TUM – Courses
IN2045
Network Analysis – statistical and formal models and methods

Chapter: Introduction to Modelling

Dr. Heiko Niedermayer

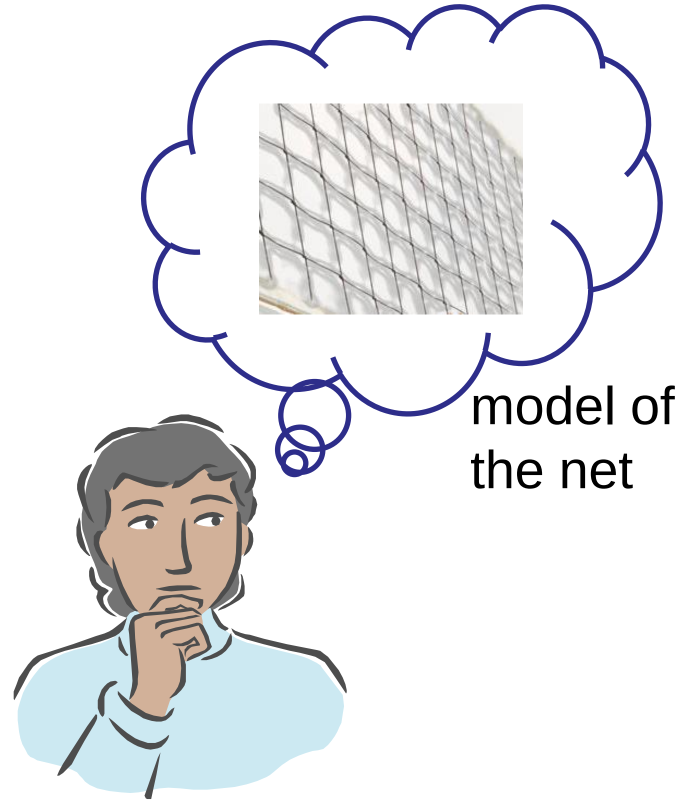
Chair for Network Architectures and Services

Department of Computer Science
Technische Universität München
http://www.net.in.tum.de

Technische Universität München

What is a model?

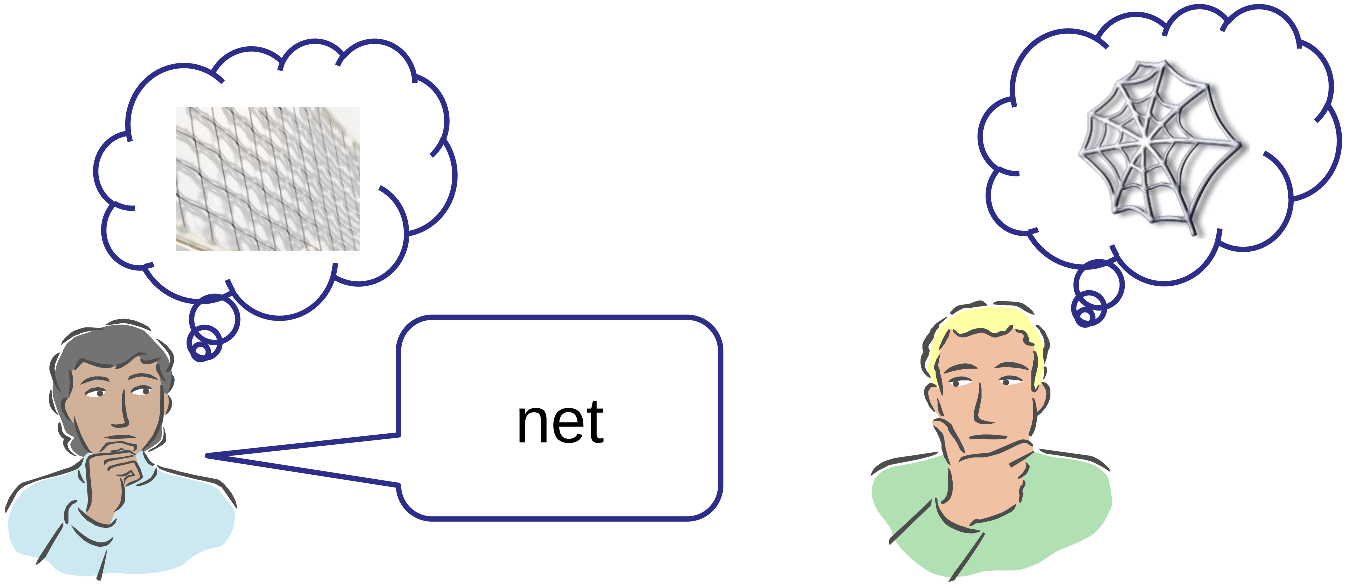
real net



model of
the net

What is a model?

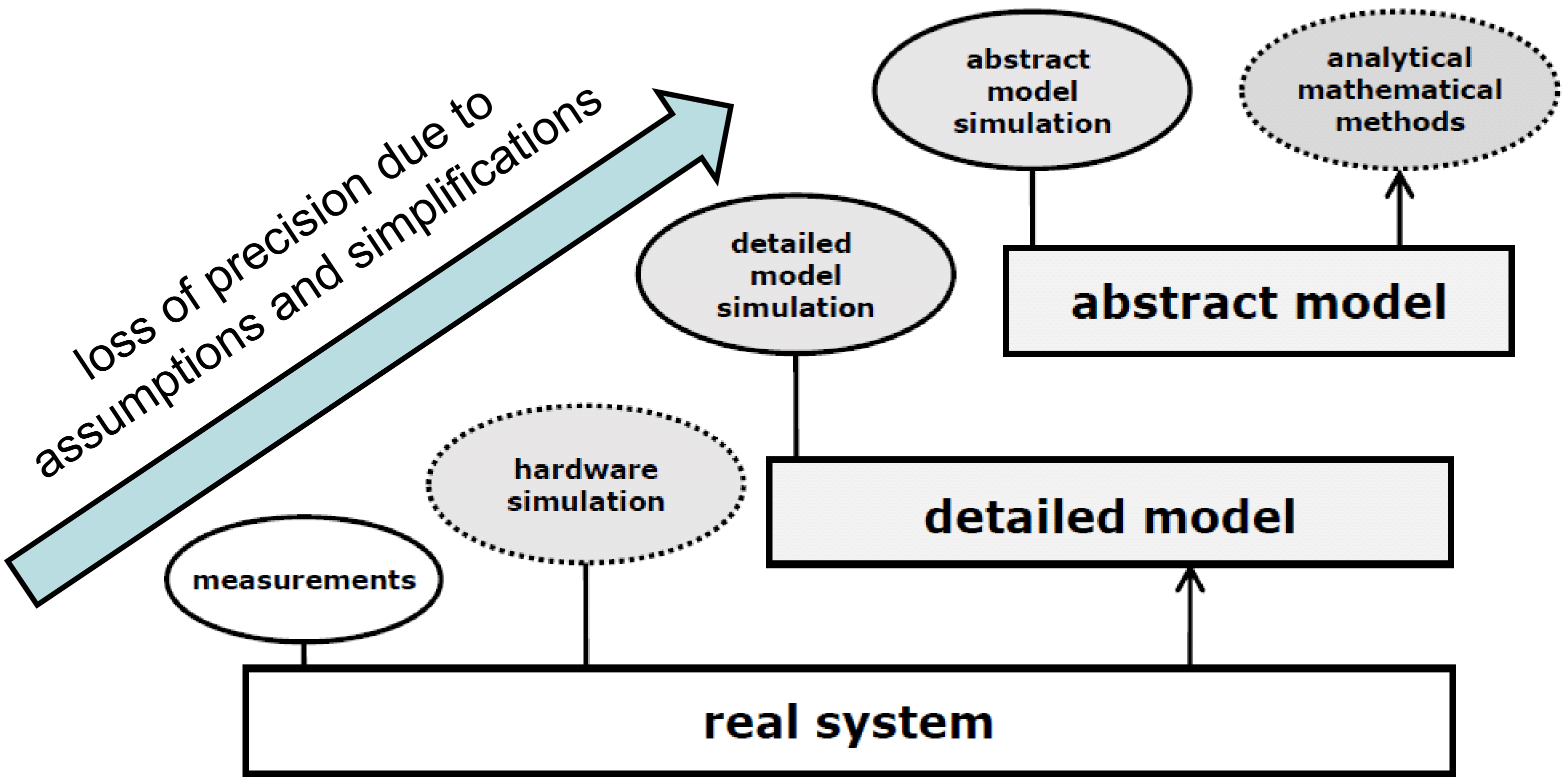
- Whenever we think about something, a model of the thing appears in our mind.
  - The model in our mind is not identical to the thing.



  - The model in other‘s minds is neither identical to the thing nor necessarily identical to our model… (even if both were volleyball nets)

What is a model?

- So, a model refers to some thing (real system), but abstracts from certain details.
  - e.g. due to complexity, but also for conceptual reasons
  - e.g. common details of all system instances vs individual instances



What is a model?

- A Model
  - Is based on some original
  - Simplifications
    - We do not cover all aspects, e.g. the ones that may differ among different instances of a system.
  - Pragmatism / „The Power of Modelling“
    - It may be applicable to other things than the original as well.
    - The concept of a network may apply to spider webs, volleyball nets, communication networks, social networks, the Internet!

- In Mathematical Logics
  - A model usually sets free variables of the logical formulas to a model-specific value.
  - This allows to prove new theorems under the conditions set by the model.

What is a model used for?

- To set a certain set of assumptions
- To find a certain set of assumptions
- To explain
- To explore
- To draw conclusions
- To prove
- To quantify
- To visualize
- To know how to obtain a result
- …

How models appear….

Usage of models

- Implicit
  - Most of the time, models are not described as such, but people and researchers assume a model and then argue on that basis.
  - They often will not name the model nor describe it and all its assumptions.
  - People might not be aware of using a model.
- Explicit
  - When the model is described.

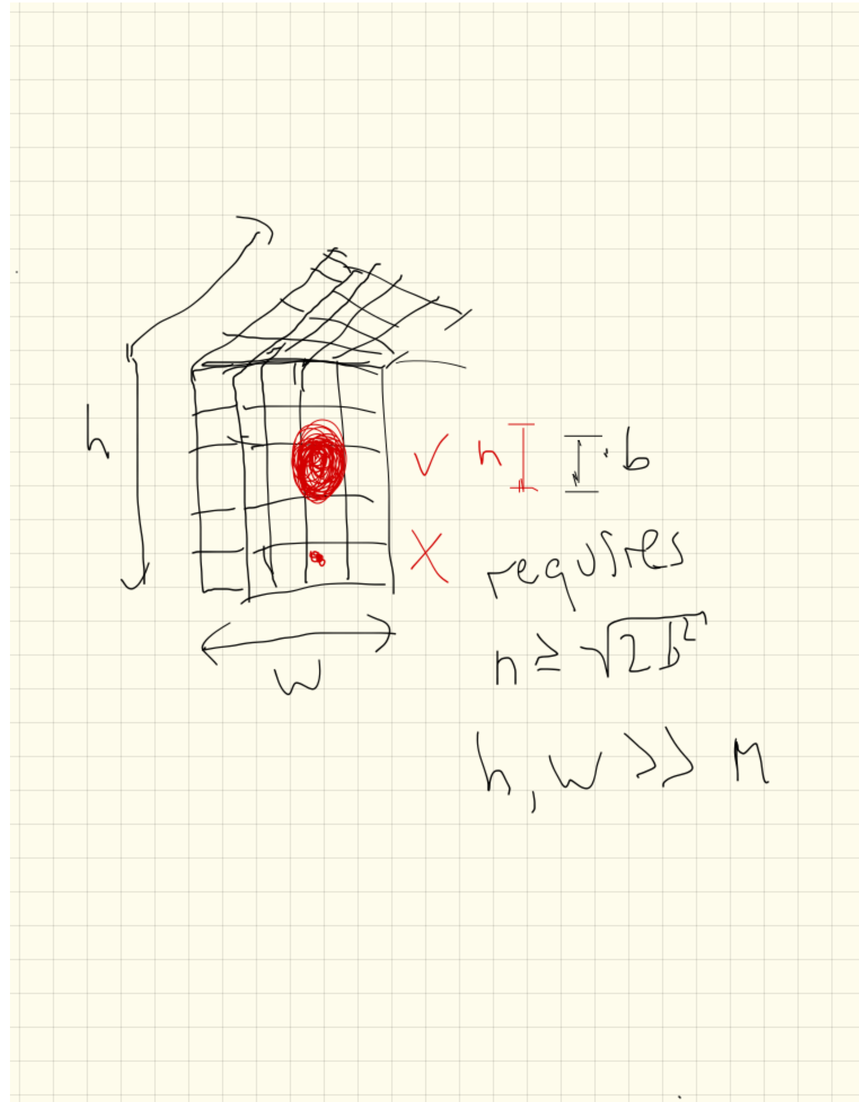
How to describe and discuss a model?

- Informal Text
  - Describe a model in natural language.
  - Most common form.
  - The semantics of real languages can be exploited.
  - Tends to be less structured, and prone to ambiguities.
  - Analyzed via reasoning in natural language

- A made-up example: „Ball Stop Net“
  - A ball stop net is a net with a limited size and only local geographic scope. It is related to a net-specific item called ball. This ball has a certain size. The net consists of nodes and edges where nodes are intersections of edges. Locally, this construction is 2-dimensional, but the construction may be distorted to generate a 3-d form. The difference between two neighboring nodes has to be smaller than the size of the ball. The overall net has to be significantly larger than the ball.

How to describe and discuss a model?

- Mathematical Formulas
- Logics
- Drawings, Visualization
  - Components, their relations, etc.
- Modelling Language
  - Instead of natural language.
  - Do not mix up a language to write down a model with the modelling!
- Pseudo Code
- Program Code for Simulation or Emulation
- …

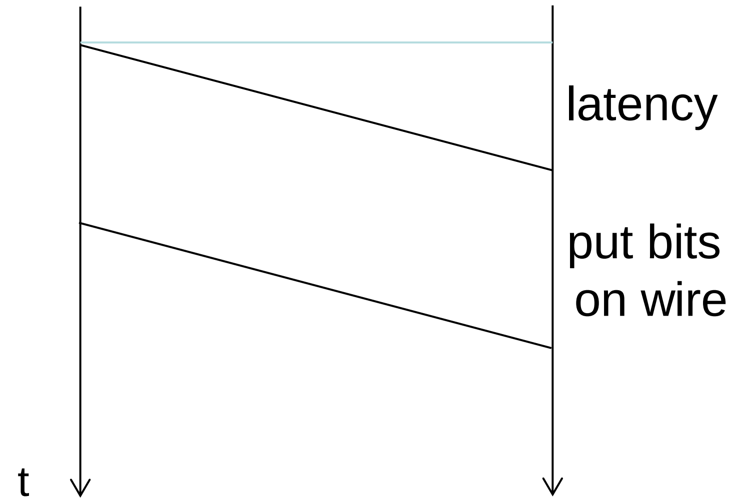


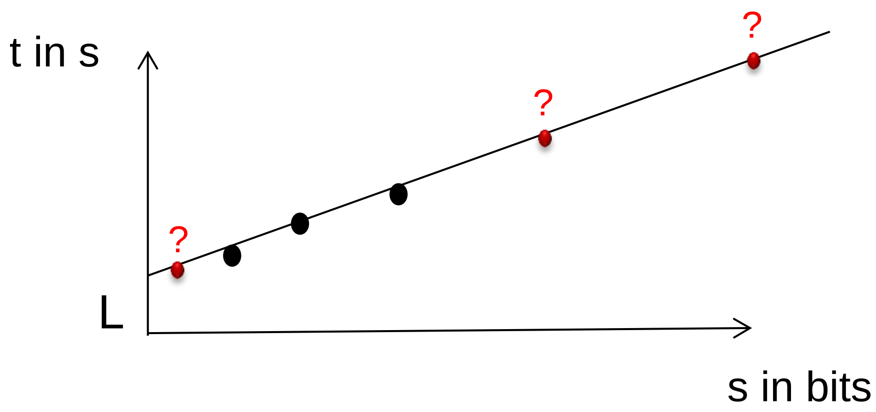
Benefits of modelling

- In the following, we give examples for benefits of generating models and applying them.

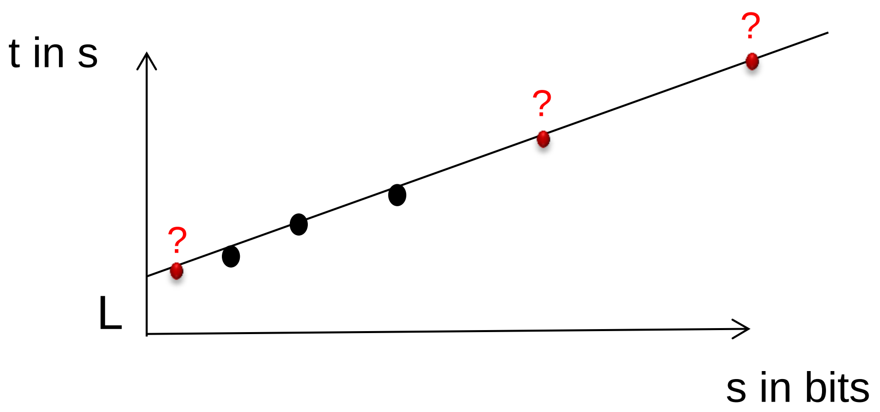
Benefits of Modelling – Predict yet unknown value

- Cycle time of a packet a line
  - Data rate: R = 100 Mb/s
  - Latency: L = 3.5 ms
  - CycleTime(packetsize s)= L + s / R
  - Theory suggests the linear formula
  - Values R and L have to be determined, how many measurements are required?





Benefits of Modelling – Predict yet unknown value



- Thus, having a model gives us the possibility to measure a few values and predict all the others.
  - Possible infinite others!

- Beware, models tend to have assumptions that limit their applicability.
  - E.g. for large packets of, say, 10^20 bits transmission errors may become important for the analysis… or storage operations …..

Benefits of Modelling – Become aware of meaning and consequences

- **Write down all principles, knowledge, facts, … that you know.**
- **Everybody knows how students are. So, here the facts (not serious):**

All students are lazy, they only party.

Students give terrible answers in exams.

Students will earn lots of money
because all of them will have a bachelor or master degree.

Students who give terrible answers, will not pass the exam.

If you did not learn, you will give terrible answers in exams.

Students who party, do not learn.

To earn a bachelor degree, a student has to pass exams.

Benefits of Modelling – Become aware of meaning and consequences

- **Lets apply some reasoning**

All students are lazy, they only party.

Students who party, do not learn.

∴ All students do not learn.

If you did not learn, you will give terrible answers in exams.

∴ All students will give terrible answers in exams.

Students who give terrible answers, will not pass the exam.

To earn a bachelor degree, a student has to pass exams.

∴ No student will ever earn a bachelor degree.

Benefits of Modelling – Become aware of meaning and consequences

- **So, who of you has a Bachelor Degree?**

XYZ has a bachelor degree and was student when he/she earned it.

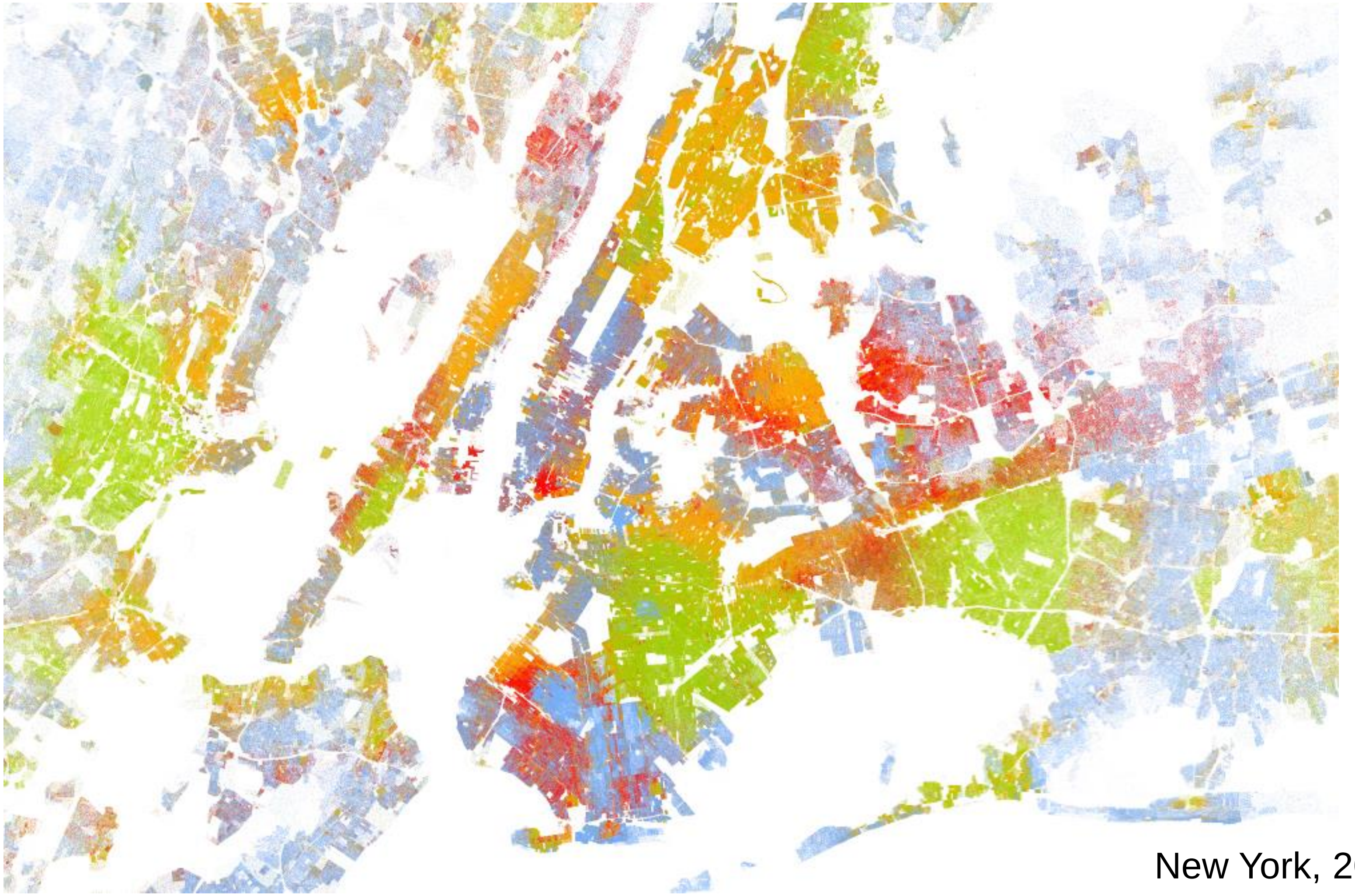
Concluded Fact from Model:
No student will ever earn a bachelor degree.

∴ CONTRADICTION !

So, we are not only able to draw conclusions when we apply modelling. We can also find that our model assumptions may lead to contradictory statements or conflicts with reality.

Benefits of Modelling – Find non-obivious explanations

A classic example from sociology: Segregation Modelling



New York, 2013

Source: http://findingjustice.org/new-maps-highlight-racial-segregation-america/

Benefits of Modelling – Find non-obivious explanations – Segregation

- As the map on the previous slide shows, people do not tend to live equally-distributed over the map according to the percentage of their peer group, but peers seem to cluster together.

- Peer groups can be:
  - Ethnic groups
  - Rich / Poor
  - …

- So, whatever the real cause for that effect is (and we won‘t give a final answer), there is a process in society that leads to this situation.

Benefits of Modelling – Find non-obivious explanations – Segregation

- Analyzing the problem
  - There are people distributed over a city, generalized some 2-dimensional space.
  - People belong to peer groups.
  - There is a notion of neighborhoods.
    - A small geographic area in the 2d space.
    - There are multiple people in a neighborhood
    - We evaluate the neighborhood with respect to the peer groups.
  - People seem to have some (hidden?) preferences where they live…
    - move to, want to leave, ….

Segregaton – possible explanations (simplified)

- Lets first start with the not non-obvious ones….
- Explanation 1: People from each peer group are only allowed to live (rent home or buy home) in certain areas that are reserved for the peer group.
  - In that form not plausible today, but financial aspects and subsidized housing might have similar but weaker influence.
- Explanation 2: People do not like people from other peer groups and segregation happens because they want to live among their peers only.
  - While true for some, studies that ask people what they think and want tend to come to different conclusions.

Segregation Modelling

The purpose of the modelling is now:

- To find out if such explanations can really lead to the situation found in reality.
- To find other explanations that (also) achieve this.
- To better understand what kind of actors exist and what could drive their actions.

Measuring Segregation

- To continue our analysis, we have to understand what it means to find segregation in our data or modelling results.
- We have peer groups PG_1, PG_2, PG_3, ….
- For each group, we know their share of the city's population: GlobalAvg(PG_i)
- When we take a neighborhood, we can compute the local shares in the neighborhood and compare them with the global shares.

Measuring Segregation

- If the local shares are identical to the global one's, we find that the people are evenly distributed and that segregation would be 0.

- $LocalShare(pg) = \frac{\#(people\ from\ peer\ group\ pg)}{\#\ (people\ in\ neighborhood)}$

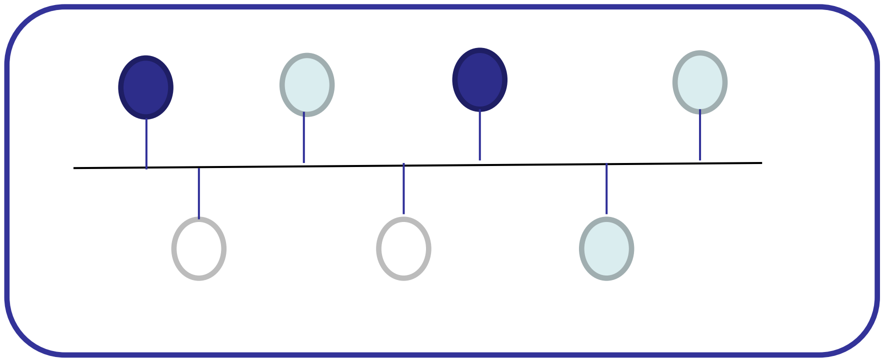
- $segMeasure = \sum_{pg\ in\ PeerGroups}(GlobalAvg(pg) - LocalShare(pg))^2$

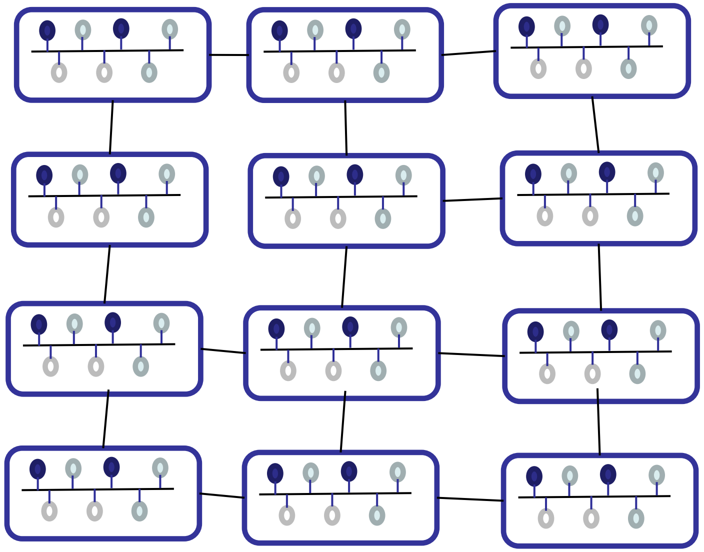
Segregation Modelling – First Try (a set of LANs)

Lets try with something network-alike:

- A neighborhood could be like a local area network (LAN)
  - The neighbors are one broadcast domain that interacts with each other.
- The city would be a network / grid of such LANs
  - So while the neighbors have a large influence on a peer‘s happiness. Peers from neighboring LANs might have some reduced influence.

LAN





Segregation Modelling – First Try (a set of LANs)

- Initially, we might start with randomly positioning the peers of the city.
  - Other options are possible.
- Now, we can implement a set of strategies that peers may follow.
  - Some strategies might need additional modelling, e.g. prices for new rents in neighborhood and incomes of peers.
  - For this chapter, we will only consider that peers are interested in the shares of the peer groups.
- Strategy of a peer from Explanation 2:
  - IF all peers in my neighborhood are from my peer group THEN stay
  - ELSE move to other place.
- Note that there are multiple options for how to implement the move.
  - It could be a random place that is free.
  - It could also involve maximizing among multiple places.
  - It could also invovle not moving when all candidate places are worse.

Segregation Modelling – First Try (a set of LANs)

- Other strategies
  - Avoid my peer group
  - Don‘t care
  - Don‘t care, but I do not want to be the only freak in my neighborhood
  - I want some of my peers around me
  - …

- We could now write a simulation script to see what happens if peers play these strategies over multiple rounds. What does it converge to? Does it converge? …

  → Shall we write some code?

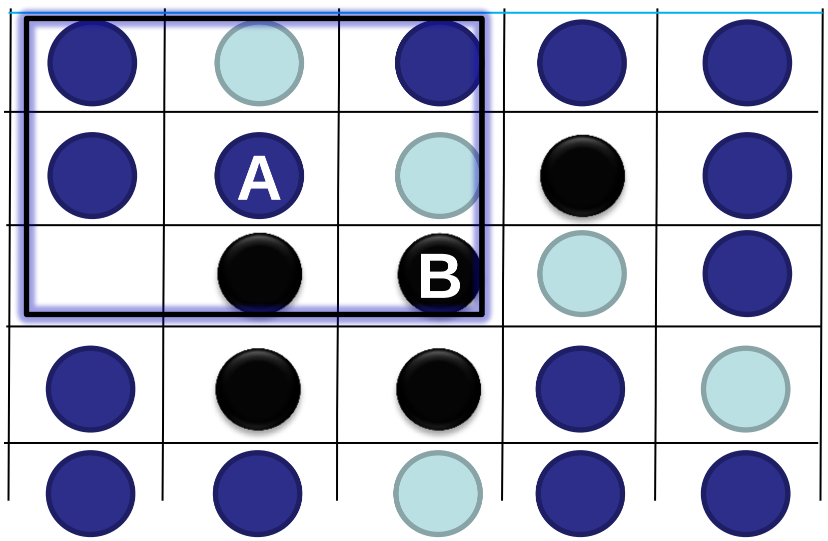
- Note that in our model all peers in one neighborhood, see the same.
  - Neighboring LANs only influence strategies that take them into account.
  - Big fences around neighborhoods to separate them are not that common.

    → so lets go to the prominent model from Schelling

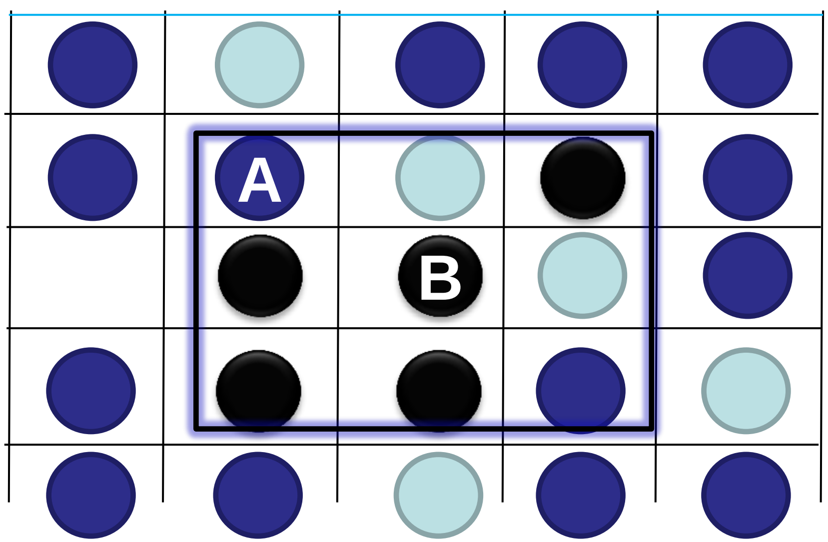
Schelling's Segregation Model

- Possible places to live form a grid. Some places taken by a peer, some not.
- The neighborhood for a peer is the set of peers around this peer on the grid.
  - In our example the 8 peers around the peer.
  - As you can see, even though A and B are neighbors, they see a different neighborhood.

Neighborhood of A

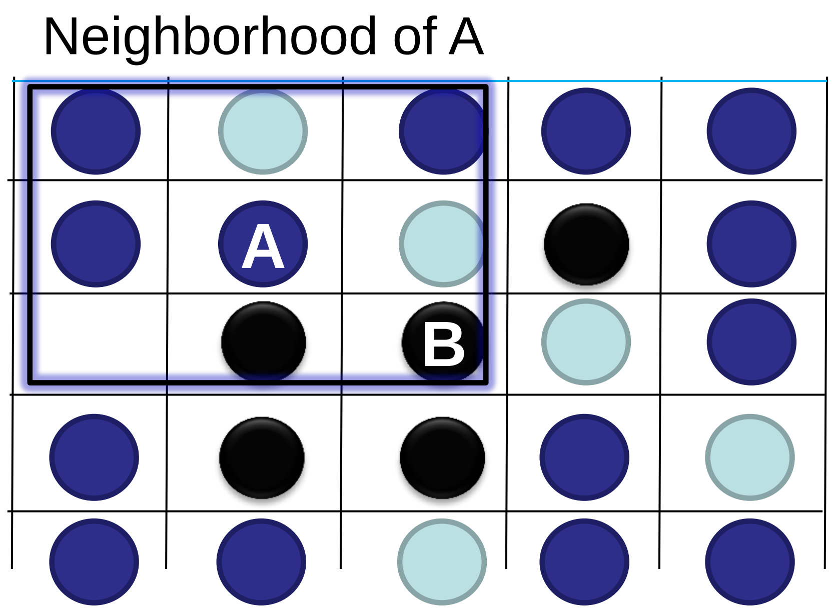


Neighborhood of B



Schelling's Segregation Model

- The strategy that we now assume is that
  - Peers more or less like diversity
  - … but they want a minimum share of peers from their peer group.
- So basically, we can set it from rather tolerant (few from my peer group) to intolerant (only my peer group).



→ Lets look at some simulations with NetLogo.

Percent Unhappy: Peers with too many non-simular peers.

Percent Similar: Average percentage of similar peers in neighborhoods.

Schelling's Segregation Model

Simulation Results:

- Segregation becomes visible even if smaller percentages of similar peers (from same peer group) are wanted.
- The percentage of similar peers on average much higher than this threshold value.
  - E.g. 70% (achieved) similar peers for threshold of 30 % (wanted).
  - This can be seen as mismatch between microscopic action of the peers and the macroscopic result.
- Unstable non-convergent behaviour with low similarity if threshold high.
  - Partially due to the random search for a new place if agent moves…

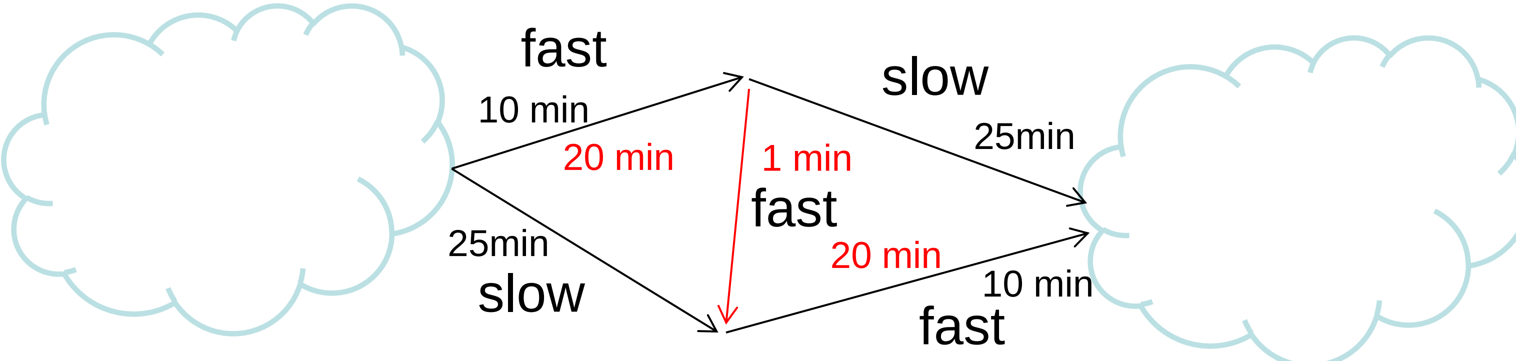
Benefits of Modelling – Conclusions

The segregation discussion has shown further modelling benefits:

- **Non-obivious explanations**
- **Microscopic** behaviour (individuals, agents, code) might differ from **macroscopic** results (globally observed situation)
  - A fundamental reason why we need modelling, simulation, etc., because unexpected things may happen and simple analysis may be wrong.
- Please note that this does not show that Schelling‘s model and the actions of the agents is the correct explanation, but it shows that from observing segregation (macroscopic behaviour) one cannot savely conclude that that is what the agents want (the microscopic behaviour)

Benefits of modelling – unexpected consequences

- Braess‘s paradox is an example from networks / transportation where adding an additional link can reduce the performance (maybe more on it at the end if we discuss game theory).



Without the red edge, drivers / packets can only chose from fast/slow and slow/fast, which is seen as same cost and they can randomly distribute over the alternatives.
With the red edge, the option fast/fast becomes optimal. However, the fast edges become slower the more they are used (10min → 20min). Given the right parameters, this will be slower than before (35 min → 41 min), but changing to a slow one is worse as both routes go through the jammed fast edges that are still a bit quicker (45 min).

Reasoning

- Before we end this chapter and start to work with formal logic and theorem provers, lets first discuss reasoning.

- Some slides before, we have already applied basic reasoning.
- We start with a set of premises, which can include assumptions (what are talking about).
- These arguments support a conclusion that we draw from them in the end.

Premise
Premise
Premise

∴ Conclusion

Deductive Reasoning

Premise	Ethernet has 1Gbps
Premise	DSL has 16 Mbps
Premise	Faster than DSL → is fast.
∴ Conclusion	∴ Ethernet is fast.

- Deductive reasoning strictly follows logical calculus.

- The main point there is that the argument is valid. This means that under no conditions can the premises be true, yet the conclusion be false.
  - And if that should happen, then the argument is not good. It is false.

- There are valid forms that are valid indepent of the semantics. E.g. Modus Ponens:

$$\frac{a \qquad a \rightarrow b}{\therefore\ b}$$

Deductive Reasoning

- Another example:

$$\frac{\neg b \qquad a \rightarrow b}{\therefore \quad \neg a}$$

- Logical syllogisms deal with membership to certain predicates. This can be come difficult when NOT or existential quantifier is used.

$$\frac{\text{All A are B} \qquad \text{All C are A.}}{\therefore \quad \text{All C are B.}}$$

$$\frac{\text{No A is K} \qquad \text{All F are K.}}{\therefore \quad \text{No F is A.}}$$

$$\frac{\text{All A are B} \qquad \exists \text{ A, A is C.}}{\therefore \quad \exists \text{ B, B is C.}}$$

Quantifiers:

- FOR ALL(x from empty) WHATVER → TRUE
- THERE EXISTS (x from empty) WHATEVER → FALSE

Inductive Reasoning

- Deductive Reasoning assumes a certainty that in reality is often not available.

- Example:
  - What if God changes physics tomorrow and ethernet will not work while DSL still works?
  - Unlikely yes, but something could happen, maybe just a device failure.

Ethernet has 1Gbps
DSL has 16 Mbps
Faster than DSL → is fast.

∴ Ethernet is fast.

- Inductive Reasoning tries to give arguments in cases when we cannot be certain.

8 out of our 20 computers from A failed.
3 out of our 14 computers from B failed.

∴ You should get one from B.

Inductive Reasoning

- Inductive arguments are usually not valid.
  - Because of their statistical nature.
  - Because of their predicting on the basis of experience.
  - Because of operating on analogies.
- Thus, it is not structure that makes inductive arguments good or bad.

- The strength of these arguments comes from the strength of the evidence they present and how moderate the conclusion is.
  - More precise: The conclusion should be certain given the evidence.

- Important: New additional facts / premises can make the strongest evidence weak.

8 out of our 20 computers from A failed.

3 out of our 14 computers from B failed.

It was a bus accident at our retreat that destroyed the 8 computers from A.

∴ You should get one from A.

Conclusion

- Whenever we analyze a system, we apply some modelling.
- Models can have many forms and purposes.
- Models allow us to reason about systems.
- This is particularly interesting when the complexity of a system makes it impossible to safely predict outcomes.
- With respect to analysis, try to reason carefully. Be aware that there are deductive (mathematics) and inductive variants (natural sciences, empirical arguments).
  - Both are important and related to the next chapters (deterministic modelling, modelling with uncertainty).