CSS Exam Questions SS22

Question Group GEN (General Social Simulation)

1. What is an analytical model? What is a simulation? What discerns simulation from other analytical techniques? When might simulations be needed?

An analytical model is a mathematical model that uses equations to describe changes in a system. They have a closed form solution which can be calculated using the equations. (e.g. calculate the values of dependent variables)

A Simulation is a type of modeling where inputs are entered at the beginning and during runtime outputs are calculated and can be observed. The focus is on the behavior of the model through time.

Simulations can be used for representing dynamic systems. For example, real world events where random processes play an important role and where it is not possible to use analytical approaches to explain these events. In such cases, Simulations can be used to make approximations for the behavior of the system.

What Simulations are needed for:

- Understanding real world scenarios better
- Predicting future scenarios
- Substitute human capabilities (e.g. simulate knowledge of experts, doctors, chemists)
- Training pilots, train drivers, racing drivers, etc.
- Entertainment (games, Sims, flight simulator, etc.)
- Discovery, Science (building models and testing theories)

2. What is an emergent phenomenon (emergence)? Give examples for emergent phenomena! Which modelling techniques can simulate emergent phenomena?

A phenomenon is emergent if a new category needs to be formed to describe it, which is not required to describe the behavior of the underlying components of the phenomena.

Examples:
Motion of Atoms → Temperature
Motion of Birds → Flocking Behavior
Motion of Cars → Traffic Jams
Motion of People in stadium → Mexican Wave

A single Atom does not have a Temperature but a collection of atoms has a temperature since it is an emergent property of the atom collection’s motion.

Modelling techniques: Only Bottom-up techniques are possible
Individuals need to be modeled (emergence is the result of interaction)
+ Multiple levels need to be possible (emergence effect on 2nd layer)

3. What are agents in the context of social simulation? What is a self-organizing system? What is stigmergy?

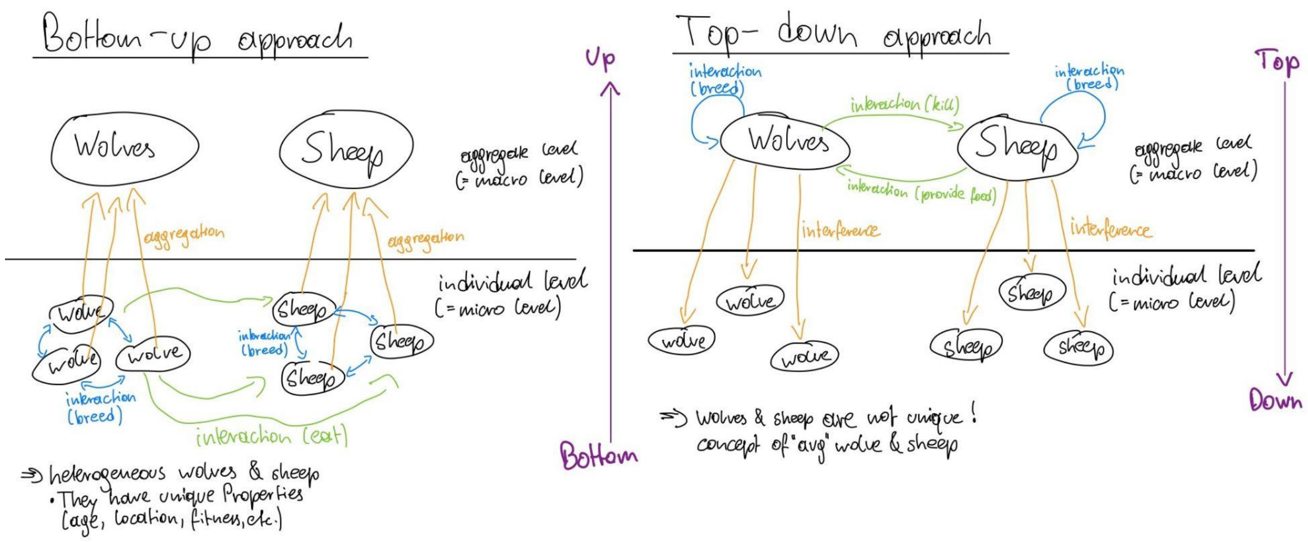
Agents are entities/simulated individuals that may interact with each other or have a certain behavior. Agents don’t have to be humans, they can be anything in the context of the simulation. Agents are heterogeneous meaning they have properties that can differ for each agent (e.g. health, age, etc.)

Self-organizing systems don’t need central planning or leader, and there are no predefined directions or patterns that they follow. However, a pattern can still occur which is in this case self-organizing (e.g. Birds forming swarms and observe the behavior of the other birds, there is no clear leader bird but still they follow the direction of the swarm)

Stigmergy is when agents self-organize through their environment. Ants for example do not follow the actions of other ants directly (compared to bird swarm for example) but they leave behind traces of pheromone when they find food on a successful path, other ants start to follow more intense paths and patterns emerge.

4. What is the difference between top-down and bottom-up modeling techniques – and which simulation methods are top-down, which bottom-up. Give an example that shows the difference between the two approaches (also draw a sketch showing the layers, agents, interaction and direction of inference; choose a different example than sheep & wolves).

Bottom-up BU	Top-down TD
More focus on individual level because heterogeneous development of agents is possible. In a BU simple rules can be sufficient to explain complex phenomena, for example flocks of birds may be very hard to describe if the focus is on the swarm pattern but are a lot simpler described though the rules of movement of individual birds.	Many processes may seem complex from a TD approach, no heterogeneous agents, aren't accurate representations of most socio-economic processes, interaction between individuals don't have a high relevance.
Microsimulation, Cellular Automata, Agent Based Modelling	System Dynamics



5. Why is prediction of social processes problematic (compare purely physical processes with social processes)? Why did the apple not hit the ground (and what does this represent)? Which social systems can be better predicted (e.g. with analytical methods or for longer periods) than others and why?

Physical processes follow exact laws of nature (e.g. Gravity) that can be known equations and that can be solved (e.g. Apple hitting the ground after 1 s after solving equation).

Social processes don't have an exact future state that can be predicted, because not all information is completely known by everyone and the behavior of many individuals has an influence on the outcome. Making predictions in social systems is therefore much more complicated.

The apple didn’t hit the ground because of interference (e.g. someone catching the apple before hitting the ground) by an individual (only this individual knew of the interference before = asymmetric information). This changed the outcome, and therefore the predictions were wrong.

For medium to large homogeneous system (= systems with huge number of individuals) short to medium term predictions are possible because often individual behavior is evened out completely. Except if there are single individuals with very high influence on the whole system. Long term predictions problematic because the structure will change over time.

6. Which four computational social simulation techniques have been presented in the lecture? Shortly describe them and (especially important) also state for which kind of process/situation each technique would be suited well!

1. System Dynamics (Top-down)

Used for policy analysis and design. The behavior of a system is a result of its structure. Tries to capture a system's structure and understand the dynamic complexity of a system. social systems are in the main focus and especially their interdependencies (e.g. lagged reactions in networks, chains). Policies that aim to stable a social system might even have a destabilizing effect due to delayed reactions. Should help to design more effective policies and lower policy resistance (unintended effects of policy changes, e.g. more traffic after road building program). Causal (Feedback) Loops together with Stocks and Flows play an important role for SD simulation to show causal influences and effects in a system.

Mathematically:
A system of coupled nonlinear first order differential (or integral) equations (this is the flow)
Simulated time is cut into discrete intervals (causes slight calculation error) of length dt.

Example for SD: Predator - Prey model

Elements are homogeneous (no individuals possible), the simulated system’s structure stays static during the simulation.

SD works good for a more abstract and simpler simulation of complex systems that have unknown "effects" that we cannot consider. Instead of including every little detail in the model, and being anxious that we do not forget anything, we can just model the aggregation e.g. we can just model the birthrate of a population rather than trying to model the underlying phenomenons or interactions between individual agents.

2. Micro simulation (Bottom-up)

Has individual units in simulation with attributes (e.g. age, sex, income, etc.) and a set of global rules for the units:

- deterministic → probability = 1
- stochastic → probability < 1

Tries to estimate the outcomes of applying these rules (e.g. for policy evaluation). Can be done for short or long periods of time.

They can become kind of similar to ABM models, but there are still some differences:

- In ABM, interaction between individuals on the same layer possible
- In ABM, Agents have strategic decision rules

Often used to build models for Taxation, pension, financial systems.

But also for Spatial Microsimulation where differences in geographical position play an important role (e.g. EU member countries; it makes a difference if you are in Greece or Germany in terms of tax, healthcare etc. policies) → Economic Policy Evaluation, Healthcare Policy Evaluation, etc.

3. Cellular automata (Bottom-up)

Built up on grid environments (1D,2D,3D) where cells are placed on.

Consists of simple individuals “Cells” which are finite automaton that behave in the same way. In every point in time, every cell has one of a few specific states (often 2 states, dead/alive). In the initial configuration, the cells are assigned a state, which is then updated for every cell at the same time at every step. The rules often state that cells react to their own current state and the state of their neighboring cells. All cells have the same rules applied to them.

It would be suited well for a simulation of a spreading virus in a network or population, or basically every situation where the state of an individual might be influenced by its neighbor's state over time.

4. Agent based Modelling (Bottom-up)

ABM can be a wide range of different computational models (can be very different from each other depending on where applied)

But all ABMs share some similarities:

- Simulating the behavior of many heterogeneous agents
- Explaining the aggregate behavior of an agent population
- Explaining unobservable individual behavior.

Common ABM Structure:

- Agents: individual decision units (=heterogeneous)
- Relations: Interactions between agents in the Environment (temporary/permanent)
- Environment: virtual world, rules, everything that is not an agent

They are suitable for many disciplines:
- Computational Economics (Theory and Applied)
- Business Process Simulation
- Geo-Information
- Physics, Biology, Medicine

Examples: (there is a lot, since they are suitable for many disciplines)
- Bird flocking (Self-organizing swarm 2D & 3D)
- Paths (Ant stigmergy)
- Covid (Virus spreading)
- etc.

Question Group SDM (System Dynamic Modelling)

7. What is (are) the world model(s) (Club of Rome)? What did it try to show? Was prediction of exact values a goal?

There are several global models that have been generated throughout the world, they are system dynamic models that deal with complex interactions of the economy, the population and the ecology of the World and try to predict the future state of the world, so the prediction of exact values is not the goal.
World3 model was produced by the Club of Rome and documented in the book "*Dynamics of Growth in a Finite World*", it added new features to Jay W. Forrester's World2 model and is the spark of all later models. The model consisted of several interacting parts. Each of these dealt with a different system of the model. The main systems were:

- the food system, dealing with agriculture and food production
- the industrial system
- the population system
- the non-renewable resources system
- the pollution system

It predicted the instability of the current global socioeconomic system and the general kinds of changes that will and will not lead to stability, economic growth could not continue indefinitely because of the limited availability of natural resources. The land is finite, so the agricultural inputs can not keep up with demand, there necessarily will be a food collapse at some point in the future; also that when nonrenewable resources are extracted, the remaining resources are increasingly difficult to extract, thus diverting more and more industrial output to resource extraction.
It failed to predict pushes of modernization leading to a structural change away from heavy industry towards a society with advanced technology and a service-based economy: the rise of the computers and the Internet

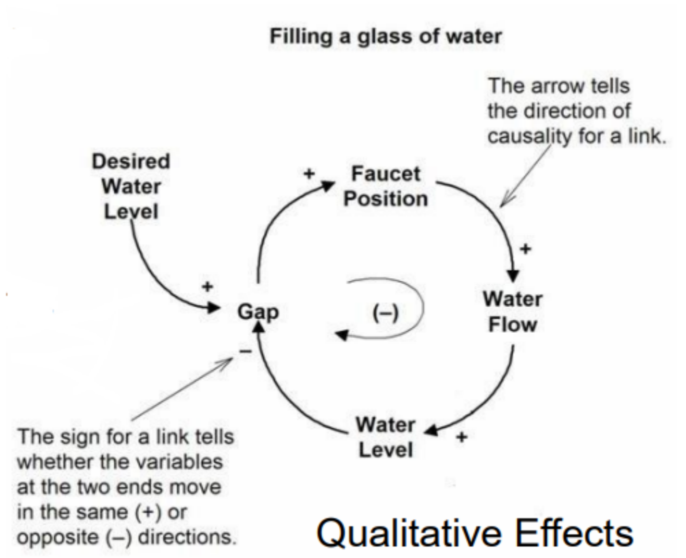
Prediction of exact values was not the main goal. Instead, main goals were:

1. gain insights into the limits of our world,
2. study the most relevant factors of these limits
3. warn people to be cautious of the economic policies (sustainable lifestyle is nice 👌)

8. What is a causal loop diagram? What are feedback loops? Give an example for a positive and a negative feedback loop! What are both methods good for? When would they better not be used?

A causal loop diagram is a visual representation of feedback loops that help to visualize how different variables are interrelated in a system dynamics modeling.
It consists of some elements, arrows and the corresponding sign (positive or negative).
The arrows tell the direction of causality for a link, and the sign for a link tells whether the variables at the two ends move in the same (+) or opposite (-) direction.

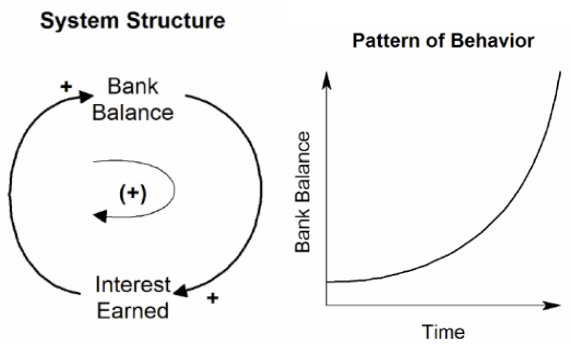
A causal loop diagram is good for presenting the basic ideas incorporated in a model in a manner that is easily understood, without having to discuss in detail. They shouldn't be used when it is important to determine the dynamics of a system.



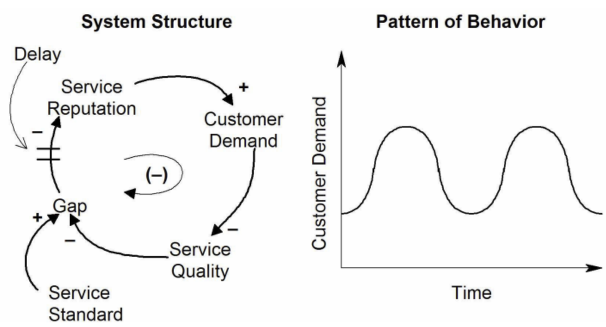
Feedback loops are logical chains of cause and effects in a system. The basic modes of behavior in dynamic systems are identified along with the feedback structures generating them. It is used when an element of a system indirectly influences itself.

Positive feedback loops generate processes like growth, amplify deviations and reinforce change. Negative loops are acting to bring the state of the system in line with a goal (defined by gap) or desired state by counteract any disturbances that move the state of the system away from the goal (balance, equilibrium, stasis)

Positive Feedback Loop



Negative Feedback Loop (with delay = Oscillation)



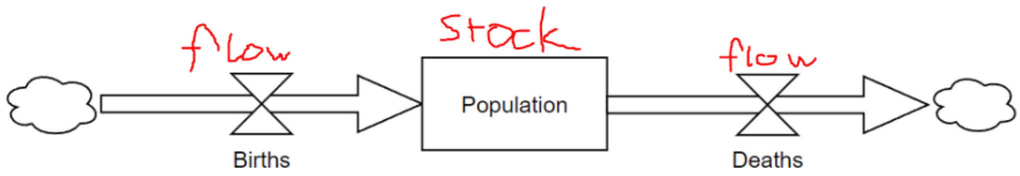
9. What is a stock and what is a flow? Give three different pairs of examples (i.e. one stock and one corresponding flow)! Give examples of common misconceptions of stock and flow (including sketches)!

A stock is a quantity at a certain point in time that is accumulated over time by inflows and/or depleted by outflows, for example: Population, Public debt, Account balance
A flow variable is measured over an interval of time (eg: a year), is a quantity that increase (in-flow)/decrease(out-flow) a stock, for example: Births/Deaths, Year‘s budget deficit, Deposit/Withdrawal
Here are the guidelines that can be used to help identify stock and flows:

- Stocks usually represent nouns and flows verbs
- Stocks do not disappear if time is (hypothetically) stopped and flows do
- Stocks send out signals (information about the state of the system) to the rest of the system.

Examples:

- Capital (stock) inc. or dec. by investment/depreciation (flows)
- Population (stock) inc. or dec. by births/deaths (flows)
- Guests in a hotel (stock) inc. or dec. by arrival/departures (flows)



An example of a common misconception of the meaning of stock and flow will be that people say "deficit" (a flow) when they mean "debt" (a stock), or they'll say the "inflation (a flow into a stock) is lower therefore the general level of prices (a stock) are falling". In fact, decreasing inflation to a lower value means that prices are rising but at a slower rate.

10. What is a system dynamics simulation model (compared to causal loop and stock flow diagrams)? How is such a simulation developed (simple words)? What are such simulations used for? Which kind of systems/processes are they well suited to simulate?

A system dynamic model is a computer-based mathematical modeling approach for policy analysis and design that helps to understand the behavior of complex systems, using stocks, flows and internal feedback loops.
Mathematically, it is a system of coupled nonlinear first order differential (or integral) equations

The simulation is developed through these steps:

1. identifying the most important stocks and flows that change
2. identifying sources of information that impact the flows
3. identifying the main feedback loops.

These simulations are used for understanding the sources of policy resistance and designing more effective policies by studying social systems, capturing his structure and predicting short run/ long run effects, this is the quantitative and qualitative analysis and what-if scenarios.
The elements in SD are homogeneous (no individuals possible), the simulated system's structure stays static during the simulation.

Example for SD: Predator - Prey model

SD works good for a more abstract and simpler simulation of complex systems that have unknown "effects" that we cannot consider. Instead of including every little detail in the model, and being anxious that we do not forget anything, we can just model the aggregation e.g. we can just model the birthrate of a population rather than trying to model the underlying phenomenons or interactions between individual agents.

11. Which types of fundamental modes of dynamic behavior can you think of (Chapter 4 of the Business Dynamics book)? Which of them do you think are likely to be found in natural systems in the long run – give examples of systems and their qualitative behavior)!

- Exponential Growth (positive (self-reinforcing) feedback)

Needs a positive feedback loop to occur and will keep on amplifying initial changes and the rate of change and therefore lead to reinforcing feedback and therefore exponential growth.

- Goal seeking (negative feedback)

Needs negative feedback loops to create balance and reach a state of equilibrium. Negative feedback loops can therefore bring the state of a system to a desired state or goal. If the state of the system does not match the desired goal, corrective action takes place until the desired state is reached.

- Oscillation (negative feedback + delay)

It is also caused by negative feedback loops. Also like the goal seeking mode the system wants to reach a desired state, and therefore a corrective action is taken to move closer to the goal. However, due to a time delay in the system the corrective actions are carried out even after the desired state has been reached, causing a new corrective behavior in the other direction.

Oscillation is the most common pattern found in natural systems, especially when thinking of physics there are a lot of examples: a pendulum, waves of different kinds (e.g. sound, ocean, electromagnetic, etc.), string instruments and more.
Exponential growth can be found in natural systems as well but is not as common as oscillation, for example the expansion of the universe.

12. From a system-theoretic view: What is the danger of a policy that only regards the short run? Why are complex socioeconomic systems hardly predictable (especially when seen from an SD perspective)? Give an example of a social system where the short run differs from the long run and why this is problematic!

A Policy that only focuses on short term effects, can cause long run effects that were not taken into consideration by policymakers. This is caused by the delay of feedback within the systems, and it can take a long time to see the unintended consequences.

Some reasons why socioeconomic systems are hard to predict:

- It is not clear what the symptoms of a problem really are because they are separated from the actual problem by time and space
- Human intuition is sometimes wrong about the expected behavior
- Policy intervention sometimes only focuses on either short term or long term effects
- External policy intervention is neutralized or counteracted by the system's internal feedback
- External shocks can hardly be predicted within the model and therefore the focus should not be on factoring them in but to structure the system to withstand uncertainty
- In the real-world, there is no equilibrium in complex systems, but continual change

Complex socioeconomic systems are hardly predictable, seen from SD perspective:
Because of the complex, nature changes could influence the rules of the model itself and this could not be modeled successfully by just using SD models because of their static structure and the lack of the possibility to adapt to changes.

Example: Road building policy to avoid traffic jams
A policy is put in place to invest and build better road infrastructure because of traffic jams. In the short term, this might have a positive effect with less traffic jams and faster commuting times due to more and bigger roads being available. Over time, due to the convenience of using this better infrastructure, more people rely on commuting to places by car and therefore start to create more traffic. In the long run the city will have more traffic than before with more space used up for road infrastructure but will have the same problem as before or even worse on roads that were not renewed to meet the new demand. This is problematic because it leads to the exact same problem as before, but now in an even bigger dimension due to the higher amount of cars.

Question Group MCS (Microsimulation, Cellular Automata and Social Networks)

13. What are microsimulation (or microanalytical simulation) models? What are they used for (give examples) / What are they good at? In what way does the microsimulation approach differ from the system dynamics approach?

Microsimulation is the first modelling technique that operates at the level of individual units such as persons, households, vehicles, firms…
Within the model each unit is represented by a record containing a unique identifier and a set of associated attributes (age, sex, marital-, employment status, income, tax payments, etc.)

A set of rules (transition probabilities) are then applied to these units, leading to simulated changes in state and behavior.

- deterministic (probability = 1): e.g. changes in tax liability resulting from changes in tax regulations
- stochastic (probability ≤ 1): e.g. chance of dying, marrying, giving birth or moving within a given time period.

In either case, the result is an estimate of the outcomes of applying these rules, possibly over many time steps, including both total overall aggregate change and the way this change is distributed in the population or location that is being modeled.

It is often used to evaluate the effects of proposed interventions before they are implemented in the real world, for example, a traffic microsimulation model could be used to evaluate the effectiveness of lengthening a turn lane at an intersection, and thus help decide whether it is worth spending money on actually lengthening the lane

In econometrics, it is used to simulate the behavior of individuals over time. The microsimulation can either be dynamic or static. If it is dynamic, the behavior of people changes over time, whereas in the static case a constant behavior is assumed. There are several microsimulation models for taxation, pensions, and other types of economic and financial activity

It is also used in health sciences and spatial fields, that allow the characteristics of individuals living in a particular area to be approximated, based on a set of constraint variables that are known about the area

The difference: in system dynamics one can model only at macro-level and can't in different levels and go towards in simulation. In microsimulation, it is possible to interact between individuals in the same and in the different levels.

14. Which kind of elements does a microsimulation consist of? How are microsimulation models initialized? What kind of data is needed/used for these simulations?

Elements:

- Individuals (Households, vehicles, firms, etc)
- Rules aka transition probabilities
    - Deterministic rules (p = 1, e.g. tax brackets)
    - Stochastic rules (p < 1, e.g. chance of dying)

Initialization based on representative (samples) empirical data

- Take real world sample
- Calculate distribution

15. What characterizes cellular automata models? What does a cellular automata model consist of (describe)? Give an example how a simple cellular automata model might work!

Cellular automata is a discrete model of computation to show self reproductive behavior
A cellular automaton consists of a regular grid of cells that have a finite number of states.
For each cell, a set of cells called its neighborhood is defined relative to the specified cell. An initial state (time t = 0) is selected by assigning a state for each cell. A new generation is created (advancing t by 1), according to some fixed rule (generally, a mathematical function). That determines the new state of each cell in terms of the current state of the cell and the states of the cells in its neighborhood.
Cellular automaton can compute functions and solve algorithmic problems, emulate Turing machine and can be used in social simulation due to the spatial possibilities.

One example is the Conway's Game of Life, is a zero player game (=NPCs only, you can just watch, basically a movie), that is, the state of the system depends exclusively on its initial state. It takes place on an infinite two-dimensional grid, in which cells can be 'on' (alive) or 'off' (dead). Each cell interacts with his eight neighbors (the cells that are vertical, horizontal and diagonal adjacent). The time is simulated discretely and cells are updated once per time step. There are some cell's rules:
- Under population: every cell with less than two live neighbors dies
- Overpopulation: every cell with more than three live neighbors dies
- Reproduction: every dead cell with three live neighbors become a live cell
- Any live cell with two or three live neighbors lives on to the next generation.

16. Are cellular automata applicable for social simulations? What are they good at? What are the limitations of cellular automata models?

Cellular Automata models are not so relevant in computational social sciences as there are available more elaborate techniques (ABM, agent based model)
They are good at investigating the outcomes at the macro scale of millions of simple microscale events, and it can be calculated very fast due to simple structure
The limitations depend on the CA variant

- In the more classic variants that have only 1 state, all the cells are identical with passive interaction (no movement), the problem is the extreme variance of results with small changes in the initial configuration, rule or timing

- But in all variants the spatial arrangement is always a central issue

17. What are checkerboard models? Give an example for such a model! Are these checkerboard models CA models or ABM models and what do they have in common with these techniques?

The checkerboard model is a computer simulation of social interaction among members of two groups. The checkerboard represents a social field on which two groups of checkers move on the board on the basis of positive, neutral or negative attitudes toward one another assigned to them. The resulting pattern of positions of the pieces represents the social structure
The checkerboard model is capable of demonstrating the intimate connection between attitudes of group members toward their own group and toward others to a continuous social interactional process and to the resulting social structure.
They are and advancement of CA and the precursor/simple form of ABM
An example of checkerboard model is Sakoda's model or the Schelling's model of segregation

18. What is a social network? Also give examples! How can social networks be represented / analyzed (graphically and analytically)? What is social network analysis?

A social network is a social structure that can be represented by graphs, this is a set of nodes (social actors like individuals, firms, organizations..) and edges (ties, birth, marrying, relocating, finding friends) between these nodes that form social contacts. The sum of all (incoming and outgoing) ties of a unit represents the convergence of the various social contacts of that unit.
Social networks are dynamic, self-organizing, emergent, and complex phenomena, they can be either real life or digital; for example a social network of friendship like Instagram or Facebook

Social Network Analysis is the process of investigating social structures and understanding social interaction through the use of networks and graph theory via:

- a static analysis, this is a snapshot of a Social Network's Structure at a point in time, E.g. Sociogram
- a dynamic analysis, investigating the rules of changing micro- or macrostructures

19. What is the small worlds phenomenon (what does it describe)? Describe the initial experiment that showed/measured the phenomenon! How is it connected to power-laws? Which kinds of networks are likely to show the small worlds phenomenon? How can a small worlds network form?

The small world phenomenon was found out while examining the average path length for social networks of people in the United States. The research was groundbreaking in that it suggested that human society is a small-world-type network characterized by short path-lengths, it is normally associated with the phrase:" six degrees of separation".
Milgram's experiment developed out of a desire to learn more about the probability that two randomly selected people would know each other. Imagine the population as a social

network and attempt to find the average path length between any two nodes. Milgram's experiment was designed to measure these path lengths by developing a procedure to count the number of ties between any two people.

Social networks are an example of the small world phenomenon

The small world network initially begin with some people that are socially very active, as previous “social success“ is an attractor, eventually, those with many friends eventually get more, these people eventually form hubs – few hubs are enough ; and also retain long-term connections; already highly interconnected individuals get more connections

The distribution of links to nodes in a small-world follow a power-law. The power law means that the vast majority of nodes have very few connections, while a few important nodes (we call them Hubs) have a huge number of connections.

20. What are strong ties and weak ties in social networks? Why can weak ties be important – explain and give examples! What does this have to do with network structures?

Strong ties are close family and close friends, they have a strong influence on decisions, common behavioral norms and tend to move in common circles; high overlaps of information

Weak ties refer to the acquaintances, remote family and colleagues, they tend to move in other circles, have access to “new” information and build bridges to other parts of the wider social network (society)

However, these weak ties are crucial in binding groups of strong ties together. They bring circles of networks into contact with each other, strengthening relationships and forming new bonds between existing relationship circles.

21. Explain the following measures and how they are calculated: Characteristic Path Length & Clustering Coefficient!

Characteristic Path Length L:

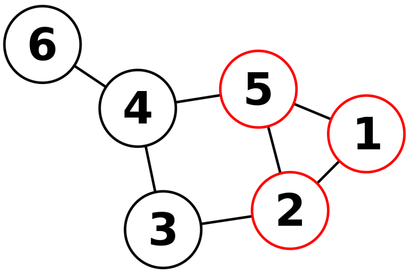
- Calculate the shortest path (number of edges) from each node to every other node
- L is the average of all these shortest paths

→ L = Average shortest path length

$$l_G = \frac{1}{n \cdot (n-1)} \cdot \sum_{i \neq j} d(v_i, v_j)$$

Clustering Coefficient:
Is a measure of the degree to which nodes in a graph tend to cluster together
Evidence suggests that in most real-world networks, and in particular social networks, nodes tend to create tightly knit groups characterized by a relatively high density of ties; this likelihood tends to be greater than the average probability of a tie randomly established between two nodes

How close are nodes to being a clique with its neighbors.
A clique is a subset of nodes such that every two distinct nodes in the clique are connected to each other.



The red nodes are a mfing gang (clique)

if a node has kv neighbors, the local clustering coefficient is:
ci = 2 Nv / kv (kv - 1)
where:

- Nv = **actual number of links** that exist between neighbours of v
- kv (kv - 1) / 2 = number of **links that could exist** between neighbours of v

and the global clustering coefficient is the average of all ci, this is number of closed triplets/number of all triplets

Question Group ABM (Agent Based Modelling Questions)

22. What are the main features (modelling elements) of an agent-based model - what characterizes an agent-based model? How does an agent-based model differ from other simulation models/methods?

Agent based modeling (ABM) is a modeling concept, where individual agents interact with each other in some kind of environment. It is more of an idea than a fixed modeling approach, and is generally very flexible.

<u>Main features:</u>

- *Agents:* represent social actors, have individual attributes like age, and interact with each other. E.g. a person, household, company, country.
- *Environment:* The agents all exist in an environment, which has some properties (e.g. grid, geographical map) and can interact with agents (e.g. taxes).
- *Bounded Rationality:* The agents cannot sense everything and know everything, their senses are limited to some extent.
- *Interactions:* The agents have some behavior, and interact with each other and the environment.
- *Learning:* Agents are able to learn from experiences. There are 3 kinds of learning. (1) The individual agents learns, (2) The population as a whole learns, because agents teach other agents their experiences, and (3) the whole population learns through natural selection, meaning “dumb” agents die
- *Ontological Correspondence:* The agents belong to a certain category, which also exists in the real world e.g. persons, students, households, companies, governments.

<u>Differences to other approaches:</u>

Cellular Automata (CA) only allow for 1 attribute per cell, have a restrictive environment (e.g. cells can only be on a grid), and the cells adhere to global rules to determine the next state of the model. In AGM the individual agents can have multiple attributes, are able to interact with each other and the environment, and thus can even influence the rules of the environment and behavior of other agents. It also allows modeling the relationships with other agents very flexibly with graphs.

Microsimulations (MSM) offers, similarly to ABM, some heterogeneity between individual actors. But the rules are fixed globally. However ABM allows for more complex behavior, and the behavior of the agents might even influence the rules and vice versa. ABM also offers the individuals to learn from experiences.

System Dynamics (SD) aims to model complex, isolated systems from a bird's eye view. So in comparison to ABM there are no individual actors and ABM. One important aspect of SD is that the model can influence itself through feedback. However this is limited, and the individual elements cannot adapt to changes. ABM is very similar, but more powerful, since all the components of ABM (actors, environment, …) can influence each other. Thus with ABM the rules and behaviors of the simulation can change overtime.

23. What are the most important aspects of agents in the context of ABM models? Furthermore, describe the main characteristics that agents may have in an ABM model!

Agents represent social actors, have individual attributes like age, and interact with each other. E.g. a person, household, company, country.

Main characteristics of an agent:

- *Perception*: Agents who are able to perceive the surrounding environment.
- *Performance*: Agents can perform some behavior. They move within the environment, can send and receive messages from other agents, and can interact with the environment like picking up food.
- *Memory*: Agents can remember experiences.
- *Policy*: Agents have rules and heuristics that determine what behavior they perform.

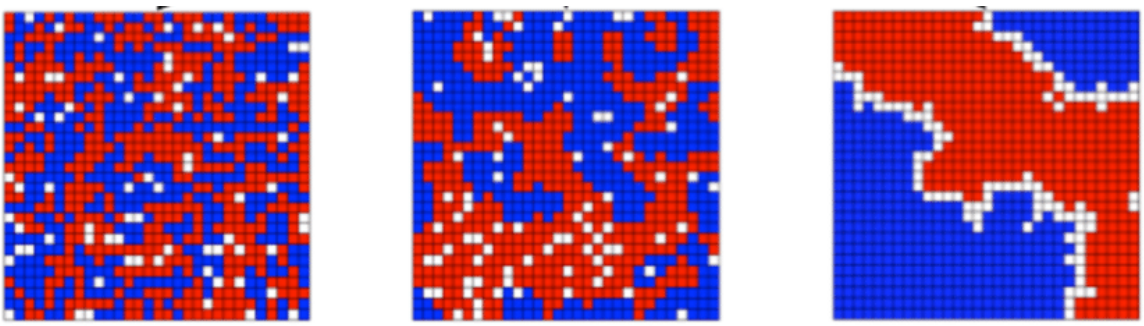
24. Describe Schelling's Segregation Model, how and what it tries to explain. What does the model predict? What did Schelling use to simulate it originally (computers?)?

Schelling's Segregation Model aims to explain how different ethnicities of people often group together in US cities. This grouping phenomenon is also called *segregation.*

The model consists of a grid, on which people of 2 ethnicities are placed on. These individuals want a certain share of their neighbors to have the same ethnicity as them. The share is fixed and also called the tolerance value (or Threshold). If the individuals are not satisfied, they move to an empty space on the grid.

The model found that even if individuals only want a very small share of their neighbors to have the same ethnicity, this is enough to lead to segregation.

Originally Shelling simulated the model with a checkerboard as the grid, and with zinc and copper coins, for the two ethnicities of people.



15% Threshold   30% Threshold   75% Threshold

25. Describe the simple bird-flocking model! What does this simple model show (i.e. why is it interesting for our purposes)? What are the behavior rules (micro-level behavior) of the birds and to which behavior do they lead (which macro level regularities)?

It shows how complex behavior (self-organization) can emerge from simple rules.

3 simple rules of the individual birds:

- Alignment: adapt direction to the nearest bird
- Separation: keep minimum distance (prevent collisions)
- Coherence: approach birds in sight (limited)

Behavior: Complex Movements/Patterns of the swarms on the macro level.

26. Describe the demographic prisoner's dilemma (DPD) from the second assignment! Describe the rules of the game and the expected dynamics – especially the differences to the regular prisoner's dilemma (non-iterated, non-demographic).

In the demography, there are two types of individuals

- cooperators who cooperate
- defectors who act selfishly

PLAYER A \ PLAYER B	COOPERATE	DEFECT
COOPERATE	A: 1 year jail B: 1 year jail	A: 10 years jail B: 0 years jail
DEFECT	A: 0 years jail B: 10 years jail	A: 5 years jail B: 5 years jail

The individuals interact with each other according to the classical prisoner's dilemma. Over time, they can gain or lose wealth according to their interactions (prisoner dilemma matrix). If they gain enough wealth they can reproduce and if they lose too much wealth they can die out.). Prisoners have a fixed role (confess or cooperate) and an initial position on the map. With the initial parameters, there are more cooperators. In the traditional dilemma, the confessing strategy is dominant.

In the classical prisoner's dilemma, there are just two individuals who try to find an optimal solution for minimizing their jail time. After the two individuals made their separate choice of either cooperating or defecting, there is an outcome state and the game is over. The demographic prisoners dilemma consists of multiple individuals and goes over multiple rounds (iteration), and the goal of the individuals is wealth maximization.

Not finished yet, or is it?

27. Why is randomness needed in most agent-based models – i.e. which problems does a random element address. What problems might arise due to the random element? How can these problems be alleviated again?

Randomness is very useful:

- emulate noise/uncertainty (e.g. Flawed communication, errors of judgment, random decisions)
- Sometimes values can't be observed (e.g. unknown attribute values of the agents)
- For the initial layout (e.g. spatial position)

Since we now have random elements, this means every run of the simulation will be different → single run might not be so helpful, repeated runs (e.g. 100 times)

28. What are experimental designs with regard to computer simulations? What kind of designs did we mention in the lecture? Which experimental design does a basic Netlogo Behaviorspace experiment use? How could we employ other experimental design techniques?

Experimental Designs:
- Factorial Design:
  - *Full factorial design:* every combination of each meaningful parameter values, the default approach
  - *Factorial Design:* Specific partitions of the multidimensional parameter space
- Random Design:
  - *Latin Hypercube design*: Random but meaningful combinations
  - Many other alternative methods

Behaviourspace = built-in Netlogo tool for unattended simulation runs
- Uses only **full factorial designs** (disadvantage: huge number of runs)
- Runs a model multiple times with a different model's settings and records the results of each test run

Other more complex experimental designs can be run using the R-Netlogo package
- Allows running Netlogo simulations from within R
- Allows analysis during runs

29. What is the difference between model verification and validation regarding agent-based models? Describe both shortly, then explain techniques for validation – which techniques are needed for which type of agent-based model (three types were mentioned with respect to scale)!

Verification & Validation:

Model Verification is used for checking if the model was programmed correctly and does not contain any bugs.

Model Validation is used to check if the "right" model was created. In other words, we want to check if our model even describes the social phenomenon we want to study, or if we created a model that unintentionally describes something different.

Techniques for model validation:

1. *Sensitivity Analysis:* Is used to determine how good the model and theory fit together. This is done by testing the model with different parameters and analyzing the output. Due to the randomness of ABM the model should be run multiple times with the same settings. Since there often are also many parameters and possible values for these parameters, often a huge number of runs have to be done. Using knowledge

about the parameter values like a practical minimum value, can be used to reduce the runs. Additionally it might be beneficial to randomly sample the parameter values to further reduce the amount of runs.

2. *Empirical Validation:* Another approach is to compare the model with empirical data. This might not work for every model, since sometimes empirical data is just not available, especially when doing predictions.

<u>Which validation fits to which model:</u>

- Abstract Models (the very general bois):
  - check if the general phenomenon makes sense
- Middle Range Models (the in between bois):
  - empirically check if the *overall results* of the model similar to the phenomenon we want to observe (qualitative similarity)
- Facsimile Models (the very detailed bois):
  - empirically check if *the values* of the model are similar to the values we expect in the real world. (quantitative similarities)

<u>Types of Models:</u>

- Abstract Models:
  - Are used to demonstrate some general social process with the goal of getting overall patterns from some simple behavioral rules. The results can then be used to further develop theories and maybe create more specific models
  - An example would be Schelling's segregation model, which shows how ethnic segregation in US cities could emerge, based on some simple rules. The model uses very simple rules and is not specific to any US city.
- Middle Range Models:
  - Try to describe the characteristics of certain social phenomena, and are more specific than Abstract Models. They should still offer some generality, so that the insights can for example be applied to not only 1 industrial district in a certain city, but generally to all industrial districts
  - An example would be the modeling flow of knowledge in "innovation networks". In this model companies trade with each other, and knowledge can be gained through internal research or by using knowledge obtained from other companies. With better knowledge these companies are able to make better products. Analyzing the number of partners a company has in such a model shows that the number follows the power law, a very common empirical phenomenon seen in social networks.
- Facsimile Models:
  - Are very detailed models that aim to describe a quite specific social phenomenon as exactly as possible. Since they are so detailed, they can be used for reproducing historic developments and for predicting future developments. However, the real world still has some element of randomness, so the results should not be seen as an absolute truth.
  - An example would be a model of the inventory of a business, with the goal of evaluating the consequences of restocking their items a bit later or earlier.

30. Shortly describe sensitivity analysis? Why is sensitivity analysis needed? What are the main challenges of conducting a sensitivity analysis? What is the connection between experimental designs and sensitivity analysis

Sensitivity Analysis is used to determine how good the model and theory fit together. Basically to test how robust a model is. This is done by testing the model with different parameters and analyzing the output.

<u>A big challenge is the number of runs:</u>

1) Due to the randomness of ABM the simulation must be run multiple times with the same settings.
2) This has to be done for each and every parameter setting we want to test

=> huge number of runs

<u>Restrict number of runs:</u>

- Using knowledge about the parameter values like a practical minimum value, can be used to reduce the runs.
- Additionally it might be beneficial to randomly sample the parameter values to further reduce the amount of runs.

<u>experimental designs and sensitivity analysis</u>
The are quite similar both use empirical knowledge to decide what we want to simulate (which parameters, which parameter ranges, etc)