# Using Elementary Cellular Automata to Model Different Research Strategies and the Generation of New Knowledge

**Georg Jäger**
*Institute of Systems Sciences, Innovation and Sustainability Research*
*University of Graz, Merangasse 18, 8010 Graz, Austria*

In this study, elementary cellular automata (CAs) are used to model the process of generating new knowledge. Each research goal is formulated as a target state of an elementary cellular automaton, while the scientific method used to reach this goal is represented as a rule. This system has many similarities to the actual process of knowledge generation, mainly caused by the possible complex behavior of CAs. The proposed model is then used to compare different strategies of scientific research like inter- and intradisciplinary cooperation in different scenarios. The obtained results are in agreement with reality and therefore substantiate the assumption that CAs are suitable to model the process of scientific research.

*Keywords*: knowledge generation; modeling complex systems; interdisciplinary cooperation; elementary cellular automata

## 1. Introduction

The generation and diffusion of new knowledge are of great significance for today's knowledge-based economy [1–4]. According to the triple helix concept [5–7], the main actors in these processes are governments, industry and of course universities, which are all involved in producing and sharing new knowledge. While diffusion of knowledge has been investigated very thoroughly [8–11], the actual process of generating knowledge is not yet completely understood.

Knowledge generation is closely related to learning, yet these two processes are not identical. Learning is the process of absorbing information that is already available somewhere. The process of learning can be observed in laboratory conditions [12–14], and there are various models that describe the learning process, like statistical learning theory [15], connectivism [16], transformative learning [17] or social learning theory [18]. The generation of new knowledge, that is, the process of doing research, is fundamentally different. This process is much more complex, and finding a suitable model is challenging. The main reason for this is that generating new knowledge is a creative

https://doi.org/10.25088/ComplexSystems.27.2.145

process that does not follow simple rules, and therefore the success of a research attempt is more difficult to predict than the success of an attempt at learning new information. In order to fill this research gap, we propose to use a model based on cellular automata (CAs) to model the process of generating new knowledge.

CAs are discrete models consisting of a grid of cells with a finite number of states. They can be used in many scientific fields to solve diverse problems [19] like traffic simulation [20–22], urban development [23], understanding complex social systems [24], medical models [25], energy-transport models [26], lattice gas models [27], cryptography [28], studying artificial life [29] or reservoir computing [30]. CAs can even be used as an alternative to differential equations [31].

The concept of CAs was proposed by Ulam and von Neumann in the 1940s [32]. One of the most popular CAs, Conway's Game of Life [33], introduced a broader audience to the concept of CAs by showcasing the immense complexity that can arise from simple rules. Although there are many different forms and classes of CAs, this study is mainly concerned with elementary cellular automata (ECAs), first systematically studied by Wolfram [34–37]. ECAs are one-dimensional CAs that only allow two different states for each cell, commonly labeled 0 and 1. Given an initial state of an ECA, the next state of each cell can be determined by the current state of the cell, the current state of its left neighbor and the current state of its right neighbor. Since each of these three cells can have two possible states, there are eight different combinations of those three states. Each of these combinations can result in the new state being either 0 or 1, leading to $2^8 = 256$ different rules for ECAs. These rules can be represented in binary notation [34] or as an integer (e.g., rule 00000011 corresponds to rule 3). Analyzing all these rules systematically reveals that some of them lead to a stable state, while others show periodical behavior. There are also rules that feature complex behavior. For rule 110, it was even proven that it is Turing complete [38]. An example of such complex behavior is given in Figure 1, which shows the time development of an arbitrary state using rule 110. ECAs are simple systems, yet they can show complex behavior under certain circumstances, so they are suitable candidates to model the equally complex process of generating new knowledge.

An abstract picture of a research process is given in the following. Each research process has a certain starting point (previous knowledge and experience) and a research goal (e.g., finding the answer to a specific research question). Researchers then apply a scientific method (computer simulations, statistical analysis, surveys, …) to reach this goal. Some attempts are successful right away, while for other

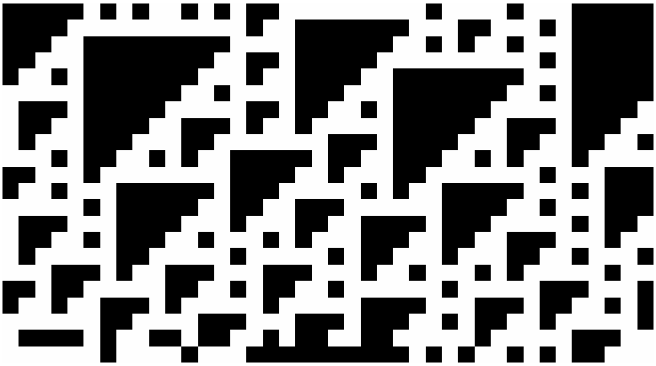

**Figure 1.** Time development of an arbitrary state using rule 110. Each line represents a discrete time step in this time development, with time running from top to bottom. In this visualization, the complex nature of the patterns that arise is clearly visible.

problems, more experience and therefore more attempts are necessary to find a solution. This system has astonishing similarities to ECAs. If we interpret both the current state of knowledge and the research goal as states of an ECA and the applied rule as the applied research method, both systems share the following properties:

- Some rules (methods) are better suited to reach the target state than others.
- Some rules (methods) are more flexible than others and can be used to reach many different target states.
- Not every rule (method) can be used to reach the target state, given a specific initial state.
- Similar target states (research goals) can be reached by similar rules (methods).
- Similar rules (methods) can show similar behavior.
- Whether or not a target state (research goal) can be reached, given a rule (method) and an initial state (previous knowledge), cannot be intuitively understood, but it is rather a complex process with unclear outcome.

These similarities can be utilized to construct a model of knowledge generation using ECAs, as detailed in the following section.

## 2. Methods

The purpose of this model is to use ECAs to simulate the process of research and knowledge generation and compare different research strategies in various situations. The research process is modeled in the following way: the research goal (which could be solving a specific problem, finding the answer to a specific research question or advancing the field in some other way) is represented as a state of an ECA, the so-called target state. The objective of the research process is now to reach this target state $S_T$, starting from an initial state $S_I$, within exactly $k$ steps. To reach this objective, different approaches and methods could be used. They are represented as the usual rules of ECAs [34]. So the overall research objective can be formulated as

$$S_T = \hat{R}^k(S_I), \tag{1}$$

where the operator $\hat{R}$ applies the rule $r$ ($0 \leq r < 256$) to a state. In order to satisfy this equation, the simulated researcher has to select a method (rule) and can then try to produce knowledge by choosing an arbitrary initial state $S_I$. If equation (1) is satisfied, the research attempt was successful. A visual representation of such a successful research attempt is given in Figure 2.

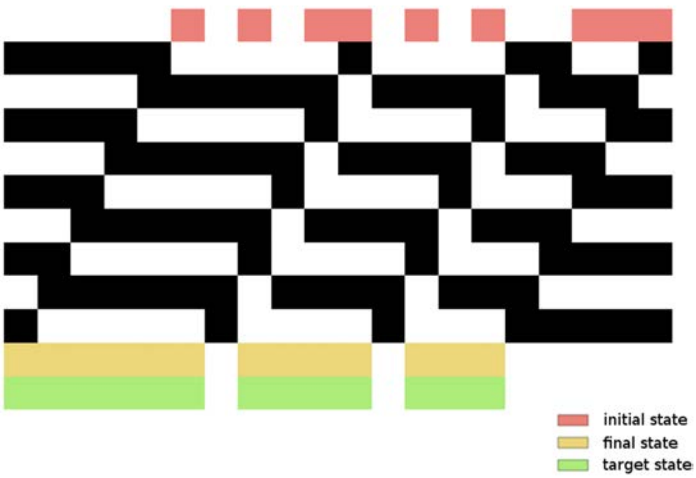


**Figure 2**. Successful attempt at solving a research problem. A certain rule (scientific method) is applied to an initial state (red) for a certain number of time steps (here, 10). The final state (yellow) matches the target state (green) perfectly, indicating complete success.

In general, however, the first attempt will not be successful, so an iterative optimization process is used, representing the researcher's efforts to arrive at a satisfying solution. Therefore, we need a way to measure the success of a research attempt. This is done by comparing the final state $S_F$, that is, the state that results from applying rule $r$ $k$ times to state $S_I$, to the target state $S_T$. Using a binary representation of both states, the proximity to the target $P$ can be calculated as

$$P = \frac{\sum_i \left(i - \left|S_T^{(i)} - S_F^{(i)}\right|\right)}{\sum_i i}, \tag{2}$$

leading to $P = 1$ for exact matches. Using this proximity, a simple optimization algorithm is used to model the research process. Starting from an initial state $S_I$, a modified state $S_M$ is generated by changing a random bit from $S_I$. Then the proximity to the target state of both states is calculated. If $S_M$ leads to a lower proximity, it is discarded; otherwise, $S_M$ replaces $S_I$. This process leads to a monotonically increasing proximity. We allow for 500 such iteration steps, after which the final success of this research process is evaluated as the proximity to the target state of the last (and therefore best) of the 500 attempts. One possible development of the proximity during the research process is depicted in Figure 3. Note that selecting a different number for the allowed steps has no influence on the qualitative results, as long as it big enough for the optimization process to make significant progress (≥100).

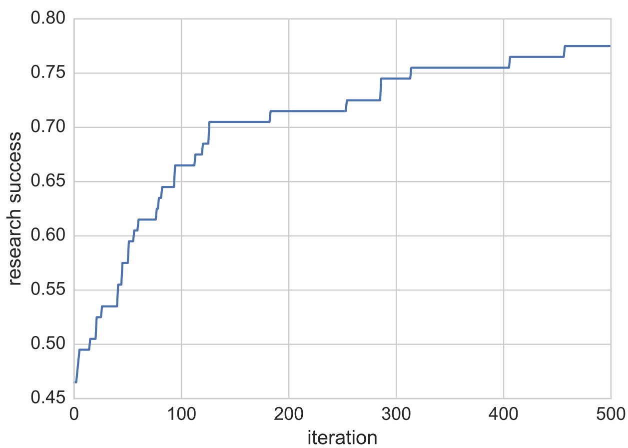


**Figure 3**. Proximity development during an iterative research process. The proximity is increasing monotonically, but $P = 1$ is not necessarily reached within 500 iteration steps.

This process represents the standard, uncooperative research strategy. One method (rule) is chosen and the method is applied to the problem until the best possible solution (given the allowed time) is found. However, there are other strategies that can also be investigated using this model. Intradisciplinary research benefits from the experience of other researchers in the same field, which can drastically improve the results. The methods used or generated by intradisciplinary research are similar to the ones researchers would use without cooperating, but they are not the same, and could be better suited to solve the problem at hand. Here, intradisciplinary research is modeled in the following way. After successive attempts to solve the research problem fail (i.e., if the proximity does not improve after several iteration steps), the applied method (rule) is slightly modified by changing one random bit. So rule 00001111 could change to 10001111, which is a different but closely related rule. If this new rule produces better results, it is used; otherwise, it is discarded.

Interdisciplinary research is depicted in a similar way. After a certain number of failed attempts, a new method is found by cooperating. Since here the cooperators come from a completely different field, the method (rule) is not related to the old rule, but chosen as a random rule of ECAs. Again, it is only used if it provides a benefit; otherwise, it is discarded. An overview of how the different research strategies are modeled is given in Figure 4.

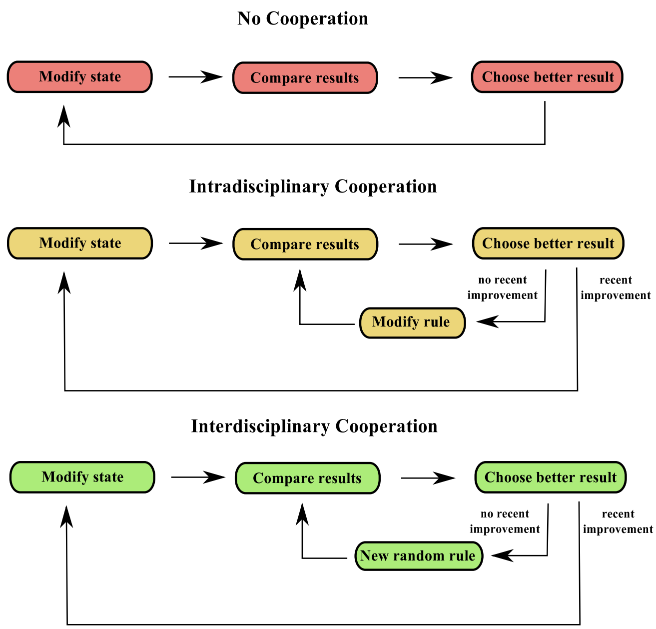


**Figure 4**. Modeled research process for different research strategies. While uncooperative research stays with the initial method, cooperative research can lead to modifications or changes of the method used.

To investigate the benefits of these research strategies, each of them is applied to random problems (random target states) of various difficulties. The difficulty of a problem is here defined as the size of the

ECA, so a target state with only 10 cells represents a simple problem, while a target state with 200 cells represents a difficult problem. Each strategy is used to solve the same 1000 research problems, and we evaluate the performance by looking at the distribution of final proximities and the percentage of fully successful attempts, defined by $P = 1$. In addition to this investigation, we simulate a scenario in which the solution of a similar problem is already known. Here, the starting point of each research strategy is the insight that a specific initial state leads to a known target state using a specific rule. The target state that needs to be reached is only a slight modification to the already-solved target state. All three research strategies try to solve the new problem with the solution of the old problem as the initial guess, modified in the usual way for cooperative research. The number of time steps after the target state should be reached (i.e., the parameter $k$ in equation (1)) can be chosen arbitrarily, since it has no significant influence on the final results, as long as $k \geq 5$. Simulations were performed with $k = 10$, and the results are presented in the following section.

## 3. Results

Results for solving simple research problems (the size of the target state is 10 cells here) are presented in Figure 5(a). Here it is evident that interdisciplinary research has the highest potential to solve the problem. Nearly 30% of all problems were solved with $P = 1$, so the target state was reached exactly. Intradisciplinary research performed slightly worse, with about 12% total success. For uncooperative research, fewer than 10% of research questions were solved with $P = 1$. The uncooperative strategy shows the biggest variance, which corresponds to the fact that some scientific methods are simply not suitable to solve certain problems.

The results of a medium-sized problem (20 cells) are presented in Figure 5(b). The findings are similar to the simple problem, yet the difference between the strategies is less pronounced. Fully successful attempts were reached in 6.1% of all cases for interdisciplinary research. Intradisciplinary research and uncooperative research reached 5.2% and 2.0%, respectively.

Figure 5(c) shows the results of attempts to solve difficult problems (a target state with 100 cells). Here, the results are qualitatively different from previous scenarios. While uncooperative research has the biggest variance regarding research success, it also has the highest percentage of fully successful attempts (16.5%). In this case, this strategy is therefore the most successful, when compared to intradisciplinary

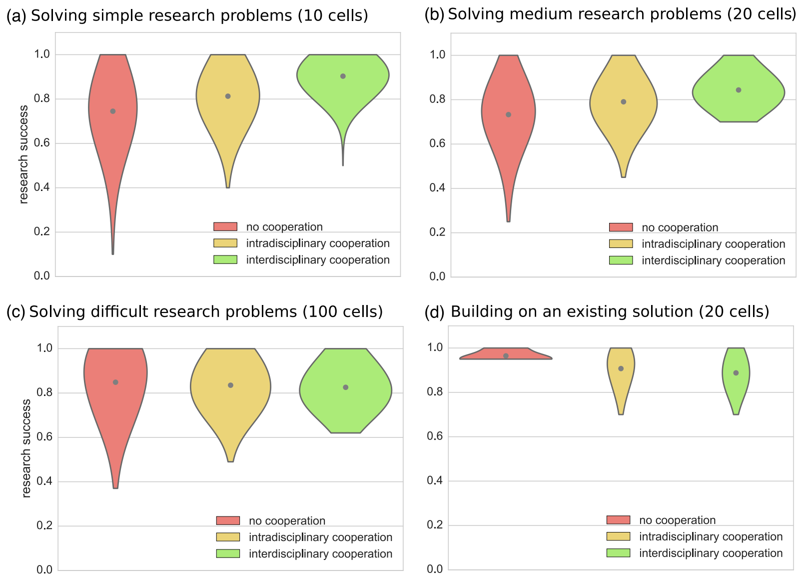


**Figure 5**. Success rate for solving research problems. (a) Simple research problems (10 cells). Here, interdisciplinary cooperation is more successful than the other strategies. The uncooperative strategy performs worst. (b) Medium research problems (20 cells). The situation is similar to the one for simple problems. The differences between strategies are, however, less pronounced. (c) Difficult research problems (100 cells). Here, the uncooperative strategy performs best on average and has the most 100% successful attempts. (d) Medium research problems (20 cells), when the solution to a similar problem is already known. In this scenario, the uncooperative strategy outperforms all other strategies, since changing the research method used due to cooperation is a disadvantage in this case.

cooperation (8.7%) and interdisciplinary cooperation (6.5%). A possible interpretation of this result is that difficult problems simply take more time and effort to solve. Changing the applied method or using a completely different one can therefore be a disadvantage in this situation. Staying with one method may sometimes produce bad results, but there is also a significant chance that it will lead to achieving the research goal, if one tries long enough. For very difficult problems, failures are somehow necessary in the scientific process. However, by gathering knowledge and experience in one specific method, even these problems can be solved.

The situation changes drastically, if one is interested in the solution of a problem, when the solution to a similar problem is already known, as depicted in Figure 5(d). These problems were of medium difficulty (20 cells). Since an approximate solution was already known, all strategies performed better than for an ordinary problem of medium difficulty. The percentage of perfect solutions was 11.3% for interdisciplinary research and 14.9% for intradisciplinary research. For uncooperative research, 27.9% of all attempts were fully successful, and the average research success was above 95%. This result is in agreement with the real scientific process. If a similar problem has already been solved, using a similar method is very successful. Changing a method that is proven to work by combining the original method with methods of other researchers can be a big disadvantage, and it may be more beneficial to rely on methods that are well established in the scientific field in question.

## 4. Discussion

This study shows that it is indeed possible to use elementary cellular automata (ECAs) to model the process of research and knowledge generation. Obtained results are in agreement with the observed reality, and the model can serve as a starting point for further investigations of the advantages and disadvantages of different research strategies. While most problems benefit from cooperation, there are two special cases in which cooperation might not be beneficial. If the research problem is too difficult, not cooperating was the best strategy. Focusing on one single approach and gathering experience there has a higher success chance than using new, cooperative approaches. Also, if the solution to a similar problem is known, it is better to rely on the established method than to try new, cooperative methods.

The presented model relies on ECAs to model the research process, which makes it quite abstract. Each research goal is formulated as a random state of an ECA of an arbitrary size that determines the difficulty of the posed problem. The process of reaching this research goal is then a simple optimization. Even though this depiction of the research process is simplistic, it can serve as an elementary model. In that sense, its simple nature is a big advantage. The model is also suitable for investigating different research strategies, by slightly modifying the optimization process. This is a significant simplification, since here the research strategies only differ in terms of method used, even though intra- and interdisciplinary cooperation can have more effects on the research process, but are not included here. Nevertheless, the presented model produces plausible results and can serve as a starting point for a more advanced model.

Various expansions are conceivable to make the model more realistic. While the simple optimization process used here is a viable way to depict the research process, more elaborate optimization techniques that also allow for decreasing proximity, like simulated annealing [39], could also be used and might be closer to reality. However, too-elaborate techniques would lead to a perfect solution for most problems and can therefore not be used to model the process of generating knowledge. So the choice of the optimization algorithm is arbitrary, but the chosen technique does lead to realistic results. Another interesting expansion would be to restrict the allowed rules to rules with class 3 or 4 behavior [40], to better account for the complexity of the process of knowledge generation.

The problem of finding a suitable way of modeling the generation of new knowledge is deeply connected to our lack of understanding of how research exactly works on a fundamental level and what processes and effects are responsible for the success or failure of a research attempt. One has to accept that this may never be completely understood, since the system is simply too complex. This may also be the reason why ECAs are so well suited to model the research process. They can feature astonishing complexity as well and seem to share many properties of the investigated complex system. This makes them promising candidates for modeling complex systems in general and underpins the need for further research in the field of CAs.

## References

[1] D. Foray and B. Lundvall, "The Knowledge-Based Economy: From the Economics of Knowledge to the Learning Economy," *The Economic Impact of Knowledge* (D. Neef, G. A. Siesfeld and J. Cefola, eds.), Boston: Butterworth-Heinemann, 1998 pp. 115–121.

[2] B. Godin, "The Knowledge-Based Economy: Conceptual Framework or Buzzword?," *The Journal of Technology Transfer*, **31**(1), 2006 pp. 17–30. doi:10.1007/s10961-005-5010-x.

[3] L. Leydesdorff, *The Knowledge-Based Economy: Modeled, Measured, Simulated*, Boca Raton: Universal Publishers, 2006.

[4] G. Dosi, "The Contribution of Economic Theory to the Understanding of a Knowledge-Based Economy," *The Economic Impact of Knowledge*, (D. Neef, G. A. Siesfeld and J. Cefola, eds.), Boston: Butterworth-Heinemann, 1998 pp. 123–129. doi:10.1016/B978-0-7506-7009-8.50012-4.

[5] L. Leydesdorff and H. Etzkowitz, "The Triple Helix as a Model for Innovation Studies," *Science and Public Policy*, **25**(3), 1998 pp. 195–203. doi:10.1093/spp/25.3.195.

[6] L. Leydesdorff, "The Triple Helix Model and the Study of Knowledge-Based Innovation Systems," *International Journal of Contemporary Sociology*, **42**(1), 2005 pp. 12–27.

[7] L. Leydesdorff, "Triple Helix of University-Industry-Government Relations," in *Encyclopedia of Creativity, Invention, Innovation, and Entrepreneurship* (E. G. Carayannis, ed.), New York: Springer-Verlag, 2013 pp. 1844–1851.

[8] R. Cowan and N. Jonard, "Network Structure and the Diffusion of Knowledge," *Journal of Economic Dynamics and Control*, **28**(8), 2004 pp. 1557–1575. doi:10.1016/j.jedc.2003.04.002.

[9] O. Sorenson and L. Fleming, "Science and the Diffusion of Knowledge," *Research Policy*, **33**(10), 2004 pp. 1615–1634. doi:10.1016/j.respol.2004.09.008.

[10] C. Autant-Bernard, J. Mairesse and N. Massard, "Spatial Knowledge Diffusion through Collaborative Networks," *Papers in Regional Science*, **86**(3), 2007 pp. 341–350. doi:10.1111/j.1435-5957.2007.00134.x.

[11] R. Cowan and D. Foray, "The Economics of Codification and the Diffusion of Knowledge," *Industrial and Corporate Change*, **6**(3), 1997 pp. 595–622. doi:10.1093/icc/6.3.595.

[12] H. Hydén and E. Egyházi, "Nuclear RNA Changes of Nerve Cells during a Learning Experiment in Rats," *Proceedings of the National Academy of Sciences*, **48**(8), 1962 pp. 1366–1373. doi:10.1073/pnas.48.8.1366.

[13] A. S. Reber, "Implicit Learning of Artificial Grammars," *Journal of Verbal Learning and Verbal Behavior*, **6**(6), 1967 pp. 855–863. doi:10.1016/S0022-5371(67)80149-X.

[14] R. G. M. Morris, E. Anderson, G. S. Lynch and M. Baudry, "Selective Impairment of Learning and Blockade of Long-Term Potentiation by an N-methyl-D-aspartate Receptor Antagonist, AP5," *Nature*, **319**(6056), 1986 pp. 774–776. doi:10.1038/319774a0.

[15] V. N. Vapnik, *Statistical Learning Theory*, New York: Wiley, 1998.

[16] G. Siemens. "Connectivism: A Learning Theory for the Digital Age." (Apr 25, 2018) elearnspace.org/Articles/connectivism.htm.

[17] P. Cranton, *Understanding and Promoting Transformative Learning: A Guide for Educators of Adults*, San Francisco: Jossey-Bass Publishers, 1994.

[18] I. M. Rosenstock, V. J. Strecher and M. H. Becker, "Social Learning Theory and the Health Belief Model," *Health Education & Behavior*, **15**(2), 1988 pp. 175–183. doi:10.1177/109019818801500203.

[19] S. Wolfram, *Cellular Automata and Complexity: Collected Papers*, Reading, MA: Addison-Wesley, 1994.

[20] K. Nagel and M. Schreckenberg, "A Cellular Automaton Model for Freeway Traffic," *Journal de physique I*, **2**(12), 1992 pp. 2221–2229. doi:10.1051/jp1:1992277.

[21] D. A. Rosenblueth and C. Gershenson, "A Model of City Traffic Based on Elementary Cellular Automata," *Complex Systems*, **19**(4), 2011 pp. 305–322. complex-systems.com/pdf/19-4-1.pdf.

[22] C. Stoica-Klüver and J. Klüver, "Simulation of Traffic Regulation and Cognitive Developmental Processes: Coupling Cellular Automata with Artificial Neural Nets," *Complex Systems*, **17**(1), 2007 pp. 47–64. complex-systems.com/pdf/17-1-4.pdf.

[23] M. Batty, *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*, Cambridge, MA: MIT Press, 2007.

[24] R. Hegselmann and A. Flache, "Understanding Complex Social Dynamics: A Plea for Cellular Automata Based Modelling," *Journal of Artificial Societies and Social Simulation*, **1**(3), 1998 pp. 1–30. jasss.soc.surrey.ac.uk/1/3/1.html.

[25] T. Alarcón, H. M. Byrne and P. K. Maini, "A Cellular Automaton Model for Tumour Growth in Inhomogeneous Environment," *Journal of Theoretical Biology*, **225**(2), 2003 pp. 257–274. doi:10.1016/s0022-5193(03)00244-3.

[26] V. J. Peridier, "Estimating Transient Surface Heating Using a Cellular Automaton Energy-Transport Model," *Complex Systems*, **16**(2), 2006 pp. 139–153. complex-systems.com/pdf/16-2-3.pdf.

[27] D. A. Wolf-Gladrow, *Lattice-Gas Cellular Automata and Lattice Boltzmann Models: An Introduction*, Berlin, Heidelberg: Springer-Verlag, 2000.

[28] S. Wolfram, "Cryptography with Cellular Automata," in *Advances in Cryptology (Crypto '85 Proceedings)*, Berlin, Heidelberg: Springer, 1985 pp. 429–432. doi:10.1007/3-540-39799-X_ 32.

[29] C. G. Langton, "Studying Artificial Life with Cellular Automata," *Physica D: Nonlinear Phenomena*, **22**(1–3), 1986 pp. 120–149. doi:10.1016/0167-2789(86)90237-X.

[30] S. Nichele and M. S. Gundersen, "Reservoir Computing Using Nonuniform Binary Cellular Automata." arxiv.org/abs/arXiv:1702.03812.

[31] T. Toffoli, "Cellular Automata as an Alternative to (Rather than an Approximation of) Differential Equations in Modeling Physics," *Physica D: Nonlinear Phenomena*, **10**(1–2), 1984 pp. 117–127. doi:10.1016/0167-2789(84)90254-9.

[32] J. von Neumann, "The General and Logical Theory of Automata," *Cerebral Mechanisms in Behavior: The Hixon Symposium* (L. A. Jeffress, ed.), New York: Wiley, 1951 pp. 1–41.

[33] M. Gardner, "The Fantastic Combinations of John Conway's New Solitaire Game 'Life'," *Scientific American*, **223**(4), 1970 pp. 120–123.

[34] S. Wolfram, "Statistical Mechanics of Cellular Automata," *Reviews of Modern Physics*, **55**(3), 1983 pp. 601–644. doi:10.1103/RevModPhys.55.601.

[35] S. Wolfram, "Cellular Automata as Models of Complexity," *Nature*, **311**(5985), 1984 pp. 419–424. doi:10.1038/311419a0.

[36] S. Wolfram, "Computation Theory of Cellular Automata," *Communications in Mathematical Physics*, **96**(1), 1984 pp. 15–57. doi:10.1007/BF01217347.

[37] S. Wolfram, "Universality and Complexity in Cellular Automata," *Physica D: Nonlinear Phenomena*, **10**(1–2), 1984 pp. 1–35. doi:10.1016/0167-2789(84)90245-8.

[38] M. Cook, "Universality in Elementary Cellular Automata," *Complex Systems*, **15**(1), 2004 pp. 1–40. complex-systems.com/pdf/15-1-1.pdf.

[39] S. Kirkpatrick, C. D. Gelatt, Jr. and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, **220**(4598), 1983 pp. 671–680. doi:10.1126/science.220.4598.671.

[40] S. Wolfram, *A New Kind of Science*, Champaign, IL: Wolfram Media, Inc., 2002.