

MAGNETS, SPINS, AND NEURONS:
THE DISSEMINATION OF MODEL TEMPLATES ACROSS DISCIPLINES

Tarja Knuuttila and Andrea Loettgers

One of the most conspicuous features of contemporary modeling practices is the dissemination of mathematical and computational methods across disciplinary boundaries. We study this process through two applications of the Ising model: the Sherrington–Kirkpatrick model of spin glasses and the Hopfield model of associative memory. The Hopfield model successfully transferred some basic ideas and mathematical methods originally developed within the study of magnetic systems to the field of neuroscience. As analytical resource we use Paul Humphreys’ discussion of computational and theoretical templates. We argue that *model templates* are crucial for the intra- and interdisciplinary theoretical transfer. A model template is an abstract conceptual idea associated with particular mathematical forms and computational methods.

1. Introduction

Theoretical work in many, if not most, scientific disciplines is driven today by mathematical and computational modeling. Philosophers of science have taken notice of this development. However, despite the quickly accumulated literature on modeling, philosophical discussion has barely touched upon one of the most conspicuous features of contemporary modeling practices: the transfer of mathematical and computational

methods across disciplinary boundaries. The reason for this neglect of the interdisciplinary features of modeling can be related, at least in part, to the inclination of philosophers to concentrate on the representational relationship of a single model and its supposed real-world target system. Yet another important reason is the lack of appropriate conceptual tools to study the cross-disciplinary exchange that takes place in contemporary modeling. Namely, if one wants to cover various fields of study and even various disciplines, the obvious question then becomes: What draws all these seemingly related yet different models together? From the perspective of the actual modeling practices, the more traditional answers given by theoretical unification (Kitcher 1981) and interfield theories (Darden and Maull 1977) do not seem to pay enough attention to the important role that mathematical formulas and computational methods play in theoretical transfer. A proposal that targets precisely these aspects of modeling has been put forth by Paul Humphreys (2002; 2004).

In discussing present computational science, Humphreys asks what would provide us a proper unit of analysis to unravel its characteristic traits. He finds such a unit in what he coins a computational template. Computational templates are genuinely cross-disciplinary mathematical formulas and methods, such as sets of equations, algorithms, or computational methods, which can be applied to different problems in various disciplines. To date, Humphreys' important insight on computational templates has not yet attracted too much philosophical discussion (see however Knuuttila and Loettgers 2011). What seem to be needed are more case studies on model transfer between different disciplines detailing what it is that actually travels between the different fields and how. In the following we will offer one such case study: the transfer of spin glass models in

statistical physics to modeling neural networks in neurosciences. The Ising model that is one of the most famous highly idealized and simplified models in physics provides the basic template for both models (Ising 1925).

2. Computational Templates and Interdisciplinary Transfer

Paul Humphreys' notions of theoretical and computational templates are developed in the context of his analysis of the contemporary computational science (2002; 2004).

According to Humphreys, to fully appreciate the use of computational methods in science one needs to switch attention from the problems of representation to the problems of computation. Instead of models, theories, or paradigms, he wants to start from something that is "simple and well-known" (2004, 60), so simple and well-known in fact that it has escaped the explicit attention of philosophers of science. He notes the "the enormous importance of the relatively small number of computational templates in the quantitatively oriented sciences" adding that "science would be vastly more difficult if each distinct phenomenon had a different mathematical representation" (Ibid., 68). These computational templates that Humphreys adopts as his basic units of analysis are cross-disciplinary equations, or other mathematical and computational methods, that are typically but not necessarily derived from theoretical templates. They may have been introduced as theoretical models of a certain system like the Ising model and the Lotka–Volterra model, being only subsequently applied to different domains. On the other hand, computational templates may have their origin in formal disciplines like the Poisson distribution in probability theory. In the former case, the theoretical model becomes a

genuine computational template first when it is separated from the original theoretical context and used to model other, often different types of phenomena.

What is it then that distinguishes theoretical templates like the Ising model or the Lotka–Volterra model from computational templates? According to Humphreys, the motivation for the use of computational templates is due to their tractability and generality. Both the Lotka–Volterra model and the Ising model are very general, so general in fact that they have been applied to all kinds of systems. Thus, it seems that it is not generality that distinguishes theoretical templates from the computational templates. Indeed, it is tractability that is crucial for a template to become a computational template: “The most important kind of computational template is found at the first level at which a tractable template occurs as a result of substitutions into the theoretical template” (Humphreys 2004, 61). Consequently, computational templates are the result of intricate construction processes: “more than one step is involved in moving from a theoretical template to an application because specific parameter values as well as more general features, such as the forces must be determined” (Ibid.). What does Humphreys mean by this?

Humphreys uses Newton’s Second Law as an example. In the case of the mathematical form of Newton’s Second Law ($F = ma$), specifying the particular force function (F) using Hooke’s law for a spring ($F = -kx$) produces a computational template for a harmonic oscillator ($-kx = ma$). In order to apply this template to some specific system one would need to specify the value of k in terms of observations or experiments. Because of this, although he claims that syntax matters, Humphreys also underlines that computational templates come with intended interpretation and associated justification.

And yet (much) more work is needed to turn a computational template into a proper model, as it involves subject-specific construction assumptions (e.g., ontology, idealizations, and approximations) and other adjustments and corrections.

We appreciate Humphreys' insight on computational templates, yet it seems to us that tractability may not be as crucial for interdisciplinary model transfer as Humphreys' analysis may lead one to think. To be sure, independence from specific subject matter makes a template genuinely transferable. But what is independent is not just the piece of mathematics but also the general conceptual idea(s) that enable the application of the associated equations and computational methods. In the case of the Ising model, as we will show in the next sections, conceptual ideas such as the critical point phenomena, phase transition and especially the central notion of cooperative behavior guided model transfer. Tractability was certainly important, but on the other hand, various mathematical methods for handling the model and getting it to yield interesting results were invented as a part of the further model development. Consequently, what made the Ising model into what we call a model template is the conceptual framework which renders certain kinds of patterns as instances of cooperative phenomena coupled with mathematical forms and tools suitable for articulating and developing this framework.

3. Modeling Cooperative Phenomena

The Ising model provides a striking example of model transfer. In a series of articles on the history of the Ising model, the historian of science Martin Niss showed how the Ising model, which was constructed in the 1920's as a model for the study of ferromagnetism,

became during the 1960's one of the most studied models of cooperative phenomena (e.g., Niss 2005). Stephen Brush, an author of an early version of the history of the Ising model—right at the time when the change in application and perception of the model took place—wrote about this development:

Whereas previously it appeared that the greatly over-simplified representation of intermolecular forces on which this model is based would make it inapplicable to any real systems, it is now being claimed that the essential features of cooperative phenomena (especially critical point) depend not on the details of intermolecular forces but on the mechanism of propagation of long-range order, and the Ising model is the only one which offers much hope of an accurate study of this mechanism (1967, 883).

Indeed, there has been a marked shift in what was considered the important property of the Ising model. Detailed understandings of the intermolecular forces gave way to more general considerations. The new focus was placed on the mechanism of the propagation of long-range order, that is, understanding of phenomena such as the transition from a paramagnetic to a ferromagnetic phase. With this change in what counts as the important property of the model, the status of the model also changed. It gained a much higher degree of generality and with it the possibility of an applying it in different scientific contexts and disciplines: “Whether or not it does eventually turn out that gas-liquid critical phenomena, magnetic Curie points, order-disorder transitions in alloys, and phase separation in liquid mixtures can all be described, to a good first approximation, by the same model, the problem of a generalized description of cooperative phenomena now deserves serious attention” (Brush 1967, 883).

The further development of the Ising model certainly shows that the main reason why the Ising model was applied to a wide variety of systems was not related to its mathematical tractability alone but to the fact that it provided a structure and mechanism to study cooperative phenomena.

3.1. The Ising model

Macroscopic magnetic substances usually consist of atoms or molecules, both of which are equipped with a small magnetic moment. In the most common example, magnetic iron, the forces between the neighboring magnetic moments of the iron atoms tend to align those moments in the same direction resulting in a macroscopic net magnetic moment, the ordinary magnetism of an iron magnet. Thermal motion counterfeits this tendency and above a certain critical temperature T_C , for pure iron above some $770^\circ C$, no net magnetic moment for the macroscopic piece of iron is found. The critical temperature separates a paramagnetic phase with no net magnetic moment from a ferromagnetic phase.

If the magnetic system approaches the critical temperature T_C in the process of cooling down or heating up, it undergoes what is called a phase transition. Experimentally, such phase transitions display certain characteristic features. Thus, one observes a very pronounced increase in the specific heat and the magnetic susceptibility just at T_C . This behavior is based on what Brush called the mechanism of propagation of a long-range order in the magnetic system that is discussed in more detail below.

The Ising model consists of N magnetic moments, so-called spins S_i , which can take only two values, $S_i = +1$ or -1 , corresponding to their two possible discrete orientations up and down. In the one-dimensional Ising model the spins S_i are placed on a string and each spin interacts only with its nearest neighbor. The interaction is given by the interaction energy J_{ij} describing the interaction strength between *nearest* neighbor spins S_i and S_j . In the original Ising model J_{ij} is constant and in the case of a ferromagnet where all the spins are aligned, the interaction energy J is a constant and positive ($J > 0$). Ising's original analysis of the one-dimensional case did not show any phase transitions. From this negative result, he drew the conclusion that there would not be any phase-transitions either in higher dimensions. However, in 1942 Lars Onsager showed that phase-transitions occur in the two-dimensional Ising model (Onsager 1944).

In the two-dimensional case the spins are located on the sides of a lattice and in the three-dimensional case on the sides of a cube. Again, the interactions only take place between nearest neighbors and the interaction energy is positive ($J > 0$) when the two neighboring spins point in the same direction or negative ($J < 0$) when the spins point in opposite directions. To sum up, the interaction energy depends on the configurations of neighboring spins and furthermore tends to align neighboring spins.

With each magnetic moment S_i comes an internal magnetic field h_i created by magnetic moment S_j :

$$h_i = \sum_{j, j \neq i}^N J_{ij} S_j$$

with $J_{ij} = J_{ji}$.

The energy of the system depends on the state and the interactions and is given by the following energy function:

$$E\{S\} = -\frac{1}{2} \sum_{i,j \neq i}^N J_{ij} S_i S_j .$$

The overall energy of the system decreases if S_i and S_j point in the same direction. In this case the interaction energy J_{ij} is making a positive contribution, which, together with the $-$ sign in front of the sum, leads to the decrease of the overall energy E . If S_i and S_j point in different directions, the overall energy of the system increases because the interaction energy J_{ij} is making a negative contribution.

Methods and concepts have been developed to calculate equilibrium states as well as phase transitions and critical exponents of the Ising model. The so-called partition function is particularly important in this respect:

$$Z = \sum_{\{S\}} e^{-E/kT}$$

where k is the Boltzmann constant. The sum runs over all the 2^N possible states of the system and to each state $\{S_i, i = 1, \dots, N\}$ a certain probability is assigned. This probability is given by the so-called Boltzmann factor $e^{-E/kT}$. If one knows the energy E for every state of the system, then the properties of the system, in equilibrium at a temperature T , can in principle be studied as if one had an ensemble of identical systems (canonical ensemble). The probability for finding one of them in any of the possible states is described by the Boltzmann factor $e^{-E/kT}$.

The partition functions has been one of the main reason why it took such a long time to solve the two-dimensional case. Mathematical tools needed to be developed to

handle the sum over all the possible states of the system. Being able to calculate the partition function allows for example to determine the magnetization M of the system at a given temperature T . M would be calculated by:

$$\langle M \rangle = \frac{\sum_{\{S\}} M\{S\} \cdot e^{-E/kT}}{\sum_{\{S\}} e^{-E/kT}} .$$

As already pointed out, the application of the Ising model has not been limited to magnetic systems. Its property of providing a generalized model for cooperative phenomena allowed and triggered its application to different systems in and beyond physics. In the next section we discuss the utilization of the Ising model in studying critical behavior in disordered magnetic systems, so-called spin-glasses.

3.2. Variations of the Ising model

There are several ways in which the ‘simple’ Ising model can be changed. For example, the structure or dimension of the lattice can be changed, an external magnetic field can be introduced, or the interactions between the spins can be varied. In the following we focus on the variation of the interactions between the magnetic moments because it became important for the transfer of the model into the field of neural networks. As discussed above, in the Ising model, the interaction strengths had been chosen to be constant and restricted to nearest neighbors. But there are more complicated cases. An example is provided by the so-called spin glasses. In the 1970’s spin glasses caught the attention of physicists in the field of condensed as well as statistical mechanics. Spin glasses differ from ordinary glasses by containing a small number of magnetic moments interacting

with each other. These interactions lead to interesting cooperative behavior due to the fact that both ferromagnetic and antiferromagnetic couplings are present in the system. The simultaneous presence of ferromagnetic and antiferromagnetic couplings, in general, does not allow for the establishment of a conventional long-range order (of ferromagnetic or antiferromagnetic type).

Using the Ising model as a starting point, Samuel F. Edwards and Philip W. Anderson designed a model for spin glasses, the so-called Edwards–Anderson model (Edwards and Anderson 1975), which then got extended by David Sherrington and Scott Kirkpatrick leading to the Sherrington–Kirkpatrick model (SK model) (Sherrington and Kirkpatrick 1978). The main difference between the two models is that, like the Ising model, the Edwards–Anderson model is limited to nearest neighbor interaction whereas the Sherrington–Kirkpatrick model takes infinitely long-range interactions into account. In the latter model each spin is connected to all the other spins in the system. Despite the difference in how the interactions are modeled, both models share the basic structure taken from the Ising model. We will concentrate on the Sherrington–Kirkpatrick model because of the important role it played in modeling the properties of neural networks such as pattern recognition.

3.3. The Sherrington–Kirkpatrick model for spin glasses

Like in the Ising model the magnetic moments are represented by binary variables S_i in the SK model. Again the magnetic moments S_i take either value +1 or -1. The coupling

between two spins, S_i and S_j , is, as before in the case of the Ising model, represented by the coefficient J_{ij} and the overall energy of the system is also of the form:

$$E = - \sum_{i,j} J_{ij} S_i S_j .$$

At first sight the interaction energy is of the same form as in the Ising model. The main difference becomes visible once one takes a closer look at the couplings between the magnetic moments. In the SK model the couplings are modeled as a function of the distance between the magnetic moments $J_{ij} = J(R_i - R_j)$, with R_i and R_j as the positions of the magnetic moments on for example a lattice. Positive values of J_{ij} correspond to ferromagnetic and negative values to antiferromagnetic couplings.

In the SK model the spins cannot at the same time satisfy both ferromagnetic and antiferromagnetic couplings. Ferromagnetic couplings tend to align the spins; antiferromagnetic couplings make them point in opposite directions. The consequences of these competing interactions between the ferromagnetic and antiferromagnetic couplings become apparent at low temperatures when the system undergoes something like a phase transition. The system exhibits a ‘freezing transition’ to a state with a new kind of ‘order’ in which the magnetic moments are aligned in random directions (Binder and Young 1986). The topology of the energy landscape of spin glasses after undergoing this ‘freezing transition’ is rather complex, consisting of a large number of valleys, representing metastable or stable spin configurations.

The spin glasses stick out because they do not exhibit the same kind of phase transition as has been observed in so many other systems, such as for example in ferromagnetic systems, gases and liquids, which had been so successfully modeled by the

Ising model. The phase transition observed in the case of the spin glasses does not lead into an ordered phase (Cannella and Mydosh 1972). Indeed, one of the most debated questions concerning the spin glasses is whether the transition is a phase transition at all, and if so, whether it could be a new kind of phase transition. Such questions sparked an active field of research in condensed matter physics and statistical mechanics. The place to look for an explanation seemed to be the interactions between the components of the system, which had already turned out to be critical in the case of the Ising model. The SK model differed from the Ising model only with respect to the form of interactions and their infinite range.

As in the case of the Ising model, calculating the partition function posed a challenge. Here the sampling over different configurations of magnetic moments in order to calculate equilibrium states and possible phase transitions needed to be extended in order to account for the disorder in the system. A novel mathematical method was developed for this purpose: the so-called replica method. In order to get an idea of this mathematically rather difficult method, one only needs to know that in addition to calculating probability for the occurrence of the 2^N possible states of the system, one also has to take into consideration the different possible realizations of disorder. But as it turned out, taking into account the disorder leads to a serious problem: The form of phase transitions varies depending on the distribution P_{ij} of the interaction between the magnetic moments. This means that there exists a correlation between disorder and the form of the phase transition. In order to get a more representative result, scientists performed an average of a large number of different realizations—*replicas*—of interactions.

With the introduction of this new mathematical method, scientists started to discuss whether some properties of the model system were merely mathematical artifacts. Thus the extension of the Ising model brought about a tension between the emergence of interesting new properties and behaviors, and the characteristics of novel mathematical tools. It seems that there was more at stake than mere mathematical tractability. The Ising model as a model for studying cooperative phenomena comes with a conceptual framework that cannot be detached from the model. It consists of concepts such as equilibrium states, phase transitions and critical points. The Ising model was crucial in formulating a definite meaning to these concepts that had already been discussed in terms of order-disorder transformations (Kramers and Wannier 1941). In this context the Ising-model was treated as a mathematical object, and already in 1936 Rudolph Peierls showed the existence of a phase transition by making use of the Ising model. In the case of the spin glasses it was not clear if such systems would ever develop into equilibrium states and if they did so, what would the associated phase transition look like. The efforts the scientists put into the clarification of the question of whether the mathematical method leads to an artifact or not, shows the importance of the conceptual framework that comes with the Ising model. It also shows that in applying the model template the main idea of cooperativity is crucial, being conceived as a general property of systems consisting of a large number of interacting components.

In the following years, the replica method, and the meaning of the results gained by it, became one of the main foci in the study of the SK model (Parisi 1983). Interestingly, before the problems around the replica method were solved, parts of the SK model had already been applied to the study of neural networks.

4. The Hopfield Model

The Hopfield model was introduced in 1982 by the theoretical physicist John Hopfield. It is one of the most famous neural networks models and contributed to the flourishing of neural network modeling in 1980s. However, it was not the first model that derived from the Ising model the general idea that the interaction between the large numbers of neurons in the brain (cooperative phenomena) could give rise brain functions (collective phenomena).

Already in 1954 B.G. Cragg and H.N.V. Temperley reflected on the possible organizational principles among neurons:

The organization of the neurons in central nervous tissues is not understood, and while it is generally agreed that each neuron probably interacts with many of its neighbors, it has not been possible to predict any experimentally testable properties which would arise from interaction. Examples occur in physics, however, of unexpected properties that are entirely due to interaction. This paper examines a possible analogy between the organization of neurons and the kind of interaction among atoms which lead to co-operative processes. (Cragg & Temperley 1954, 85.)

The work of Cragg and Temperley provides an early example of a transfer of the Ising model into neuroscience: it takes from the Ising model the very general idea that the properties of neural networks result from interactions between neurons.

In a similar but more general way theoretical physicist E.R. Caianiello argued in his paper *Outline of a Theory of Thought-Processes and Thinking Machines*: “Our main guiding principle has been the conviction [...] that the human brain, tremendous in its complexity, yet obeys if one looks at the operation of individual neurons, dynamical laws that are not necessarily complicated, and that these laws are such as to engender in large neuronal assemblies collective modes of behavior” (Caianiello 1960, 240). Caianiello’s use of the concept ‘collective mode’ is interesting. He links cooperative to collective behavior in such a way that the former gives rise to the latter. The use of collective behavior also hints at a serious problem underlying the approach of using the Ising model as a kind of model template. Concepts such as phase transitions, critical points and equilibrium states to which a specific meaning is assigned in physics are devoid of a well-defined meaning when it comes to neuroscience. The more general notion of collective phenomena accounts for the emergence of the phenomena but leaves the question of their kind open.

4.1. Reconciling structural analogy and mathematical tractability

The Hopfield model of an associative memory (1982) was an important step forward in getting the analogy between neural networks and the Ising model, especially in its SK model variant, to work. Hopfield has explained that his basic idea of modeling an associated memory consisted in conceiving a system that develops into an attractor.¹

¹ In private communication.

This process can be visualized by water flowing from different directions into a sink. Hopfield “played” around with different ideas such as the Game of Life, which had been introduced by John Conway in 1970 for simulating the birth, reproduction, and death of living organisms. Making use of the Game of Life, Hopfield re-described the complex biological network as a network of interconnected binary variables performing simple rules. Hopfield’s focus was on the dynamics of the system but, as it turned out, his efforts were not successful. He could not find dynamical rules that would lead to the envisaged dynamic behavior. In the framework of the Ising model such a dynamic would correspond to the propagation of the interaction between the components and the establishment of a long-range order. In a later publication Hopfield reflected on this problem of finding general dynamical rules in biological systems:

The dynamical equations of quantum mechanics and quantum electrodynamics (and their classical equivalents when appropriate) are the essential elemental laws of physics which lead to biology. [...] But the real mysteries of biology lie in the way in which these dynamical laws of physics and the substrate of electrons, photons and nuclei on which they operate, produce the complex set of counter intuitive phenomena labeled with the term biology. (Hopfield 1986, 300.)

Such considerations did not come up in relation to the original work on the Ising model. Here the partition function was used in calculating the energy of the different states of the system as well as their probabilities. The problem faced by Hopfield was different. He was interested in the question of how the system would develop into an attractor associated with a stored pattern, and what kind of dynamic would guarantee such propagation into a stored pattern.

Hopfield's next attempt was to draw an analogy to the SK model. He formulated the question he wanted to answer in the following way:

In physical systems made from a large number of simple elements, interactions among large numbers of elementary components yield collective phenomena such as the stable magnetic orientations and domains in a magnetic system or the vortex patterns in fluid flow. Do analogous collective phenomena in a system of simple interacting neurons have useful "computational" correlates? (Hopfield 1982, 2554.)

And he added:

All modeling is based on details, and the details of neuroanatomy and neural function are both myriad and incompletely known. In many physical systems, the nature of the emergent collective properties is insensitive to the details inserted in the model [...]. In the same spirit, I will seek collective properties that are robust against change in the model details. (Ibid.)

Here we see how Hopfield makes use of the central property of the Ising model, which consists of it being independent from the details of the interactions and properties of the interacting components. In this quest of modeling an associative memory inspired by the modeling of physical systems, two elements turned out to be essential: 1. The introduction of a general dynamical rule, not necessarily related to the biological system but such that makes the energy function of the model system to decrease in the course of the dynamic. This behavior Hopfield interpreted as pattern recognition. 2. The couplings

between the neurons were chosen according to the SK model. This led to a great number of energy minima, which in the Hopfield model represent patterns stored into the system.

4.2. The mathematical structure of the Hopfield model

At first sight, the mathematical structure of the Hopfield model cannot be distinguished from the SK model. In the Hopfield model, the magnetic moments are turned into neurons, but they are similar to the magnetic moments in that they are conceived of as binary variables. Neurons in a biological network are non-linear units, summing up activation signals. They are rendered binary by introducing a threshold: if the summed up signals exceed a threshold, the neuron fires an action potential and goes over into a quiet state. The model neuron σ_i takes the value 1 in case it is active, and the value 0 if it is inactive. The state of each of the neurons is determined by its post-synaptic potential (PSP) h_i , produced by the activating signals arriving from all the other neurons to which it is connected. It is given by:

$$h_i = \sum_{j, j \neq i}^N J_{ij} S_j$$

As is already evident on the basis of the preceding quotes, Hopfield used a dynamical approach in exploring how cooperative phenomena could give rise to collective phenomena such as associative memory. As the partition function was not helpful in answering those questions, he introduced into his model a dynamical rule, the so-called Glauber dynamic (Glauber 1963). The Glauber dynamic is an algorithm not unique to neural networks. It models the effect of external agencies such as a heat bath or

a magnetic field on the Ising model. The version of the Glauber dynamic that Hopfield was using is a deterministic one in which the temperature $T=0$. In later work, the temperature was included to the model making the dynamic stochastic (e.g., Amit, Gutfreund and Sompolinsky 1987).

The synaptic efficiencies between the neurons are described by the set of parameters J_{ij} in which the information is stored using the Hebb learning rule (Hebb 1949). According to it the simultaneous activation of two connected neurons results in a strengthening of the synaptic coupling between the two neurons. This rule is formalized in the Hopfield model as follows:

$$J_{ij} = \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu}$$

The ξ_i^{μ} are variables that describe a pattern, i.e., a given configuration of active and inactive neurons. The number of patterns stored into the network is given by p , and in each pattern the number of neurons is equal to the total number of neurons in the network, N . Each of the patterns is associated with an energy minimum. The topology of the energy landscape shows a similar complexity as in the case of the SK model. This complex topology had been an important reason for drawing an analogy to the SK model. The large number of energy minima were potential patterns. In the Ising model, in contrast, only two patterns can be stored due to the maximum of two energy minima. The state of the whole system at time t is given by the set of all variables $\sigma = \{\sigma_1, \dots, \sigma_N\}$, which can take different values.

As in the case of the SK model, Hopfield introduced an energy function into his model:

$$E = -\frac{1}{2} \sum_{i \neq j} J_{ij} S_i S_j$$

It assigns an energy value E to each system configuration $\sigma = \{\sigma_1, \dots, \sigma_N\}$. This energy function has the same structure as the interaction energy in the case of the magnetic system. It can be shown that under the assumption of symmetric coupling $J_{ij} = J_{ji}$ and $J_{ii} = 0$, the energy function is governing the dynamics so that, starting from an incomplete input, the system develops under the course of the dynamics into an energy minimum, which is associated with one of the stored patterns.

Hopfield made use of the assumed structural similarities between the neural system and the SK model and combined them with a dynamical rule. According to it, the case of symmetric couplings would lead to a decrease of the energy function, meaning that the system approaches an attractor representing a stored pattern. The Glauber dynamic, even though it is not describing in a realistic way the dynamic in the biological neural networks, nevertheless made the model mathematically tractable. This had not been accomplished by the earlier attempts of modeling neural networks inspired by the Ising model.² Importantly, the Hopfield model did not result from an effort to draw a direct analogy to the Ising model, which clearly distinguishes it from the earlier attempts. Hopfield preserved in his model the general interactions between the components, which give rise to the observed phenomena (associative memory). In addition he selectively choose properties from the SK model such as the randomly distributed couplings and brought it together with a dynamic rule, that had the property of decreasing over time.

² An important exception is the Little model (Little 1974).

5. Model Templates in Model Transfer

The Ising model, quite against any expectations of its original author, became one of the most successful models in physics that is nowadays used to study phenomena so diverse that the branch of physics built into it is sometimes called ‘universal physics’ (Hughes 1999, 97–98).³ Although what seemed to be transmitted between different fields in the case of the Ising model were mathematical structures, the conceptual side of model transfer was equally important. The physicist Daniel Amit has described this conceptual fruitfulness of the Ising model in the following way:

[The Ising model] has been a birthplace and the testing ground for a treasure of new concepts in essentially all fields of physics. Such fundamental ideas as symmetry breaking, cooperative phenomena, order parameters, disorder parameters, critical exponents, symmetry restoration etc., have had their first explicit, precise articulation in the framework of this apparently simple, naïve model. (1989, 105.)

As the quote hints at, it was the precise mathematical formulation of the Ising model that enabled the consequent conceptual and theoretical development. In the 1930s there were discussions on ordered-disordered transitions but the Ising model was the first model to deliver a formal articulation of these ideas. It provided the springboard for further theoretical and mathematical innovation in terms of the development of the Ising model

³ On universality, see Batterman 2000.

into two or more dimensions, and into different off-springs such as the spin glass models (i.e., the Edwards–Anderson model and the SK model).

The notion of a model template aims to capture the intertwining of a mathematical structure and associated computational tools with theoretical concepts that, taken together, depict a general mechanism that is potentially applicable to any subject or field displaying particular patterns of interaction. In the case of the Ising model, the model template consists of such notions as cooperative phenomena, phase transition and long-term order embodied into the equations describing the interactions between the components of the system, the energy, and the order parameters. The mathematical and computational methods that were developed along the Ising model provided a closely integrated toolbox for the study of magnetic systems and properties. But what made the model applicable to a wide class of systems spanning, for example, from physics to chemistry, biology, and economy, is the way it describes diverse phenomena such as ferromagnetism, correlations in social structures (e.g., Stauffer and Aharony 1994), and cancer (e.g., Torquato 2011) as cooperative phenomena. Some important general characteristics of these kinds of phenomena motivated the application of the Ising model to them.

Let us unpack on a less formal level those features of the basic Ising model that made it a successful model template, as well as the changes on it that enabled its transfer into neuroscience via the SK model. As a model template, there are three central features in the Ising model. Firstly, it consists of magnetic moments pointing up and down, and taking values $+1$ and -1 . Secondly, the magnetic moments are not independent from each other but interact with their nearest neighbors that is described by a coupling term (which

in the case of the Ising model is constant). The interaction between the components leads to an internal magnetic field h . Thirdly, an energy function can be ascribed to the system that depends on the configuration of magnetic moments and the temperature. The resulting simple model could be used to model cooperative phenomena on a very general level. The interaction between the magnetic moments leads to the establishment of a long-range order resulting in a phase-transition. Even though parts of the Ising model were changed when applied in modeling disordered magnetic systems (i.e., spin glasses) or neural networks, the basic idea of cooperative phenomena and the associated mathematical core were preserved.

In the SK model, the spins are still modeled as binary variables. The interaction between the magnetic moments leads to an internal magnetic field h and an energy E is ascribed to the model system. The most important change in the SK model is associated with the couplings, which are not constant anymore. The simultaneously occurring non-ferromagnetic and ferromagnetic couplings in the SK model lead to a different kind of overall behavior. This shows that the general idea of cooperativity formulated in the Ising model is flexible enough to allow for changes in the couplings. What needed to be adjusted were the mathematical tools and the concepts associated with the notion of phase-transition. The partition function turned out to be much more complicated in the case of the SK model. The development and refinement of the methods to calculate the partition function was needed to attain a mathematically tractable model. Interestingly enough, even though the calculation of the partition function was complicated and the status of the method itself questioned, the SK model nevertheless provided a model

template for modeling neural networks. This shows that tractability does not need to be the crucial issue in model transfer.

As in the case of the model transfer between the Ising model and the SK model, what travelled in the case of the Hopfield model was the model template comprised of a particular type of a mathematical structure and its interpretation in terms of a general cooperative mechanism leading to collective phenomena. In the Hopfield model, the neurons were modeled as binary variables σ , and their interaction was modeled in the same way as in the case of the SK model, leading to a large number of energy minima. Yet the interpretation of the coupling was different. Making use of the Hebb rule, the Hopfield model stores the pattern in the couplings, which should be retrieved by the model. But in order to achieve the same properties as in the SK model, the patterns consist of randomly chosen values +1 and -1. The main difference between the SK model and the Hopfield model is due to the way the latter focuses on dynamics. Hopfield was not interested in the particular structure of the energy minima, which in his model became attractors. He was after the dynamic through which the system develops towards attractors. Because of this, he did not calculate the partition function but introduced instead a dynamic rule, the Glauber dynamic. It is of interest that, in order to make use of this dynamic, Hopfield had to restrict the couplings to symmetric couplings $J_{ij} = J_{ji}$, and as the result, the energy is decreasing over time. This is not a realistic assumption, but it makes the model computationally tractable.

6. Conclusions

What catches the eye in contemporary modeling practices is the remarkable similarity of the mathematical forms and computational methods used across different, even distant fields of study. Paul Humphreys has approached this phenomenon in terms of what he has coined as theoretical and computational templates. The distinguishing feature of computational templates is their tractability, which makes them attractive to use across different scientific fields and disciplines. However, the model transfer between the Ising model, the Sherrington–Kirkpatrick model and the Hopfield model cannot be accounted for by tractability alone. In this case, apart from mathematical forms and methods, a very general conceptual idea of the kind of mechanism of interaction involved was also transferred. Despite the obvious danger of inflating the notions of templates that Humphreys has introduced to the discussion on modeling, we call this abstract conceptual idea embedded into a mathematical form or method a *model template*. It is model templates, we suggest, that enable cross-disciplinary transfer, sensitizing us to perceive similar patterns across wide variety of different kinds of empirical systems, and simultaneously providing mathematical and computational tools for their analysis. As such they offer resources for further investigation and new theoretical insights.

The case of the Ising model also highlights three important aspects of the use of model templates in model transfer: Firstly, apart from their conceptual-mathematical core, the model templates need to be to some extent flexible, too. They are *evolving entities*, whose application to new fields typically requires adjustments at the theoretical and mathematical level. In the cases discussed above, this involved the introduction of

new methods that were important for the tractability of the model. Thus, secondly and in accordance with Humphreys, the tractability considerations are important drivers of model construction. The methods introduced, such as the replica method and the Glauber dynamic can themselves be considered as computational templates. Last but not least, it seems that one of the virtues of the Ising model is precisely it having been developed with a particular phenomenon in mind that helps understanding the cooperative phenomena. Other less-known and more complex systems are more easily understood when related to the well-understood simplification of ferromagnetic system. Thus model templates are not altogether detached from the empirical systems that they are used to study. Indeed, what seems characteristic of the contemporary modeling practice is the way it oscillates between considering model templates as models of some specific systems and bracketing the empirical content away in studying their formal properties and transferring them to new fields.

Acknowledgements:

We wish to thank the anonymous referee and Paul Humphreys for their extremely helpful comments.

References

- Amit, D. 1989. *Modeling Brain Function: The World of Attractor Neural Networks*, Cambridge, MA: Cambridge University Press.
- Amit, D., H. Gutfreund, and H. Sompolinsky. 1987. "Statistical Mechanics of Neural Networks Near Saturation," *Annals of Physics* 173 (1): 30–67.
- Batterman, R. 2000. "Multiple Realizability and Universality," *The British Journal for the Philosophy of Science* 51 (1): 115–45.
- Binder, K. and A.P. Young 1986. "Spin Glasses: Experimental Facts, Theoretical Concepts, and Open Questions," *Reviews of Modern Physics* 58 (4): 801–976.

- Brush, S. 1967. "History of the Lenz–Ising Model." *Reviews of Modern Physics* 38 (4): 883–93.
- Caianiello, E.R. 1960. "Outline of a Theory of Thought-Processes and Thinking Machines," *Journal of Theoretical Biology* 1 (2): 204–35.
- Cannella, V. and J.A. Mydosh 1972. "Magnetic Ordering in Gold-Iron Alloys," *Physical Review B* 6 (11): 4220–37.
- Cragg, B.G. and H.H. Temperley 1954. "The Organization of Neurons: A Cooperative Analogy," *Electroencephalography and Clinical Neurophysiology* 6 (1): 85–92.
- Darden, L. and N. Maull. 1977. "Interfield Theories," *Philosophy of Science* 44: 43–64.
- De Boer, J., E. Dal, and O. Ulfbeck, eds. 1986. *The Lesson of Quantum Theory*, Proceedings of the Niels Bohr Centenary Symposium, New York: Elsevier.
- Edwards, S.F. and P.W. Anderson 1975. "Theory of Spin Glass," *Journal of Physics F: Metal Physics* 5: 965–78.
- Glauber, R.J. 1963. "Time-Dependent Statistics of the Ising Model," *Journal of Mathematical Physics* 4: 294–307.
- Hebb, D. 1949. *The Organization of Behavior: A Neurophysiological Theory*, Mahwah: Erlbaum Books.
- Hopfield, J. 1982. "Neural Networks and Physical System with Emergent Collective Computational Abilities," *Proceedings of the National Academy of Sciences of the USA* 79 (8): 2554–58.
- — —. 1986. "Physics, Biological Computation, and Complementarity," in de Boer, Dahl, and Ulfbeck (1986, 295–314).
- Hughes, R.I.G. 1999. "The Ising Model, Computer Simulation, and Universal Physics," in Morgan and Morrison, eds. (1999, 97–145)
- Humphreys, P. 2002. "Computational Models," *Philosophy of Science* 69 (3): S1–S11.
- — —. 2004. *Extending Ourselves. Computational Science, Empiricism and Scientific Method*, Oxford: Oxford University Press.
- Humphreys, P. and C. Imbert, eds. 2011. *Representations, Models, and Simulations*. New York: Routledge.
- Ising, E. 1925. "A Contribution to the Theory of Ferromagnetism," *Zeitschrift für Physik* 31 (1): 253–58.
- Kitcher, P. 1981, "Explanatory Unification," *Philosophy of Science* 48: 507–31.
- Knuuttila, T. and A. Loettgers 2011. "The Productive Tension: Mechanisms vs. Templates in Modeling the Phenomena", in Humphreys and Imbert, eds. (2011, 3–24)
- Kramers, H.A. and G.H. Wannier 1941. "Statistics of the Two-Dimensional Ferromagnet. Part I," *Physical Review* 60: 252–62.
- Little, W. 1974. "The Existence of Persistent States in the Brain," *Mathematical Bioscience* 19: 102–20.
- Morgan, M.S. and M. Morrison, eds. 1999. *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge, MA: Cambridge University Press.
- Niss, M. 2005. "History of the Lenz–Ising Model 1920–1950: From Ferromagnetic to Cooperative Phenomena," *Archive for History of Exact Sciences* 59 (3): 267–318.
- Onsager, L. 1944. "Crystal Statistics: A Two-Dimensional Model with Order-Disorder Transition" *Physical Review* 65: 117–25.

- Parisi, G. 1983. "Order Parameter for Spin-Glasses," *Physical Review Letters* 50 (23): 1946–48.
- Peierls, R. 1936. "Statistical Theory of Absorption with Interaction Between the Absorbed Atoms," *Proceedings of the Cambridge Philosophical Society* 32: 471–502.
- Sherrington, D. and S. Kirkpatrick 1978. "Infinite-Ranged Models of Spin-Glasses," *Physical Review B* 17 (11): 4384–403.
- Stauffer, D. and Aharony, A. 1994. *Introduction to Percolation Theory*, 2nd ed., London: Taylor and Francis.
- Torquato, S. 2011. "Toward an Ising Model of Cancer and Beyond," *Physical Biology* 8: 1–22.