Chapter 3
Agent-Based Simulation in Archaeology: A Characterization

Felix Breitenecker, Martin Bicher, and Gabriel Wurzer

3.1 Motivation and Classification of Computer Models

The phrase "dare to know" (Latin "*sapere aude*", originally brought forward by the Roman poet Horace and later made popular by Kant (1784)) is famously known as the motto of the enlightenment, representative for the intellectual attitude of the western civilisation during seventeenth and eighteenth century. From this era also come some of the core concepts used in simulation, *experiment* and *observation*: An experiment is a procedure that seeks to prove or reject a hypothesis (Kant 1787, p. 23) or observe an emergent outcome. However, there are cases in which questions cannot be answered by conducting classical experiments, because

- these would be too expensive (e.g. investigate the earthquake resistance of a skyscraper)
- these would be too dangerous or not possible for ethical reasons (e.g. investigating the spread of an epidemic)
- these would not be possible for logical reasons (e.g. measuring the kernel temperature of the sun)
- these would not be possible, as information from the future is required (e.g. meteorological weather prognosis)
- these would not be possible, as information from the past is required (e.g. proving or falsifying historical hypotheses if archaeological findings are missing)

Therefore, instead of conducting "real" experiments, researchers have turned to *experimentation in silico*, in which *simulation* plays the role of a tool (or method,

F. Breitenecker (✉) • M. Bicher • G. Wurzer
Vienna University of Technology, Karlsplatz 13, 1040 Vienna, Austria
e-mail: felix.breitenecker@tuwien.ac.at; martin.bicher@tuwien.ac.at;
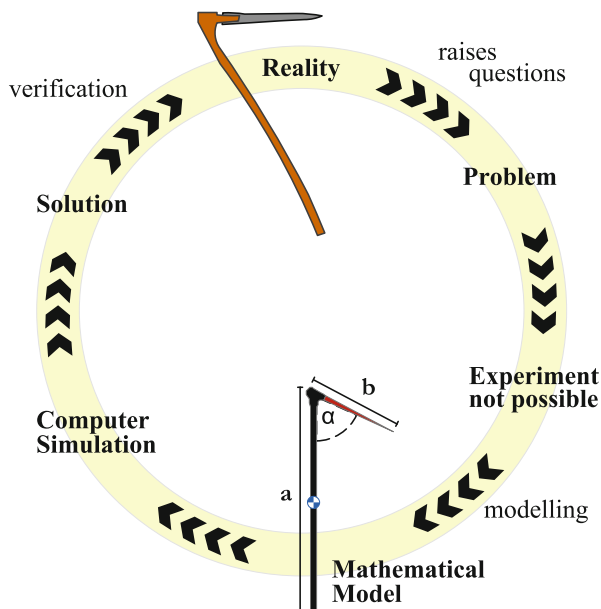gabriel.wurzer@tuwien.ac.at

Fig. 3.1 Motivation and life cycle of mathematical modelling

see Fig. 3.1): A real entity raises questions, which leads to a problem that is either defined explicitly or (as it often is in archeology) informally. The problem cannot be answered using classical experimentation, and therefore, a mathematical or theoretical model is built, which imitates the behaviour and structure of the real system as closely as possible. Based on this model, a simulation is run to solve the given problem, supported by initial data in order to receive "satisfying" results. The level of detail and precision within the modelling process and the amount and quality of the initial data (consisting of input and initial conditions of the simulation) are mainly responsible for the validity of the simulation results and thus the reliability of the answer to the given problem. However, this does not free the researcher of a verification phase, which examines how congruent the model is in respect to reality. A closer look at this aspect is given in Chap. 2.

3.1.1   The Model: Creating an Abstraction of Reality

The word "model" derives from the Latin word "modulus", which was used as a measure in architecture until the eighteenth century. To be more precise, "modulus" was the base unit used in architectural planning on which plans and miniature versions of buildings were based. Its meaning slowly transitioned into the now-common "model", which signifies a simplified replica of the real world. With respect

to Stachowiak (1973), a model is characterized by at least three different features (translated from German):

1. Mapping—A model is a representation of a natural or an artificial object.
2. Reduction—A model is simplified and does not have all attributes of the original object.
3. Pragmatism—A model is always created for a certain purpose, a certain subject and a certain time-span.

Similarly to models made out of real material (wood, stone, paper,...), mathematical models used for simulation also fulfil these three characteristics: On the one hand, a detailed mapping of the real world guarantees reliable results. On the other hand, it is still necessary to do simplifications in order to be able to perform the simulation at all. Although processing power and memory resources of computers are increasing rapidly, the ability to calculate simulations in-depth is still limited (also see high-performance computing, Chap. 4): One must not forget that the more complex the model gets, the more data is required to calibrate the simulation correctly. Regarding reduction, the following techniques can be used to obtain reasonable models:

- Distinction: Neglect those parts of reality which are not relevant for the purpose of the model.
- Cutback: Neglect details that are irrelevant.
- Decomposition: Decouple segments of an object which are not or weakly linked and treat them as own models.
- Aggregation: Couple segments which act alike and threat them as one.
- Abstraction: Create technological terms, classes and clusters in order to simplify the overall problem.

Surprisingly, the history of mathematical modelling is older than the history of the computer. For example the famous Lotka-Volterra Differential equations, a mathematical model formerly designed to simulate the cohabitation of a predator and a prey species, was initially developed by Alfred J. Lotka and Vito Volterra between 1925 and 1926 (Lotka 1956). Compared to that, the first (useful) computers were developed during the Second World War. Surely, the two scientists had problems to simulate their own derived differential equation model by hand, but they had some success regarding the qualitative analysis of their results. So, the idea of simulation is basically not linked to computers, although most simulations can only be executed with their support. Even after the introduction of computers, some simulations such as the evacuation model by Predtetschenski and Milinski (1971) were based on hand calculation methods, and some are even used to this day (Rogsch et al. 2008).

One of the most common definitions of the modern meaning of the word simulation is finally given by Shannon (1975): "Simulation is the process of designing a model of a real system and conducting experiments with this model for the purpose either of understanding the behaviour of the system and its underlying causes or of evaluating various designs of an artificial system or strategies for the operation of the system" Shannon (1975, p. 2).

Shannon defines the modelling process as a sub-process of the simulation, as the purpose of the model is solely given by the purpose of the simulation. On the one hand, it can be seen easily that the word computer is not part of this definition. On the other hand, since the invention of the computer the fields of application for mathematical models is steadily increasing as due to growing computational resources, a growing number of complex simulations can be executed.

Agent-based modelling is just one of many modelling approaches. In order to fully understand the basic mechanics of this specific approach, some specific terms related to modelling and simulation have to be discussed in the forthcoming sections. Very often, the explanation of these terms is supported by example models. These examples were only developed theoretically (i.e. for didactical reasons) and their validity and value must thus not been overestimated.

3.1.2   Black-Box and White-Box Models

Within system theory, the word *black-box* describes a basic input-output system in which the correlation between input and output is not known. Such an "opaque box" somehow translates the input signal into some output and does not reveal how the translation takes place. Regarding modelling and simulation, the term black-box is used if modelling approaches are mainly based on observations and explanations but not on strict laws and proven formulas. The model has to be developed only knowing about input data, a little bit of information about the behaviour of the real system and some output data required for the calibration of the model. If done properly, those kinds of models are also working reliably, even though everything in between the input and the output part can hardly be explained in a reasonable way; nevertheless, there are some authors who seek to reverse-engineer black-box models, in order to give users confidence and also some control over the work they are doing with them (Rogsch and Klingsch 2011).

White-box modelling, i.e. modelling based on proven laws and axioms, is preferred to black-box modelling. When this is not possible, a mixture between the approaches has to be used. The gradual transition between white-box and black-box techniques can be seen in Fig. 3.2b: Most problems that can be described by proven laws or axioms can be simulated using white-box models, whereas problems that are based on expert knowledge and observations will need to be approached by black-box modelling:

• *How fast does the pendulum swing?*
  This problem can be solved using physical basics like Newton's law of gravitation, friction formulas and the law of conservation of energy, resulting in the so called pendulum-equation. This second order differential equation can be solved by most of the now-common numerical computing and simulation packages. The derivation of the model was solely based on physical laws, which makes it a *White Box model*.
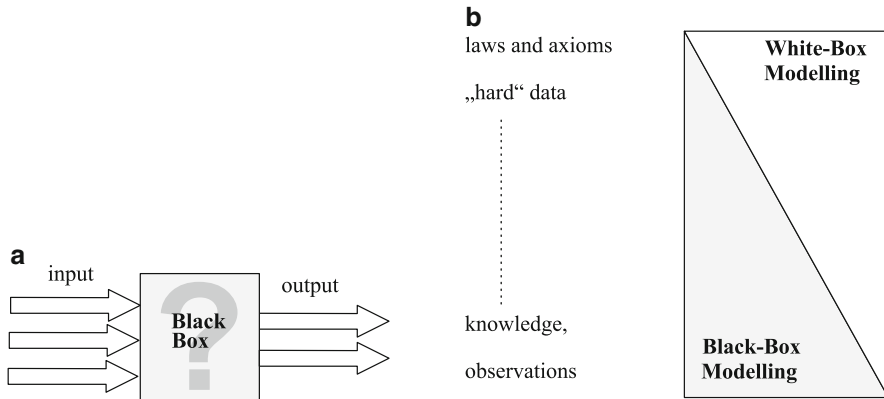
Fig. 3.2 (**a**) Black box model. (**b**) Fields of application for black box and white box models

- *How fast does a population grow?*
  Before being able to solve that problem, a lot of data has to be procured and investigated. Aside from the data, it can be observe that the growth rate of the population is proportional to the number of individuals (the higher the number of fertile individuals, the more individuals are born). Furthermore, the size of the population is limited due to resources and lack of space. In order to model these constraints governing the size of the population, a so-called logistic differential equation can be applied. The modelling process is basically motivated by data and observations, and thus is an example of a *Black Box model*.

Summing up, both examples result in mathematical formulae. The main difference being the derivation of the equation: In the first case, the formula is solely based on already-existing, proven laws. In the second case, the equation is gained by observation of the behaviour and data of a real system. Agent-based models are mostly used as black-box models, but are not limited to this field of application. Before we continue to the description of basic concepts used in agent-based modelling, some principles of microscopic and macroscopic modelling have to be discussed.

3.1.3 Microscopic and Macroscopic Approach

The larger and more complex the investigated system, the higher the need for a computer simulation (instead of a classical solution to a problem). Usually, very big complex systems consist of (or can at least be partitioned into) a large number of less complex subsystems which can be investigated in isolation. The final complexity of the overall system is governed by the interaction of all subsystems (also called *components* of the system). If those subsystems share the same basic properties and act similarly, the system can be described by an agent-based model. Examples for

Table 3.1 Examples of
systems made out of similar
components

System	Components
Population	Individuals
Company	Subsidiaries
Traffic	Cars
Paper	Wooden fibres

such systems and their components are given in Table 3.1. Interestingly, the scale at which these systems are modelled may differ, as we may distinguish models as being either *microscopic* or *macroscopic*:

- *Microscopic models* decompose a system into its subsystems. Each of these subsystems is modelled separately, taking into consideration that they are allowed to communicate with each other, in order to model the structure of the whole system. Because the subsystems are similar, they are usually modelled in the same way, but then supplied with different parameters. In order to decouple the system, simplifications have to be done regarding the often complex relationships between the subsystems. These rich relationships also make "the whole [...] greater than the sum of its parts" (Aristoteles), a property that is often referred to as *emergence*.
- In *Macroscopic models*, the influence of the subsystems upon the whole system is investigated and generalized, in order to be able to model the whole system at once. Often, black-box approaches have to be used in that context, leading to so-called "behavioural models" (i.e. models developed by studying the behaviour and causal relationships of the system). Although these models do not take into account that the whole system consists of subsystems, very satisfying results can be obtained if the model is developed properly.

The choice of the modelling approach is often a result of a "top-down" or "bottom-up" philosophy being chosen for the abstraction of the model. Usually, a "bottom-up" approach implies a microscopic model, in which the lowest level of the system (i.e. the "individuals" having the highest degree of detail) is analyzed first. After building up such a microscopic model, it is not unusual that the modeller has to simplify it, as it would otherwise become too complex. Vice versa, a "top-down" approach usually leads to a macroscopic model. The abstraction process then starts at the top level (looking at the whole system), and iteratively goes into more detail until the granularity is sufficient for the stated problem.

In order to further illustrate the difference between microscopic and macroscopic models, let us assume an example where the problem is to simulate the degrees of a large number of tourists visiting a museum after the occurrence of a fire. The microscopic approach would regard the whole system (museum) through its constituent parts (the visitors). These would be simulated individually insofar as each simulated person is free to choose his own exit route. The interaction between the individuals would here mean the movements, collisions and avoidance strategies that the visitors follow (also referred to as *pedestrian dynamics*), leading to emergent outcomes such as lane formation, congestive bulks before exits and collective

motion among the crowd (Helbing and Molnár 1995; Helbing et al. 2000). A macroscopic model would identify the mass of people to behave like a liquid, tending towards the areas with the least pressure (which in our case would be defined to be the exits). In this case, one would use the so-called Navier-Stokes-equations, which are used within physics for modelling fluid dynamics, to arrive at a macroscopic model for the system.

Summarizing, the decision which of the two mentioned approaches is taken requires a more detailed specification of the modelling purpose. It also depends on the number and quality of provided measurements and data that is necessary for calibrating the model. Both modelling approaches have advantages and disadvantages, as summarized in Table 3.2.

As computational resources are increasing exponentially (Moore 1965), microscopic models are getting more and more popular. Agent-based models are probably the best-known example of microscopic approaches, but not the only one.[1] A recent trend also lies in *cohort simulations*, which is a simulation technique in which individuals sharing the same properties are clustered into "cohorts". This approach lies in between the macroscopic and microscopic level and is therefore called *mesoscopic*.

Table 3.2 Advantages and disadvantages of microscopic and macroscopic models

Property	Microscopic	Macroscopic
Processing time	Usually microscopic models require lots of computational resources. So calculations are usually rather *slow*	As not all sub-models have to be evaluated separately computation is rather *fast*
Abstraction level, i.e. how difficult the explanation of the model is to non-experts	As the connection between model and reality is very obvious the necessary level of abstraction is relatively *low*	As the model is very often either developed black-box style or derived out of complex formulas the connection between model and reality is more difficult to understand. So the abstraction level is usually *high*
Number of necessary parameters	Due to the high number of sub-models *a lot* of specific parameters have to be defined which poses a difficult task for calibration of the model	Usually the model is well defined by *a few* parameters dealing with the overall dynamics of the system. Regarding commonly gained statistical data calibration is often easier too
Flexibility	Small changes within the structure of the sub-models can be performed very easily without restructuring the whole model. Thus microscopic models are *very flexible*	Small structural changes usually require redefinition of the whole model (unless it can- be corrected by parameter variations). So macroscopic models are *hardly flexible*

[1]Others types would for example be cellular automata (see e.g. Wolfram 2002, p.8) or finite elements.

3.2   Introduction to Agent-Based Modelling

The term "agent" derives from the latin word "agere", meaning "to act". So, defined by the root of the word itself, the main function of an agent is to act— individually. Hence, we can give a first characterization: *Agent-based simulation is micro-simulation based on individually (inter)acting sub-models.* One aspect of this sentence is especially noteworthy: It does not state that agents necessarily have to pose for persons. Rather, a sub-model can represent everything ranging from a car in a traffic jam to a charged molecule in a liquid or even an abstract concept that can be represented as an agent. Regardless of what is being represented by an agent, there are still some additional requirements that need to be fulfilled, in order to completely define this modelling method. We thus come to a formal description that covers these.

3.2.1   Formal Description

Unfortunately, it is not easy to fix a unique formal definition for the term "agent-based modelling", in order to be able to make a clean decision if a given theoretical model should an agent-based or not. This is due to the "problem" that agent-based modelling (as well as many other modelling approaches) is a scientific method that used in a multitude of disciplines; as a correct formal definition of a method is always linked to the corresponding scientific termini of a field of application, problems occur if the same method is used within disparate academical areas. To illustrate this point further, a definition of a method developed by mathematicians, given in a mathematical language, would not satisfy the needs of computer scientists or social scientists (and vice versa). In the specific case of agent-based modelling, scientists from different fields have developed their own formal descriptions, using their own scientific termini. It is obvious that this has cased a lot of communication problems, especially at interdisciplinary conferences. In order to solve these problems, very rough definitions were developed, sometimes even directly within the context of those conferences by interdisciplinary discourse. One of these definitions, developed at the Winter Simulation Conference 2005 and 2006 (Macal and North 2006), states that:

Definition 3.1.   An agent has to

- be uniquely identifiable
- cohabitate an environment with other agents, and has to be able to communicate with them
- be able to act targeted

- be autonomous and independent
- be able to change its behaviour

The principle meaning of these five points is now explained in full detail:

- *Unique Identification*: Although different agents in an agent-based model do not necessarily have to have different properties, they have to remain distinguishable at any time during the simulation. More specifically, all actions of one agent have to be traceable after a simulation run.
- *Environment*: All agents of a model share an environment in which they can interact both with each other and "the world" around them. The most common way of doing the latter is movement. In fact, most computer scientists define the property "movement" of an agent as a basic requirement for agent-based modelling. However, the general definition developed at Winter Simulation Conference (2005 & 2006) states that there may also be non-movable agents living in a totally abstract environment. As example for this, one could imagine geographically fixed subsidiaries of a company interacting on an abstract electronic network (see Example 3.3).
- *Act Targeted*: Within the environment, agents have to act so as to reach a certain goal or purpose. So, one could say that they have to be given a basic form of "intelligence".
- *Autonomous and Independent*: It is a very difficult philosophical question how the independent and autonomous action of an agent is justified, if any action/reaction rule is predefined by the modeller himself. The point of the definition can be interpreted in the following way: If the fictive agent would be given a personality without knowing its creator, the agent has to think of its own actions to be performed *independently from the other agents*.
- *Behaviour*: Evoked by its target (see previous point "Act Targeted"), every agent displays its own behaviour regarding communication and actions. If necessary the agent has to be allowed to change it.

Although these five ideas might pose a sufficient definition to modelling experts already knowing about agent-based modelling, non-experts will probably be no wiser now than they were before. Regarding modelling and simulation, theoretical and abstract definitions usually do not provide any concept on how to practically convert these theoretical ideas into useful and working models, as they are solely developed to determine the affiliation or non-affiliation of a given model to a modelling technique. Especially for scientific papers and conferences, this kind of classification is essential. In order to satisfy the expectations of the (probably disappointed) reader, the modelling approach is now presented in a little bit more user-friendly way, by using a step-by-step instruction. This is done via directly developing some example models in the next sections.

Fig. 3.3 Agent-based
approach towards modelling a
road junction, as given in
Example 3.1
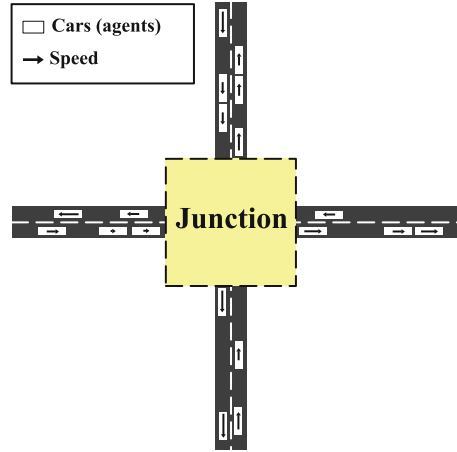
Example 3.1. Road Junction

Problem: A new road junction bringing together four streets (also see
Fig. 3.3) is to be planned. In that context, the responsible traffic planners
have to decide on either building a classical junction with traffic lights or to
employ a roundabout. The stated goal is to avoid traffic congestions as far
as possible. Experiments are not possible, so a modelling expert is called
in to answer on this question.

Abstraction: Analysis of the situation reveals that a microscopic modelling
approach is preferred, as the system can be reduced to consist of cars
(active parts of the world) and streets (passive). Since colour, shape
and type of the cars can be neglected, they can be reduced to appear
as line-shaped (one-dimensional) objects with a certain length, moving
individually on streets that are also line-shaped.

Model: Each car is represented by exactly one agent having a certain length.
To enable interaction between the agents, an environment has to be defined:
Each agent is assigned a place on one of the four one-dimensional streets
(to be precise, each agent occupies a spatial interval of its own length).
Additionally, each agent is assigned an initial velocity and a spatial target
(i.e. one of the four ends of the streets, which he tries to reach as
fast as possible). To control interactions between the agents, one has to
furthermore define how each agent adjusts its velocity to avoid collisions
with other agents while at the same time obeying the rules of the traffic
lights or roundabout. In our case, both of these two mentioned points
happen passively, by letting a would-be obstacle or traffic light send out:
"avoid colliding with me".

Example 3.1 (continued)

Conclusion: The developed agent-based model suits all five points mentioned earlier. Individually behaving agents (cars) are interacting in a goal-directed manner, within a specified environment—four linked one-dimensional lines (streets). Eventually, the success of the model will depend on the quality of the defined rules and the parameterisation. As a side-note, what we have just given as example is inspired by the famous Nagel-Schreckenberg model for simulating traffic and congestions on a one way street (Nagel and Schreckenberg 1992).

It can be observed that the strategic success of the modelling process before was precipitated by consequently answering the following questions:

1. Which parts of the system are expressed by agents?
2. What are the states of the agents? How do agents "look like"?
3. How does the environment "look like"?
4. How does communication between the agents take place? How and when do agents change their state?

To demonstrate the efficiency of these impulse questions another example is given in Example 3.2.

Example 3.2. Cave Lion

Problem: The extinction of the European cave lion (*Panthera leo spelaea*, see Christiansen 2008) is investigated. Archaeozoological findings prove the existence of the species in Europe until the onset of the first Lateglacial Interstadial, which occurred around 14.7 ka cal BP (see Stuart and Lister 2011). A possible hypothesis for the extinction is the "Keystone Herbivore Scenario" (Owen-Smith 1989), which states that the predator became extinct because of the extinction of its prey species. A second hypothesis goes into the direction of a reduction in prey numbers, possibly also connected to a geographical dispersal (Stuart and Lister 2011, p. 2337). In order to investigate these two hypotheses, a modelling expert is asked for help.

Abstraction: Based on estimations of the amount of lions and their prey, the key to an answer lies within the coexistence of two life forms, the lions and a prey species (i.e. reindeer and young cave bears, according to Bocherens et al. 2011; however, for the sake of the simplified example given here, we

(continued)

Fig. 3.4 Agent-based
approach (see Example 3.2)
modelling two cohabitating
species, cave lions (predators)
and reindeer (their prey)
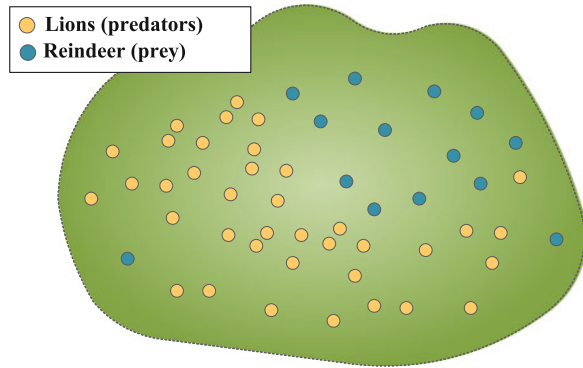
Example 3.2 (continued)

assume that the diet consisted only of reindeer). Most of the other envi-
ronmental factors can be reduced to influences on the birth and death rate
(directly or indirectly). As macroscopic approaches are usually unhandy
regarding the extinction of single animals, and consequently, a microscopic
approach is chosen.

Model: Each reindeer and each lion is represented by one agent. All reindeer
agents avoid predators as long as possible and try to reproduce themselves.
Each lion tries to catch prey in order to survive and to reproduce as
well. The two classes of agents can be simplified to being dots on
a map (see Fig. 3.4) trying to chase themselves. If height difference
does not influence the system, a two-dimensional map of the area is
a sufficient representation of the environment. Finally, rules concerning
movement, reproduction/death and the so-called "chasing radius" for all
lions complete the model definition.

Conclusion: The presented model with two different classes of agents defined
here representative of a class of models called Predator-Prey models.
This specific model was inspired by the "Wa-Tor" model (Dewdney
1984) simulating sharks and fish on a toroidal "planet". There also exist
macroscopic models for the same purpose, for example Lotka-Volterra
model based on first-order non-linear differential-equations (Lotka 1956).

To finally show that the environment can also be of very abstract manner, a third and
last example is shown.

Fig. 3.5 Agent-based
approach modelling a number
of archaeological sites, as
given in Example 3.3
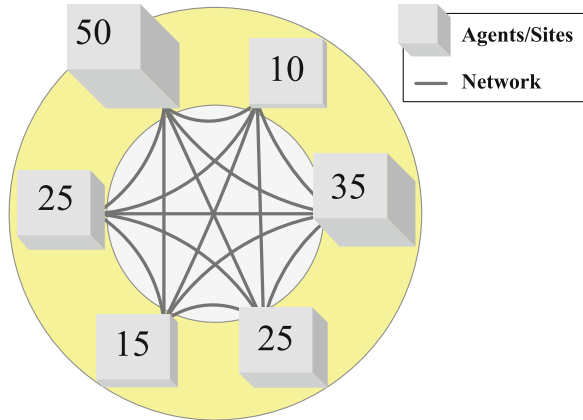
Example 3.3. Network of Archaeological Sites

Problem: A large team of archaeologists is distributed to six prospection sites.
The distribution of archaeologists should be solved optimally: Each site
has a certain capacity, which may not be surpassed. On the other hand,
each site demands a certain amount of people working there, or else the
prospection cannot be conducted effectively.

Abstraction: A first observation shows a positive correlation between the
number of archaeologists of a site and the corresponding work output (this
is, until the site gets too crowded for everyone to work efficiently, i.e. the
capacity is surpassed). The correlation coefficients and the capacities differ
from site to site. For the sake of this example, we furthermore assume
that these factors are independent from individual qualifications of each
archaeologist. Although the system itself is not dynamic (everyone keeps
working at his/her originally assigned site), it can be that the results of
a dynamic simulation in which archeologists are relocated might help to
determine the optimal distribution of workers.

Model: The "obvious" option to represent each archaeologist and site as agent
would work well in situation. However, this time, we choose to model
only the sites as agents (see Fig. 3.5) and the archaeologists as *the state
of these agents*, in the following manner: Each site stores the number of
archaeologists being occupied there, the demand and the capacity (both in
terms of people). Since sites are abstract, a detailed specification of the
environment is not necessary. All sites should be in contact at all times
during the simulation (e.g. by mobile phone, email or any other form of
communication), which leads to the introduction of links between each site

Example 3.3 (continued)

　(see network of edges in Fig. 3.5). Using this communication infrastructure, the sites can arrange an exchange in personnel so that the workload is handled dynamically, using a "supply and demand" approach.

Conclusion: If done properly, finally an equilibrium state should be achieved that distributes the research team. This model is inspired by economic models presented by Aoki (2002), which shows once more that agent-based modelling is more than dots moving (and colliding) in a spatial environment.

As demonstrated, the field of application for agent-based models is very broad. Especially if there is no scientific basis allowing white box models, agent-based approaches are getting more and more popular: They are visual, descriptive, easy to understand and very flexible. This is leading to a multitude of models not only differing in purpose and parameterisation, but also in model structure. Thus, it is sometimes necessary to subdivide the class of agent based approaches in support of a scientific metadiscourse about them. Usually, this is done by specifying a certain part of the model a little bit closer. The following sections therefore give a run-down of the most frequently used classifications in that sense.

3.2.2   Specification of Time: Discrete or Continuous

Agent-based simulations are usually dynamic simulations investigating and simulating the temporal behaviour of a system. This can be done either continuously or discretely (a fact that does not only apply to agent-based models alone, as every dynamic model can be assigned to one of these two classes). In order to be perfectly clear about these two termini "time-continuous" and "time-discrete", we now explain them for (ordered) sets. For those not interested in this rather mathematical excursus, it is recommended to skip to the definition time-continuous/discrete agent-based models further down in the text.

　It may sound trivial, but a *set* is no more than a simple assembly of any number and kind of items, called *elements* of the set. Usually the usage of curly brackets (set braces) indicates the assembly of all items in between. The order, how the elements are assembled is not important. Examples for sets:

- Room = {Couch, Table, TV, Chair}
- $A = \{1, 9, 4, 3, 2\}$
- Alphabet = $\{a, z, b, y, e, f, \ldots, m, p\}$
- NaturalNumbers = $\{1, 2, 3, \ldots\}$

Furthermore, a set is called *ordered* (or, to be precise, *totally ordered*) if every two different elements of the set can be compared and it can be determined which of them the bigger element is. For example, the previously defined set "Room" is (basically) not ordered, as it cannot be determined whether "Couch" is bigger or smaller than "TV". Obviously the sets "A" and "NaturalNumbers" are ordered. Although the set "Alphabet" is basically not ordered, instinctively one might say $a < b < c < \dots$. So it is possible to define an order in a very natural way, making the set "Alphabet" structurally equivalent to the set of all Numbers from 1 to 26. As every (totally) ordered set is structurally equivalent to a set containing numerical values, only those sets are analysed further.

The word "continuous" derives from the Latin adjective *continuous* and means *to do something steadily without interruptions*. We now try to derive the idea of continuous sets in a very intuitive way, by comparing the following two totally ordered sets:

- $X = \{1, 2, 3\}$
- $Y = \{1, 1.2, 1.4, 1.6, 1.8, 2, 2.2, 2.4, 2.6, 2.8, 2.9, 3\}$

Set Y can sloppily be seen to be a little bit "more continuous" version of set X, as the gap (respectively the interruption) between two consecutive elements is smaller. More specifically, for a perfectly continuous set it is necessary that the gap between every two consecutive numbers has to become infinitely small. As a consequence, a continuous set has to contain an infinite number of elements. More precisely, it has to contain *every* real number in a chosen interval, which is written using square brackets:

- $Z = [1, 3]$

The interval Z is a continuous set containing 1, 3 and all real numbers in between. It also contains the sets X and Y defined earlier. Finally, a proper definition of continuity can be presented:

Definition 3.2. An (ordered) set is called *continuous* if for any two different elements of the set there is at least a third element in the set that lies in between them.

The set/interval Z fulfils this condition. Regarding continuous sets, it is important to note that it is not possible to determine a "next greater" element. In case of the previously defined set A, the next larger element of 4 would be 9. In case of the continuous set Z, it is not possible to determine which element of the set is e.g. the next-larger number after 1.5. It can not be determined *by definition*: In the hypothetical case where one has found the next bigger element after 1.5, the definition of the continuous set states that there *has to be* an element lying in between 1.5 and this number! Motivated by this observation, a definition of a discrete set is given:

Definition 3.3. An (ordered) set is called *discrete* if every element has exactly one next-larger element.

This leads to three important observations:

- All finite sets (i.e. all sets with a finite number of elements) are discrete.
- If a discrete set (containing positive numbers) is not finite the elements are not bounded and become bigger and bigger.
- If otherwise all elements of a discrete set are bounded, the set is finite.

Finally, a correct definition of a time-continuous and time-discrete simulation can be given:

Definition 3.4. If the set of time-steps used for a dynamic model is a continuous, the model is called *time-continuous*. If otherwise the set of time-steps used for the model is a discrete set, the simulation is called *time-discrete*.

Some further important facts:

- Neglecting quantum mechanics, time can be seen to proceed continuously in reality. Thus, time-continuous models are usually preferred to discrete models.
- Unluckily, it is impossible to perform time-continuous simulations. Thus, a model that was designed as being time-continuous designed has to be "discretised", i.e. simulated time-discrete. As a consequence, we need to differentiate between time-continuous/discrete modelling and time-continuous/discrete simulation.
- Agent-based models are usually time-discrete models with equidistant time-steps (e.g. seconds). The decision-making process is not continuous, but happens as series of *events*: At a certain time step, the agent "decides" to take an action, resulting in a sudden state change. Some of these discrete state changes are harder to spot—for example when looking at movement: Even though it may seem that an agent moves continuously, it really performs a series of "hops" (once per time step) that we wrongly observe as being continuous. Truly time-continuous agent-based models are very rare. From the three examples presented earlier, solely the car model (see Example 3.3) would be suitable.
- The time-set does not necessarily have to be *either* continuous *or* discrete: There are also *hybrid models* where a mixture of partially continuous and discrete time intervals is possible.

3.2.3 Specification of the Environment: Discrete or Continuous

Definition 3.1 does not give any restriction regarding the shape and appearance of the environment, as long as it is inhabited by all agents. To illustrate this point further, the archaeological site model that was given as Example 3.3 did not include any topological aspects at all. Thus, we must first consider whether the environment contains topological concepts or not. If yes, a further classification into either spatially discrete or continuous can be conducted.

Fig. 3.6 Sketch of a spatial discrete compared to a spatial continuous agent-based model
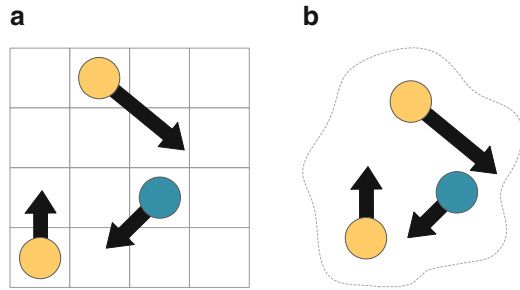

Figure 3.6 shows examples of a discrete and a continuous environment inhabited by agents. Spatially discrete models subdivide the environment, usually into equivalent cells as indicated in Fig. 3.6a. Such models are a lot easier to deal with, for instance because the calculation of movement happens in a finitely limited space.

Illustrating this point further using the predator-prey model previously presented in Example 3.2, it is clear that a reindeer chased by two lions should follow some strategy concerning an optimal direction of escape. Dealing with a spatial-discrete model, the optimal strategy can simply be determined to be one direction among a finite number of possible directions, depending on the position of the two predators. In a spatially continuous model, an infinite number of directions has to be taken into account. Likewise, distances and neighborhoods are also more complex to compute: Spatially discrete models can count the number of cells between agents and get all agents "around" a given agent (either the Von Neumann neighborhood consisting of the four cells north, south, east and west, or the Moore neighborhood additionally consisting of the north-east, north-west, south-east and south-west cell). Spatially continuous models must instead evaluate the distances (what is far, what is near?) and introduce radii (which radius is best?) around an agent.

Example 3.4. Salt Crystallisation

Problem: In order to investigate the crystallisation process of salt in a mine, a liquid model which is to be designed. Common liquid models use rather difficult mathematical formulae (e.g. Navier-Stokes-partial differential equations), which do not seem to work within this special case of modelling fine crystalline structures. Therefore, a microscopic approach is used. As we will see, this is a mixture between an agent-based approach and Cellular Automata (CAs), another type of microscopic approach[2] (also refer to Fig. 3.7).

---

[2]The transition between agent-based modelling and CAs is rather fluent: One could say with some caution that "agents" are potentially movable entities, while Cellular Automata specify a discrete environment of grid cells having rules that change their state (see e.g. Conway's "Game of Life" reported in Gardner 1971 and Berlekamp 2001).
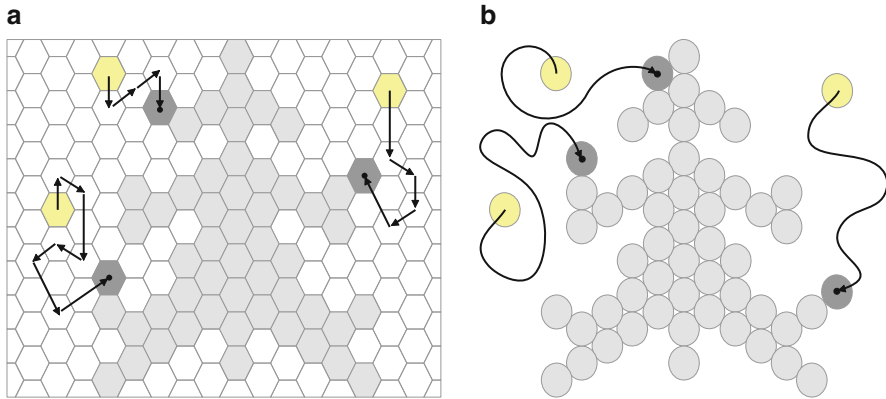
Fig. 3.7 Spatial discrete and continuous salt crystal model (see Example 3.4)

Example 3.4 (continued)

Abstraction: Observations done by microscope show on the one hand a rather random movement of dissolved ions in the liquid (Brownian motion) and on the other hand a very systematic crystallisation process happening at certain angles around the crystallisation seeds. The corresponding ions are thus abstracted as being agents behaving according to a Brownian motion in the liquid, finally sticking to one of the crystallisation seeds under certain conditions and at certain angles.

Model: The main task of this model is the specification of the environment and especially the crystallisation seeds and borders. This time, we model two complementary options: The systematic crystallisation process in which certain crystallisation angles are preferred can be implemented by using a spatially discrete model utilizing a hexagonal grid (see Fig. 3.7a). Whether one would like to call that a CA (a cell has its own behaviour for propagating crystalisation to the next cell) or an agent-based model (agents moving from cell to cell in a discretized environment) is a philosophical question leading to the same result. On the contrary, only a space-continuous model allows a correct implementation of a random-walk process (Brownian motion), which can be seen in Fig. 3.7b.

Conclusion: Typically, crystal growth is nowadays simulated by CAs. However, as this example shows, one may also choose a continuous approach, and both model types complement themselves nicely. Another point for consideration is that there is not always "a best approach", but possible a multitude of angles from which a problem can be approached. The context of this model, crystallisation, shows how simple microscopic rules can lead to a very complex, aggregated behaviour. For more detailed information about this model, see Packard (1986).

3.2.4   Stochastic and Deterministic Agent-Based Models

To give a final classification, the difference between stochastic and deterministic agent-based models is given. The word "stochastic", translated from original Greek, means "the art of probability" or "the art of randomisation". Thus, a model is generally called stochastic if the model is designed to contain random parameters or variables.[3] Such kinds of models have to be used if there is a lack of detailed information or uncertain data about specific modelling parameters. Another case is when a process cannot be modelled to such a degree of detail so as to be simulatable deterministically (also see Example 3.5).

When a model's inner mechanics are known (white box approach), it is easy to find out it contains randomisation, thus making it stochastic. But what if we have a black-box model? In this case, we can run a model twice with the same input/parameters and observe its output: If the outputs are the same, the model is deterministic, in all other cases stochastic. More formally:

Definition 3.5. A model is called deterministic, if and only if one initial configuration and parameterisation always leads to one specific simulation result. Otherwise the model is called stochastic.

Some further facts:

- Stochastic models must be simulated numerous times, until the obtained results are significant (e.g. within a confidence interval, variances, etc.). If this is not possible, simulation must be performed as often as possible (Monte Carlo Method). For agent-based models, this usually translates to high computational requirements.
- In all cases, obtained results demand proper interpretation by experts.
- Surprisingly, deterministic models sometimes lead to far more complex behaviour than stochastic ones. A perfect example is Conway's Game of Life (reported e.g. in Gardner 1971; Berlekamp 2001), in which very simple deterministic rules lead to a nearly chaotic behaviour.

> *Example 3.5.* Pool Billiard
>
> *Problem*: Pool billiard is a physical game: Success seems to not dependent on luck, but rather on controlling the angle, spin and power of a ball.
> *Abstraction*: The same angle, spin and power should lead to identical shots. Therefore, a deterministic model would seem to be an adequate solution in this case. On the other hand, pool-experienced reader might insist that

(continued)

---

[3]As a side-note, stochasticity is not a special property of agent-based models, but applies to all kinds of models.
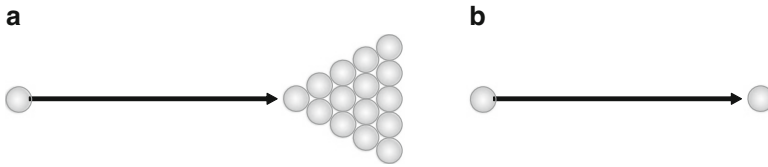
Fig. 3.8 It is not possible to say whether deterministic models or stochastic models are the best approximation for reality (refer to Example 3.5)

Example 3.5 (continued)
> the first shot (also called break-off shot, see Fig. 3.8a) contains lots of "random" effects and is usually not predictable. This is due to the fact that lots of factors influencing the 15 colliding balls (e.g. dirt-particles, surface irregularities and even quantum-mechanic effects) can hardly be measured before the shot.

Conclusion: It is not possible to say whether a deterministic (Fig. 3.8b) or stochastic model are the best approximation of reality in this specific case: Multi-body collisions behave stochastically, single-body collisions deterministically.

3.3   General Advantages and Disadvantages of Agent-Based Models

Out of all benefits that agent-based modelling brings, the ability to describe swarm behaviour is probably the most extraordinary one. It is mathematically a very surprising observation that the aggregated behaviour of a cohabitating group of agents is extremely difficult to be predicted when only investigating the cohabitation rules and parameters of the agents: This swarm behaviour that emerges unpredictably from a great number of interacting individuals (found e.g. with ants, migrating birds or humans at a soccer stadium) is now covered in more detail.

3.3.1   Emergent Behaviour

The term *swarm intelligence* or *swarm behaviour* was first introduced by Beni and Wang (1989) within the context of their work in robotics. It means that the predictable and simple behaviour of individuals leads to an unpredictable, complex behaviour for the whole group (also refer to Fig. 3.9). Today, we also use the terms *crowd intelligence* or *crowd behaviour* synonymously.

Although emergent phenomena arising from a large number individuals are well-known since long time, research in this field is still young (e.g. pedestrian and
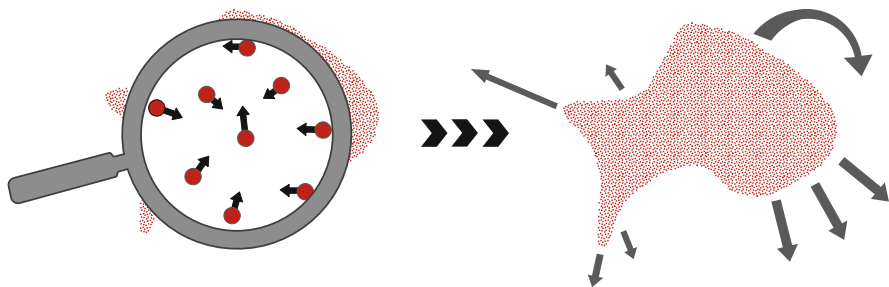
Fig. 3.9 Basic idea of emergent crowd behaviour respectively swarm-intelligence. The simple behaviour of individuals leads to (*shown left*) an unpredictable, complex behaviour of a whole crowd (*shown right*)

Fig. 3.10 "Intuitive" agent-based versus theoretical description of a system

evacuation simulation). On the one hand, it might be that the complexities of this topic have led to a slow take-up by the academic community. On the other hand, everyday computing has only recently reached performance levels at which large scale simulations are feasible. In that context, agent-based simulation is not simply used for "simulating behaviour", but as investigation tool for crowd intelligence.

3.3.2 Natural Description of a System

An agent does not know about the behaviour of the whole group but instead focuses on itself and its surroundings. Thus, the necessary level of abstraction needed is very low. This aspect has a lot of benefits: It enables easier interdisciplinary discourse between the creators of the models (technical disciplines and professionals such as archaeology) and their audience, with an emphasis on "who does what". This storytelling approach (as we would call it) may overcome the huge differences in mathematical skills: At least the structure of agent-based models can be explained very easily, the details are hidden in the code underneath (also see Fig. 3.10[4]).

---

[4]The diffusion equation at the right hand side of the picture only implies a model developed with difficult mathematical formulas. The diffusion equation is usually not used to model cohabitation of ants.

The intuitive representation of systems within microscopic approaches as well as the freedom granted by that modelling technique lead to extremely flexible models. Re-parameterisation as well as structural changes can be accommodated easily, by changing only a few lines in the implementation. From a computational design standpoint, this is called "Separation of Concerns" (SoC): Every part of the system behaves in a modular fashion, exposing information relevant to the outside (in our case e.g. the observable movement) but hides all further inner mechanics. Of course, this also implies that bugs hidden deep inside the code are hard to be found—which is one of the disadvantages mentioned in the next section.

3.3.3 Dangers and Problems

There are also some disadvantages in agent-based modelling, which can not all be attributed to the short history of the modelling method but may extend further to its underlying concept. On the surface, two important observations can be made when critically analysing agent-based models:

1. All agent-based simulations are computationally expensive. Doing serious simulation work, we may need millions or even billions of agents in order to receive useful results. So, these simulations might require a lot more powerful technologies than just personal computers—which is when high-performance computating (also see Chap. 6) comes into play. For example, parallel-computing (spread calculations onto several kernels), utilisation of the Graphics Processing Unit (GPU), cloud- and server-computing might all be used to make a simulation feasible, the decision over what method to take is, however, different in every model.
2. As mentioned before, the flexibility of the modelling technique also involves dangers: As every model is designed for a certain purpose, its field of application and especially validity is limited. Too many modifications within the model structure and wrong interpretation of the output can lead to wrong and invalid modelling result. One may also be tempted to interpret a model beyond its *border of validity*. Taking the Cave Lion example as a reference (also see Example 3.2), we may recall that this was a predator-prey model designed to predict the temporal development of the two cohabitating species. If the model is validated and calibrated the right way, it is correct to predict a certain number of animals at a specific time also in the real system. However, it is *not correct* to assume that the emerging hunting strategies of the model map to the hunting strategies in the real system, as they are (usually) too simplified!

Clearly, the "individual-based" simulation approach is what lies at the core of both scalability problems as well as dangers when every part comes together (overseen problems in the individual ruleset leading to a false aggregated result—a situation which is complex to debug). Even if modifications and extensions within the model are done properly, they usually introduce a lot of new modelling assumptions and parameters, which have to be calibrated and tested. Concluding, there are hardly

any modelling techniques as flexible as agent-based modelling, even though they are reacting quite sensitive to modifications and oversights in the specification of individual behaviour. Validation of these kind of models is a real challenge and requires a lot of time and effort. Thus, an own chapter is devoted to exactly this topic (see Chap. 4).

References

Aoki M (2002) Modeling Aggregate Behavior and Fluctuations in Economics: Stochastic Views of Interacting Agents. Cambridge University Press, New York

Aristoteles (1985) Metaphysics: Zeta, Eta, Theta, Iota. Books VII-X. HPC Classics Series. Hackett, Indianapolis. Edited by M. Furth

Beni G, Wang J (1989) Swarm Intelligence in Cellular Robotic Systems. In: Proceed. NATO Advanced Workshop on Robots and Biological Systems, Tuscany, Italy

Berlekamp ER (2001) Winning Ways for Your Mathematical Plays, 2nd edn. A.K. Peters, Natick, MA

Bocherens H, Drucker DG, Bonjean D, Bridault A, Conard NJ, Cupillard C, Germonpré M, Höneisen M, Münzel SC, Napierala H, Patou-Mathis M, Stephan E, Uerpmann HP, Ziegler R (2011) Isotopic evidence for dietary ecology of cave lion (Panthera spelaea) in North-Western Europe: Prey choice, competition and implications for extinction. Quaternary Int 245:249–261

Christiansen P (2008) Phylogeny of the great cats (Felidae: Pantherinae), and the influence of fossil taxa and missing characters. Cladistics 24(6):977–992. DOI 10.1111/j.1096-0031.2008.00226.x

Dewdney AK (1984) Computer recreations: Sharks and fish wage an ecological war on the toroidal planet wa-tor. Sci Am, 14–22

Gardner M (1971) Mathematical games: On cellular automata, self-reproduction, the Garden of Eden, and the game "life". Sci Am, 112–117

Helbing D, Molnár P (1995) Social force model for pedestrian dynamics. Phys Rev E 51(5): 4282–4286

Helbing D, Farkas I, Vicsek T (2000) Simulating dynamical features of escape panic. Nature 407:487–490

Kant I (1784) Beantwortung der Frage: Was ist Aufklärung? In: Gedike F, Biester JE (eds) Berlinische Monatsschrift. Haude und Spener, Berlin, pp 481–494

Kant I (1787) Werkausgabe, vol 3, 4th edn, Insel Verlag, Frankfurt, chap Vorrede zur zweiten Auflage, p 23. Edited by W. Weischedel

Lotka AJ (1956) Elements of Mathematical Biology. Dover, New York

Macal CM, North MJ (2006) Tutorial on Agent-Based Modeling and Simulation Part 2: How to Model with Agents. In: Proceedings of the 2006 Winter Simulation Conference, Monterey, California, pp 73–83, URL http://www.informs-sim.org/wsc06papers/008.pdf

Moore GE (1965) Cramming more components onto integrated circuits. Electronics 38(8):114–117

Nagel K, Schreckenberg M (1992) A cellular automaton model for freeway traffic. J de Physique I 2(12):2221–2229. DOI 10.1051/jp1:1992277, URL http://www.edpsciences.org/10.1051/jp1:1992277

Owen-Smith N (1989) Megafaunal extinctions: The conservation message from 11,000 years b.p. Conservat Biol 3(4):405–412

Packard NH (1986) Lattice Models for Soldification and Aggregation. In: Proceedings of the First International Syposium for Sience on Form. KTK Scientific Publishers, Tokyo, pp 95–101

Predtetschenski WM, Milinski AI (1971) Personenströme in Gebäuden - Berechnungsmethoden für die Projektierung. Verlagsgesellschaft Rudolf Müller, Köln-Braunsfeld

Rogsch C, Klingsch W (2011) To See Behind the Curtain - A Methodical Approach to Identify Cal-
    culation Methods of Closed-Source Evacuation Software Tools. In: Proceedings of Pedestrian
    and Evacuation Dynamics 2011, pp 567–576
Rogsch C, Weigel H, Klingsch W (2008) Hand-Calculation Methods for Evacuation Calculation -
    Last Chance for an Old-Fashioned Approach or a Real Alternative to Microscopic Simulation
    Tools? In: Klingsch W, Rogsch C, Schadschneider A, Schreckenberg M (eds) Pedestrian and
    Evacuation Dynamics 2008, pp 523–528
Shannon RE (1975) Systems Simulation: The Art and Science. Prentice-Hall, Englewood Cliffs,
    NJ
Stachowiak H (1973) Allgemeine Modelltheorie. Springer, Wien, New York
Stuart AJ, Lister AM (2011) Extinction chronology of the cave lion Panthera spelaea. Quaternary
    Sci Rev 30:2329–2340. DOI doi:10.1016/j.quascirev.2010.04.023
Wolfram S (2002) A New Kind of Science. Wolfram Media, Champaign, IL