistic
  - problem : how to separate communities ?
- Partition of the links
  - a link belong to exactly one community
  - a node can have links in different communities



Modelling

Modelling static networks

MODEL = RANDOM GENERATION OF SYNTHETIC NETWORKS

For simulating:

- phenomena
- algorithms
- protocols

In order to:

- design
- test
- predict
- better understand

Q: Do Internet protocols still work if Internet is 100 times larger ?

↳ Generate a synthetic network and simulate

Modelling static networks

MODEL = RANDOM GENERATION OF SYNTHETIC NETWORKS

4 classic properties:

- Low global density
- Short distances
- Heterogeneous degrees
- High local density

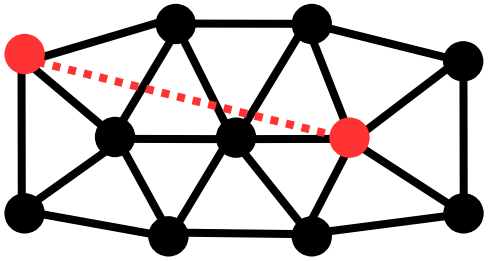
Modelling static networks

MODEL = RANDOM GENERATION OF SYNTHETIC NETWORKS

4 classic properties:

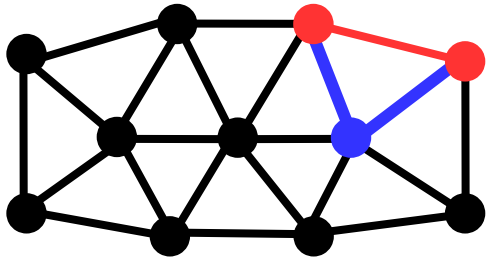
- Low global density
- Short distances
- Heterogeneous degrees
- High local density

proba ???



global density

proba ???



local density

Modelling static networks

MODEL = RANDOM GENERATION OF SYNTHETIC NETWORKS

4 classic properties:

- Low global density → parameter
- Short distances → induced by randomness — Erdös-Rényi 1960
- Heterogeneous degrees
- High local density

Modelling static networks

MODEL = RANDOM GENERATION OF SYNTHETIC NETWORKS

4 classic properties:

- Low global density → parameter
- Short distances → induced by randomness — Erdös-Rényi 1960
- Heterogeneous degrees → compatible with randomness — Molloy & Reed 1995
- High local density

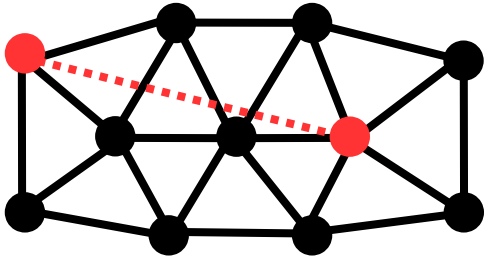
Modelling static networks

MODEL = RANDOM GENERATION OF SYNTHETIC NETWORKS

4 classic properties:

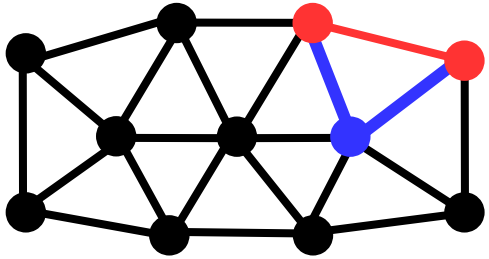
- Low global density → parameter
- Short distances → induced by randomness — Erdös-Rényi 1960
- Heterogeneous degrees → compatible with randomness — Molloy & Reed 1995
- High local density → **problem**

proba ???



global density

proba ???



local density

Modelling static networks

MODEL = RANDOM GENERATION OF SYNTHETIC NETWORKS

4 classic properties:

- Low global density → parameter
- Short distances → induced by randomness — Erdös-Rényi 1960
- Heterogeneous degrees → compatible with randomness — Molloy & Reed 1995
- High local density → **problem**

Big challenge: Generate networks having these 4 properties


low global density

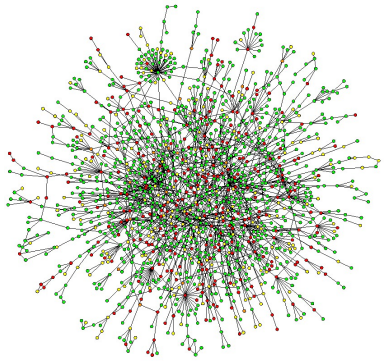

short distances


heterogeneous degrees


high local density

Algorithms

Almost structured graphs

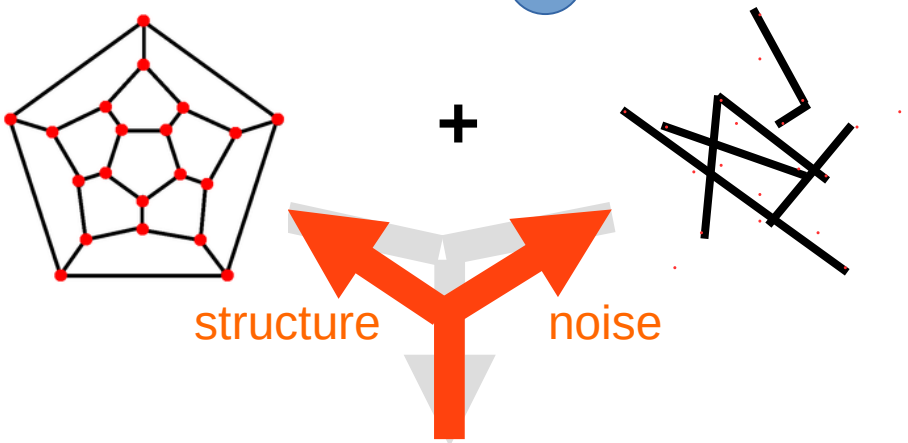


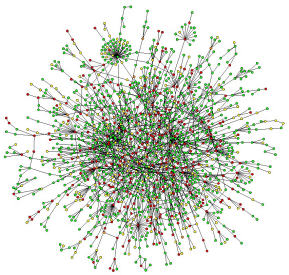
- loosely constrained → **randomness**
- strongly impacted by their context → **structure**

Complex networks = structure + randomness

1 strongly structured

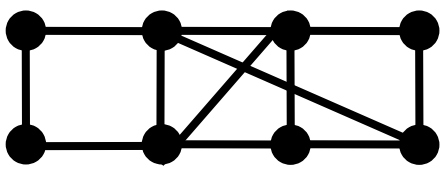
2 random modifications





Graph editing algorithms

INPUT

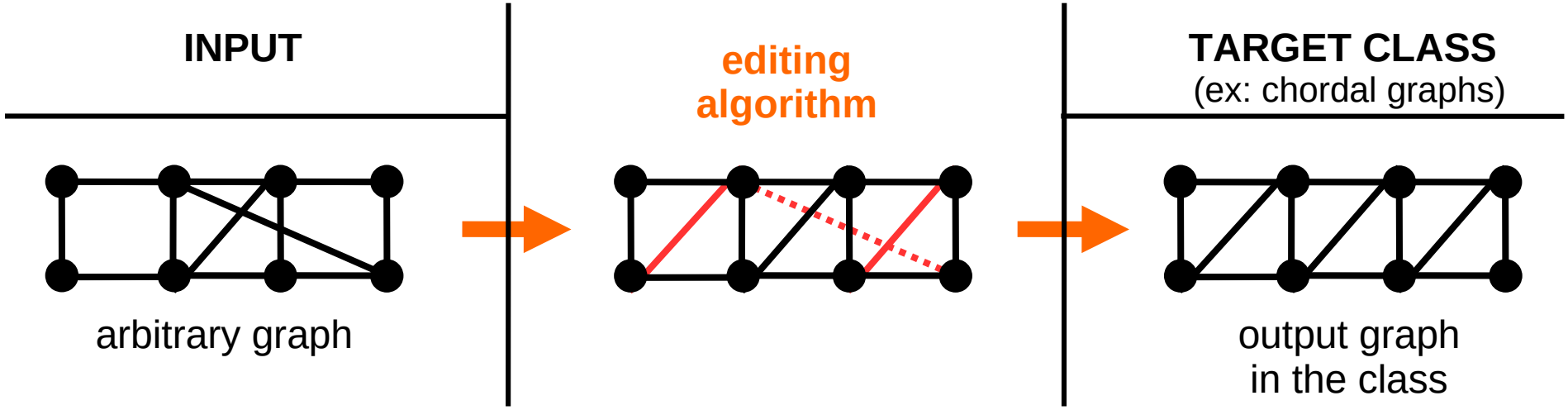


arbitrary graph

TARGET CLASS
(ex: chordal graphs)

Definition:
Chordal graphs = graphs without induced cycle on at least 4 vertices

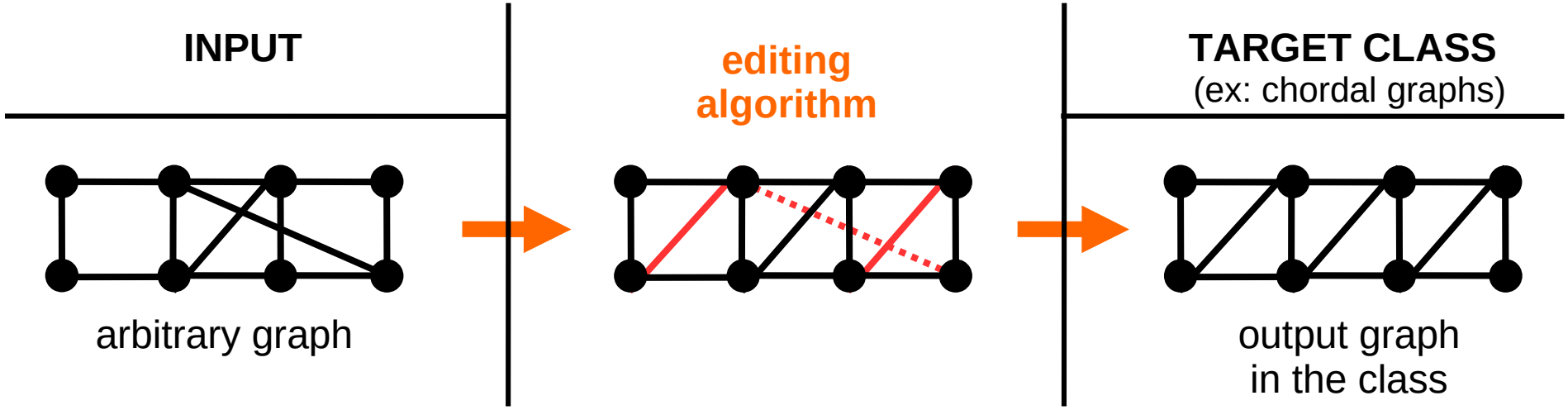
Graph editing algorithms



Definition:

Chordal graphs = graphs without induced cycle on at least 4 vertices

Graph editing algorithms



GOAL: perform as few modifications as possible

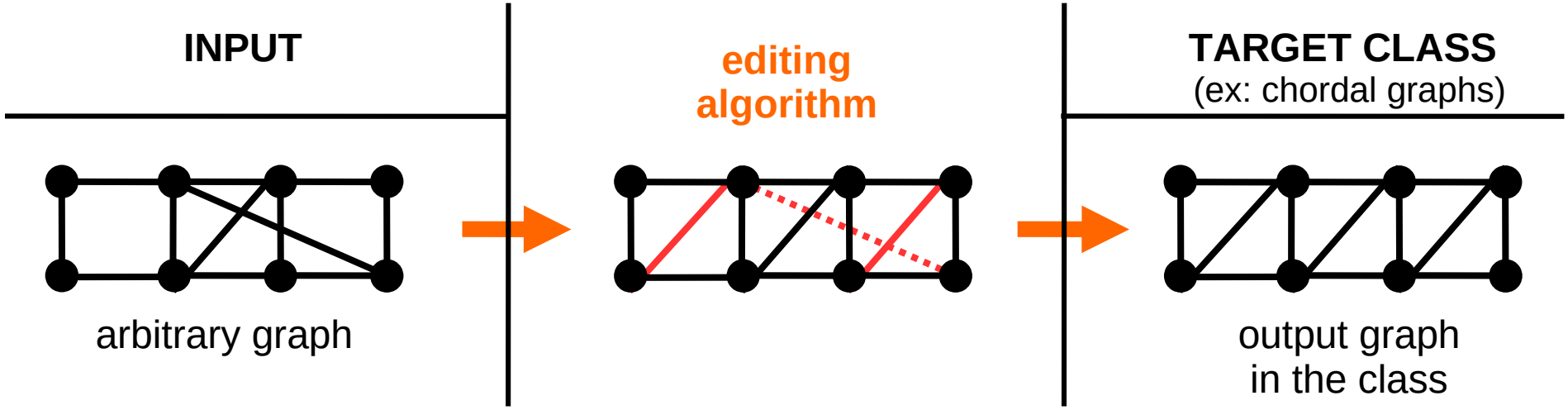
Motivations

- **Mathematics**
  Distance to and projection on a class of graphs.
  How far is a graph from having a certain property?

- **Computation**
  Natural extension of the recognition problem of graph classes.
  When the recognition fail, how to minimally correct the graph?

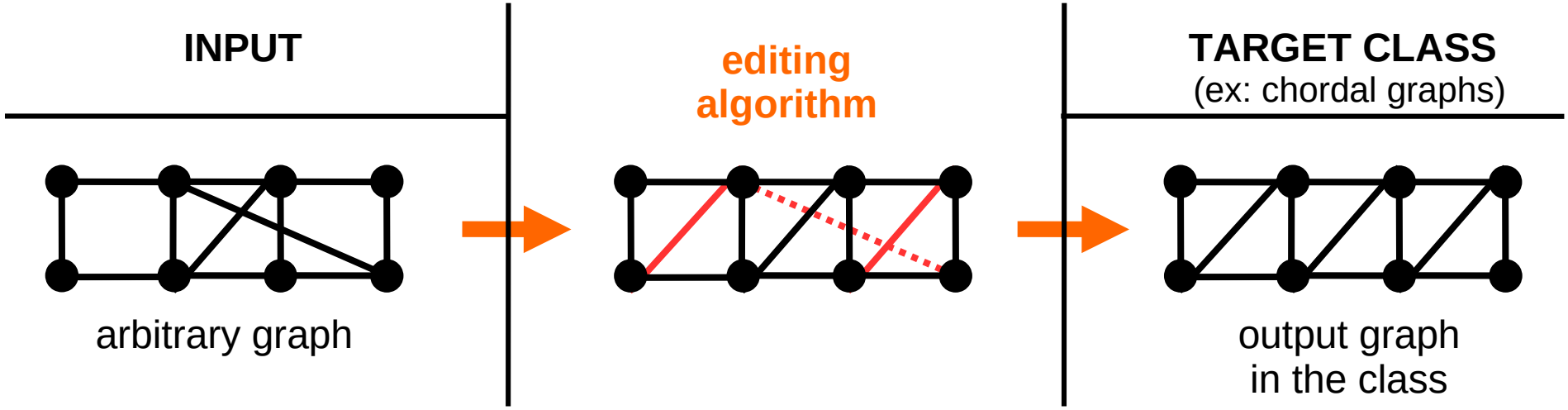
- **Data science**
  Remove noise in graph data.
  - Measurement errors
  - Randomness (non-constrained part of the data)
  - Anything deviating from the main structure

Graph editing algorithms



GOAL: perform as few modifications as possible

Graph editing algorithms

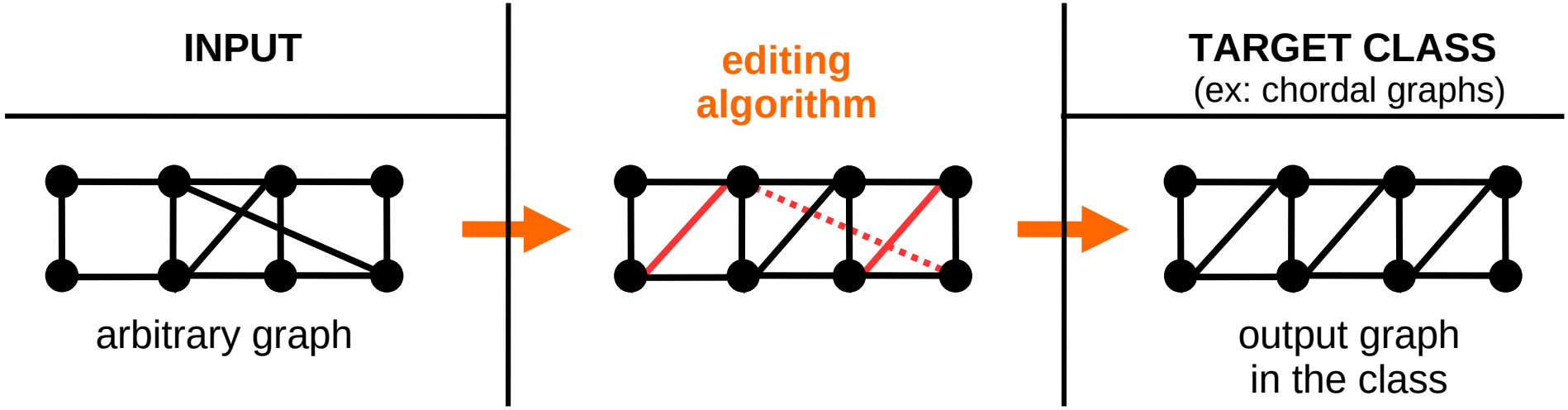


GOAL: perform as few modifications as possible

- Unfortunately: *minimum number* is ***NP-hard*** for most properties

  Even when only one type of modifications is allowed

Graph editing algorithms



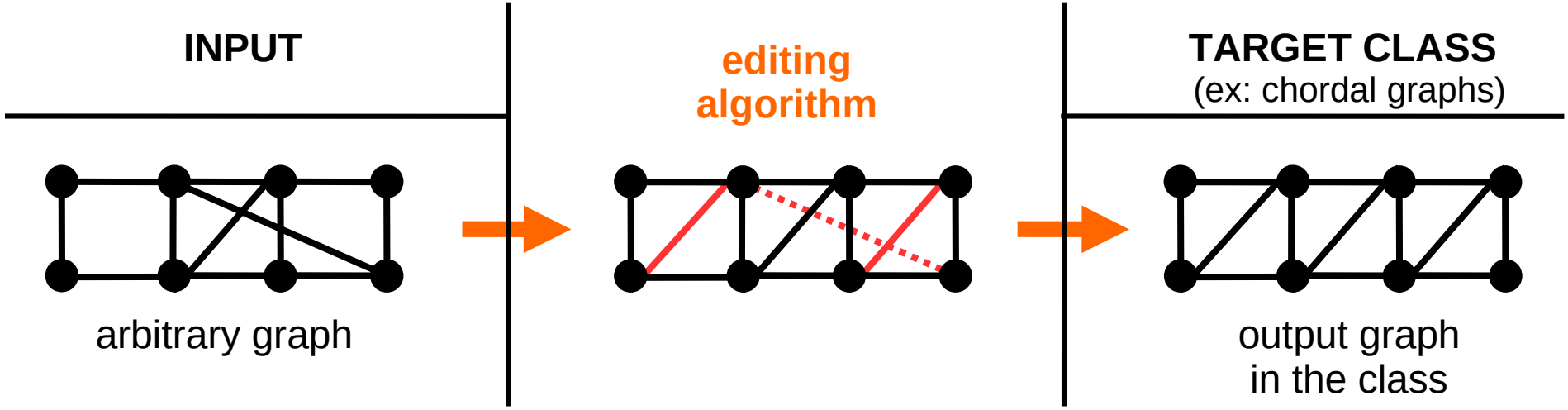
GOAL: perform as few modifications as possible

- Unfortunately: *minimum number* is ***NP-hard*** for most properties

  Even when only one type of modifications is allowed

- Different approaches:
  - Restricted inputs
  - Exact exponential algorithms
  - Parameterized algorithms
  - Approximation algorithms
  - Inclusion minimal modification

Graph editing algorithms



GOAL: perform as few modifications as possible

- Unfortunately: *minimum number* is ***NP-hard*** for most properties

  Even when only one type of modifications is allowed

- Different approaches:
  - Restricted inputs
  - Exact exponential algorithms
  - ***Parameterized algorithms***
  - Approximation algorithms
  - ***Inclusion minimal modification***

Open problems in complex networks analysis and modelling

Christophe Crespelle

christophe.crespelle@univ-cotedazur.fr

UNIVERSITÉ CÔTE D'AZUR

i3S sophia antipolis

DIGITAL SYSTEMS FOR HUMANS
GRADUATE SCHOOL AND RESEARCH