Taking Causal Modeling to a Next Level: Self-Modeling Networks Adding Adaptivity to Causality

Jan Treur
Social AI Group, Department of Computer Science, Vrije Universiteit Amsterdam
email: j.treur@vu.nl

***Abstract*— Causal modeling is an intuitive, declarative way of modeling. Due to the universal character of causality, in principle it applies to practically all disciplines. In spite of this seemingly very wide scope of applicability, there are also serious limitations and challenges that stand in the way of applicability. This concerns in particular cases where dynamics and adaptivity play a role. This paper addresses these challenges by exploiting the notion of self-modeling network that has been developed from a Network Science perspective. Adaptivity is obtained by adding to a given causal base network, a self-model which represents part of the base network's causal structure. Moreover, this construction can easily be iterated so that multiple orders of adaptation can be covered as well. This indeed takes causal modeling to a next level in more than one way. Therefore, in this way dynamics and adaptivity are also covered well, which substantially widens the scope of applicability of causal modeling.**

Keywords*— *causal modeling, self-modeling network, network reification, adaptive social network, controlled adaptation

I. INTRODUCTION

Causal modelling provides a declarative approach that has a long tradition in Artificial Intelligence; e.g., [1-4]. One of the challenges, however, is that causal modelling involving cyclic paths in causal graphs poses difficulties; therefore many approaches to causal modelling limit themselves to Directed Acyclic Graphs (DAG's). More in general, to avoid temporal complexity, dynamics is often not addressed in approaches based on causal networks, neither for the causal effects on nodes, nor for the network structure itself. The difficulty to allow cyclic paths in a causal network is one consequence of this form of abstraction from dynamics of the nodes in a causal network. Another consequence of abstracting from dynamics is that distinctions in timing and asynchrony of causal effects (i.e., how fast causal effects actually are effectuated) cannot be made, whereas often such differences in timing and asynchrony are crucial for real-world processes modelled by a causal network. Finally, within causal models, not only the nodes but also the causal relations are usually considered static, they cannot change over time. This excludes many adaptive real-world processes from the scope of applicability for causal modeling.

In the meantime, working from the perspective of Network Science, new approaches have been developed that can be used to overcome the above-mentioned limitations of causal modeling. In particular, in this paper it will be addressed how both within-network dynamics (dynamics of the node states) for causal network models and adaptivity of the causal relations can be addressed using the network-oriented modeling approach developed in [5-7].

Using this approach as introduced for within-network dynamics in [5], the dynamic perspective is based on a continuous time dimension, represented by real numbers, so that all nodes have state values (also represented by real numbers) that vary over time. The added temporal dimension enables modelling by cyclic causal networks as well, and also timing of causal effects can be modelled in detail and differently per node, so that also asynchronous processes are covered. Due to this, causal modeling can be used for causal networks that contain cycles, such as many networks modelling mental or brain processes, or networks describing social interaction processes (for example, in social media). Moreover, in [5, 7] it is shown how – supported by a dedicated software environment – networks with these within-network dynamics can be specified by declarative means, by mathematical relations and functions; the modeler does not need to address procedural descriptions nor program code.

In addition to these within-network dynamics, another useful element from the network-oriented modeling perspective is the notion of self-modeling network or reified network introduced in [6-8]. This is a network that includes a self-model for part of its own network structure in the form of nodes that represent certain network structure characteristics such as connection weights or excitability thresholds. Any (base) network can be extended by including such a self-model, which can be considered to be at a next level, compared to the base network; this step is also called network reification; e.g., [6-8]. This construction for networks in particular was inspired by another long-standing tradition in AI, namely that of meta-programming and metalevel architectures; e.g., [9-13]. Having such self-models within a network enables to model adaptation of the network structure by the within-network dynamics of the self-model representing this network structure. As the latter can be specified by declarative means in the form of mathematical relations and functions, also adaptivity of the network structure can be specified in a similar declarative manner. To support the modeler, a dedicated software environment (described in [7], Ch 9) is available that also applies to self-modeling networks.

In this paper, the perspective pointed out above will be illustrated in more detail. First in Section 2 the network-oriented modeling approach based on self-modeling networks will be briefly introduced. Next, in Section 3 it will be illustrated for an example of a multilevel second-order adaptive causal (social) network model for bonding

by (faked) homophily, while in Section 4 an example of a simulated scenario for this model is described. Finally, Section 5 is a discussion.

II. Modeling Adaptivity by Self-Modeling Networks

In this section, the network-oriented modeling approach by self-modeling networks used is briefly introduced in two steps.

A. Network-Oriented Modeling by Temporal-Causal Networks

As in this approach *nodes Y* in a network have activation values *Y(t)* that are dynamic over time *t*, they serve as state variables and will usually be simply called *states*. For these dynamics, the states are considered to affect each other by the connections within the network; therefore these connections are interpreted here as causal relations. This has been inspired partly by how in Philosophy of Mind networks of mental states and their causation relations are described; e.g., [14]. In line with this, following [5, 7], a basic *temporal-causal network structure* is characterised by:

- *Connectivity characteristics*
  Connections from a state *X* to a state *Y* and their weights $\boldsymbol{\omega}_{X,Y}$
- *aggregation characteristics*
  For any node *Y*, some combination function $\mathbf{c}_Y(..)$ defines aggregation that is applied to the impacts $\boldsymbol{\omega}_{X_i,Y}X_i(t)$ on *Y* from its incoming connections from states $X_1, \cdots, X_k$
- *timing characteristics*
  Each state *Y* has a speed factor $\boldsymbol{\eta}_Y$ defining how fast it changes for given causal impact

Here, the states X_i and *Y* have activation levels $X_i(t)$ and $Y(t)$ that vary (often within the [0, 1] interval) over time, described by real numbers *t*. These dynamics are described by the following difference (or differential) equations that incorporate in a canonical manner the network characteristics $\boldsymbol{\omega}_{X,Y}$, $\mathbf{c}_Y(..)$, $\boldsymbol{\eta}_Y$:

$$Y(t+\Delta t) = Y(t) + \boldsymbol{\eta}_Y[\mathbf{c}_Y(\boldsymbol{\omega}_{X_1,Y}X_1(t), \cdots, \boldsymbol{\omega}_{X_k,Y}X_k(t)) - Y(t)]\,\Delta t \quad (1)$$

for any state *Y* and where $X_1, \cdots, X_k$ are the states from which *Y* gets its incoming connections. The equations (1) are useful for simulation purposes and also for analysis of properties of the emerging behaviour of temporal-causal networks. The overall combination function $\mathbf{c}_Y(..)$ for state *Y* is taken as the weighted average of some of the available basic combination functions $\mathbf{c}_j(..)$ by specified weights $\boldsymbol{\gamma}_{j,Y}$, and parameters $\boldsymbol{\pi}_{1,j,Y}$, $\boldsymbol{\pi}_{2,j,Y}$ of $\mathbf{c}_j(..)$, for *Y*:

$$\mathbf{c}_Y(V_1, \ldots, V_k) = \frac{\boldsymbol{\gamma}_{1,Y}\,\mathbf{c}_1(V_1,\ldots,V_k) + \ldots + \boldsymbol{\gamma}_{m,Y}\,\mathbf{c}_m(V_1,\ldots,V_k)}{\boldsymbol{\gamma}_{1,Y} + \ldots + \boldsymbol{\gamma}_{m,Y}} \quad (2)$$

Such equations (1), (2) are hidden in the dedicated software environment that can be used for simulation and analysis; see [7], Ch 9. This software environment is freely downloadable from URL

https://www.researchgate.net/project/Network-Oriented-Modeling-Software.

Combination functions are similar to the functions used in a static manner in the deterministic Structural Causal Model perspective described, for example, in [3, 4, 15]. However, in the Network-Oriented Modelling approach described here they are used in a dynamic manner. For example, Pearl [3], p. 203, denotes nodes by V_i and combination functions by f_i (although he uses a different term for these functions). In the following quote he points at the issue of underspecification concerning aggregation of multiple connections, as in the often used graph representations the specification of combination functions f_i for nodes V_i, is lacking:

'Every causal model *M* can be associated with a directed graph (…) This graph merely identifies the endogeneous and background variables that have a direct influence on each V_i; it does not specify the functional form of f_i.' [3], p. 203

Therefore, in addition to graph representations for connectivity, at least aggregation in terms of combination functions has to be addressed, as indeed is done for temporal-causal networks, in order to avoid this problem of underspecification. That is the reason why aggregation in terms of combination functions is part of the definition of the network structure for temporal-causal networks, in addition to connectivity in terms of connections and their weights and timing in terms of speed factors.

As part of the software environment, a large number > 35 of useful basic combination functions are included in a Combination Function Library, and also a facility to easily indicate any function composition of any available basic combination functions in the library. One of the combination functions from this library used for states *Y* in the example network model described in Section 3 is:

the *Euclidean combination function* $\mathbf{eucl}_{n,\lambda}(V_1, \ldots, V_k)$ defined by

$$\mathbf{eucl}_{n,\lambda}(V_1, \ldots, V_k) = \sqrt[n]{\frac{V_1^{\,n} + \cdots + V_k^{\,n}}{\lambda}} \quad (3)$$

where $\boldsymbol{n}$ is the order and $\boldsymbol{\lambda}$ a scaling factor and $V_1, \ldots, V_k$ are the impacts from the states from which the considered state *Y* gets incoming connections.

In Section 3, it will be explained in more detail how the combination function $\mathbf{eucl}_{n,\lambda}(\ldots)$ is used to model social contagion. Social contagion makes that states of connected persons such as emotions or opinions, causally affect each other; e.g., (Levy and Nail, 1993).

The above concepts (the characteristics $\boldsymbol{\omega}_{X,Y}$, $\boldsymbol{\gamma}_{j,Y}$, $\boldsymbol{\pi}_{i,j,Y}$, $\boldsymbol{\eta}_Y$) enable to design network models and their dynamics in a declarative manner, based on mathematically defined functions and relations for them. Note that for each state *Y*, all characteristics $\boldsymbol{\omega}_{X,Y}$, $\boldsymbol{\gamma}_{j,Y}$, $\boldsymbol{\pi}_{i,j,Y}$, $\boldsymbol{\eta}_Y$ mentioned above causally affect the activation level of *Y*, as also can be seen from equations (1) and (2). Each of these characteristics do that causing in their own way from a specific role, either for connectivity, for aggregation or for timing. Below, this observation will also turn out useful in the context of self-models to address adaptivity.

B. Using Self-Modeling Networks to Model Adaptive Networks

Realistic network models are usually adaptive: often some of their network characteristics $\boldsymbol{\omega}_{X,Y}$, $\boldsymbol{\gamma}_{j,Y}$, $\boldsymbol{\pi}_{i,j,Y}$, $\boldsymbol{\eta}_Y$ change over time. For example, for mental networks often the connections are assumed to change by hebbian learning [16] and for social networks, it is often assumed that connections between persons change through a bonding by homophily principle [17-19].

Adaptive networks are often modeled in a hybrid manner by considering two different types of separate models that interact with each other: a network model for the base network and its within-network dynamics, and a numerical model for the adaptivity of the network structure characteristics of the base network. The latter dynamic model is usually specified in a format outside the context of network modeling: in the form of some adaptation-specific procedural or algorithmic programming specification used to run the difference or differential equations underlying the network adaptation process.

In contrast, by including *self-models*, a network-oriented conceptualisation similar to what was described above, can also be applied to adaptive networks to obtain a declarative description using mathematically defined functions and relations for them as well; see [6, 7]. This works through the addition of new states to the network (called *self-model states*) which represent network characteristics by network states. Then the causal impacts of these characteristics on a state Y as mentioned above can be modelled as causal impacts from such self-model states. This brings the causal impacts from these characteristics on a state Y in the standard form of a causal model where via causal connections nodes affect other nodes.

More specifically, adding a self-model for a temporal-causal base network is done in the way that for some of the states Y of the base network and some of the network structure characteristics for connectivity, aggregation and timing (i.e., some from $\boldsymbol{\omega}_{X,Y}$, $\boldsymbol{\gamma}_{j,Y}$, $\boldsymbol{\pi}_{i,j,Y}$, $\boldsymbol{\eta}_Y$), additional network states $\mathbf{W}_{X,Y}$, $\mathbf{C}_{j,Y}$, $\mathbf{P}_{i,j,Y}$, $\mathbf{H}_Y$ (*self-model states* or *reification states*) are introduced and connected to other states:

a) Connectivity self-model

- Self-model states $\mathbf{W}_{X,Y}$ are added representing connectivity characteristics, in particular connection weights $\boldsymbol{\omega}_{X,Y}$

b) Aggregation self-model

- Self-model states $\mathbf{C}_{j,Y}$ are added representing aggregation characteristics, in particular combination function weights $\boldsymbol{\gamma}_{j,Y}$
- Self-model states $\mathbf{P}_{i,j,Y}$ are added representing aggregation characteristics, in particular combination function parameters $\boldsymbol{\pi}_{i,j,Y}$

c) Timing self-model

- Self-model states $\mathbf{H}_Y$ are added representing timing characteristics, in particular speed factors $\boldsymbol{\eta}_Y$

This step of adding a self-model to a base network is also called *network reification*. If such self-model states are dynamic, they describe adaptive network characteristics. In a graphical 3D-format, such self-model states are depicted at a next level (also called *reification level*), where the original network is at a *base level*. As an example, the weight $\boldsymbol{\omega}_{X,Y}$ of a connection from state X to state Y can be represented (at a next reification level) by a self-model state named $\mathbf{W}_{X,Y}$ (e.g., for an objective representation) or $\mathbf{RW}_{X,Y}$ (e.g., for a subjective representation).

Having self-model states to model an adaptation principle in a network-oriented manner is only a first step. To fully model a certain adaptation principle by a self-modeling network, the dynamics of each self-model state itself and its effect on a corresponding target state Y have to be specified in a network-oriented manner by the three general standard types of network structure characteristics a) *connectivity*, b) *aggregation*, and c) *timing*:

Connectivity for the self-model states in a self-modeling network

For the self-model states, their *connectivity* in terms of their incoming and outgoing connections has two different functions:

Effectuating its special effect from its specific role

The *outgoing downward causal connections* from the self-model states $\mathbf{W}_{X,Y}$, $\mathbf{C}_{j,Y}$, $\mathbf{P}_{i,j,Y}$, $\mathbf{H}_Y$ to state Y represent the specific causal impact (their special effect from their specific role) each of these self-model states has on Y. These downward causal impacts are standard per role, and make that the adaptive values $\mathbf{W}_{X,Y}(t)$, $\mathbf{C}_{j,Y}(t)$, $\mathbf{P}_{i,j,Y}(t)$, $\mathbf{H}_Y(t)$ at t are actually used for the adaptive characteristics of the base network in equations (1) and (2).

Indicating the input for the adaptation principle as specified in b)

The *incoming upward or leveled connections* to a self-model state are used to specify the *input* needed for the particular adaptation principle that is addressed.

Aggregation for the self-model states in a self-modeling network

For the self-model states, their aggregation characteristics have one main aim:

Expressing the adaptation principle by a mathematical function

For the *aggregation* of the incoming causal impacts for a self-model state, provided as indicated in a), a specific combination function is chosen *to express the adaptation principle* in a declarative mathematical manner.

Timing for the self-model states in a self-modeling network

For the self-model states, their timing characteristics have one main aim:

Expressing the adaptation speed for the adaptation principle by a number

Finally, like any other state, self-model states have their own *timing* in terms of speed factors. These speed

factors are used as the means to express the adaptation speed.

As a base network extended by including a self-model is also a temporal-causal network model itself, as has been shown in [7], Ch 10, this self-modeling construction can easily be applied iteratively to include self-models of multiple (reification) levels. This can provide higher-order adaptive network models, and has turned out quite useful to model, for example, within Cognitive Neuroscience plasticity and metaplasticity (e.g., [20-23]) in a unified form by a second-order adaptive mental causal network with three levels, one base level and a first- and a second-order self-model level for causation concerning plasticity and metaplasticity, respectively, as shown in [7], Ch 4.

In the current paper, the notion of a multi-level self-modeling network will be illustrated by a higher-order adaptive social network model. In this model, in addition to the Euclidean combination function described in Section 2.1, two other combination functions from the library are used:

the advanced logistic sum combination function $\mathbf{alogistic}_{\sigma,\tau}(V_1, \ldots, V_k)$ defined by:

$$\mathbf{alogistic}_{\sigma,\tau_{\mathbf{log}}}(V_1, \ldots, V_k) = [\frac{1}{1+e^{-\sigma(V_1+\cdots+V_k-\tau_{\mathbf{log}})}} - \frac{1}{1+e^{\sigma\tau_{\mathbf{log}}}}](1+e^{-\sigma\tau_{\mathbf{log}}}) \quad (4)$$

where σ is a steepness parameter and $\tau_{\mathbf{log}}$ a threshold parameter and V_1, …, V_k are the impacts from the states from which the considered state Y gets incoming connections

the simple linear homophily combination function $\mathbf{slhomo}_{\alpha,\tau_{\mathbf{hom}}}(V_1, V_2, W)$ defined by

$$\mathbf{slhomo}_{\alpha,\tau_{\mathbf{hom}}}(V_1, V_2, W) = W + \alpha\, W\,(1-W)\,(\tau_{\mathbf{homo}} - |V_1 - V_2|) \quad (5)$$

where α is an amplification parameter and $\tau_{\mathbf{hom}}$ a tipping point parameter and V_1, V_2 are a person's representations of the two persons' states involved and W represents the weight of their connection

Here, $\mathbf{slhomo}_{\alpha,\tau_{\mathbf{hom}}}(V_1, V_2, W)$ is used to model bonding based on (faked) homophily by internal connection weight representations, and $\mathbf{alogistic}_{\sigma,\tau_{\mathbf{log}}}(\ldots)$ to model control of the bonding. Bonding based on homophily [17-19] is the social network adaptation principle that is sometimes expressed by

'Birds of a feather flock together'

This expresses how being 'birds of a feather' or 'being alike' (modeled by state values V_1 and V_2 for the two persons not differing much) causally affects the connection between two persons. Note that the homophily tipping point $\tau_{\mathbf{hom}}$ is the point where the difference between the states of the two individuals (represented by $|V_1 - V_2|$) turns an increase of bonding (outcome > W) into a decrease (outcome < W), and conversely. In Section 4 this tipping point is set at 0.25: so in that case a difference $|V_1 - V_2| < 0.25$ has as causal effect that the connection will be strengthened (increase of W), whereas a difference $|V_1 - V_2| > 0.25$ has as causal effect that it will be weakened (decrease of W).

This shows an example of how for a social application domain, within a causal network, states can have a causal effect on network connections. By applying a self-modeling network model, this form of causation (for adaptation of connections through bonding by homophily) together with the causation between states in the base network (for social contagion) is addressed in a unified manner by one overall (two-level) causal network model, in contrast to the commonly used hybrid modeling approach to adaptive networks pointed out above in the second paragraph of this Section 2.2. Moreover, in Section 3 it will be shown how also a third level for the control of the adaptation process can be incorporated within such a self-modeling causal network, thus obtaining a three-level network model unifying within one causal model the base network dynamics with adaptation of the connections of the base network and the control of that adaptation.

III. A Social Causal Network with Controlled Adaptation

To illustrate the use of self-modeling networks to incorporate in a unified manner both dynamics and (multi-order) adaptivity in a causal model, this section presents an adaptive causal network model for controlled bonding based on homophily by using subjective representations (some of which are based on fake input). The presented causal network model integrates three types of interacting processes, modeled within the causal model at three different levels:

The considered social base network itself with its (within-network) dynamics for social contagion [24]

Change of this social network over time based on bonding by homophily [17-19]: first-order social network adaptation

Control of the first-order social network adaptation: second-order social network adaptation

In contrast to what is usually done, for example, also in [19], here the bonding is not assumed to depend on the objective states for the two persons, but on how these states are perceived and represented by the persons through the formation of subjective state representation states. By controlling the formation of these subjective state representation states, indirectly the bonding is affected; contrarely, if you don't take care to acquire information about the other person, then you miss a good reason for stronger or weaker bonding. To cover this, the above three types of processes have been modeled by a second-order adaptive causal network model based on a multi-level self-modeling network using a first-order self-model (for formation of the subjective state representation states and for the bonding based on them) and a second-order self-model (for the control of the formation of the subjective representation states). That offers some room to model cheating about one's own properties, as regularly happens in real life: by faking an own state, the other person will make a false representation for it, which then will affect that person's bonding in a false manner.

The model's connectivity is depicted in Fig. 1 by an example for two persons, one of which is faking his or her properties in order to achieve successful bonding. In this 3D picture, each of the three planes models one of the

three types of processes mentioned above; for an explanation of the states, see Table 1.

Table 1 Types of states in the introduced controlled adaptive social network model

$\mathbf{S}_A$	Objective state *Z* of person A
$\mathbf{S}_B$	Objective state *Z* of person B
$\mathbf{FS}_B$	Objective state of person B faking state *Z* of person A
$\mathbf{RS}_{A,A}$	Subjective representation of person A for state *Z* of person A
$\mathbf{RS}_{B,B}$	Subjective representation of person B for state *Z* of person B
$\mathbf{RS}_{A,B}$	Subjective representation of person B for state *Z* of person A
$\mathbf{RS}_{B,A}$	Subjective representation of person A for state *Z* of person B
$\mathbf{RFS}_{B,B}$	Subjective representation of person B for his or her faked state *Z*
$\mathbf{RW}_{A,B}$	Subjective representation of person A for the weight of the connection from person A to person B
$\mathbf{RW}_{B,A}$	Subjective representation of person B for the weight of the connection from person B to person A
$\mathbf{CC}_{A,B}$	Control state for communication from A to B for the state *Z* of A: representation of the weight of the connection from $\mathbf{RS}_{A,A}$ to $\mathbf{RS}_{A,B}$
$\mathbf{CC}_{B,A}$	Control state for communication from B to A for the state *Z* of B: representation of the weight of the connection from $\mathbf{RS}_{B,B}$ to $\mathbf{RS}_{B,A}$
$\mathbf{CO}_{A,B}$	Control state for observation by B for the state *Z* of A observed by B: representation of the weight of the connection from $\mathbf{S}_A$ to $\mathbf{RS}_{A,B}$
$\mathbf{CO}_{B,A}$	Control state for observation by A for the state *Z* of B observed by A: representation of the weight of the connection from $\mathbf{S}_B$ to $\mathbf{RS}_{B,A}$

The types of connections used at and between the three levels within this network model are shown in Table 2. Here *Z* is a type of state of a person, for example, how often the person listens to a certain type of music; to keep the notations simple, this type is left out of them; if needed, the *Z* could be used as an additional subscript.

At the base level, social contagion is modelled by intralevel connections (depicted by black arrows in the lower plane in Fig. 1) such as $\mathbf{S}_A \rightarrow \mathbf{S}_B$, $\mathbf{FS}_B \rightarrow \mathbf{S}_A$, and $\mathbf{S}_A \rightarrow \mathbf{FS}_B$. Here the last connection models B faking by intentionally listening to the same type of music as A just at the moments that A can observe it. In contrast to $\mathbf{FS}_B$, state $\mathbf{S}_B$ indicates how much B normally listens to that type of music. In the simulated scenario, $\mathbf{S}_A$ will have high values and $\mathbf{S}_B$ low values, whereas by copying $\mathbf{S}_A$ also $\mathbf{FS}_B$ gets high values.

Within the first-order self-model, each person has subjective internal representation states of other persons' states *Z* and the of state *Z* of her or himself, and also of his or her connections to others. This first-order self-model is modeled in the middle plane. For example, person A's internal representation state for person B having state *Z* is modeled by state representation $\mathbf{RS}_{B,A}$, and A's subjective representation of his or her connection to B is modeled by connection weight representation $\mathbf{RW}_{A,B}$.

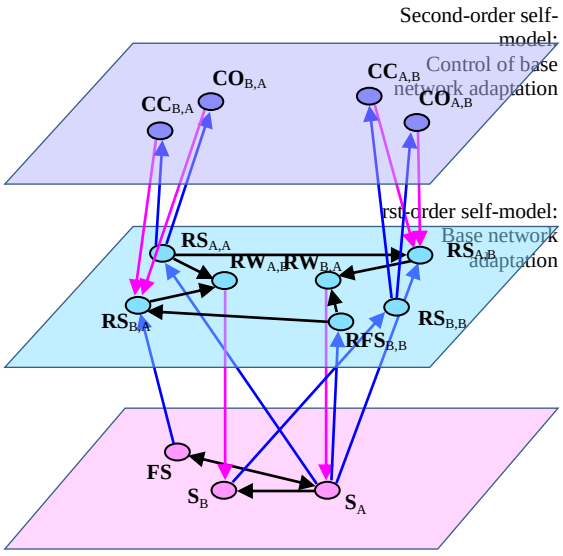


Figure 1 Overview of the connectivity of the second-order adaptive social network model for bonding by homophily for two persons A and B, where B is faking the homophily for A.

There are two pathways that contribute to formation of state representations such as $\mathbf{RS}_{A,B}$. First, these representations can be obtained through observation of $\mathbf{S}_A$ by B. This is modeled by an upward interlevel connection $\mathbf{S}_A \rightarrow \mathbf{RS}_{A,B}$ from the base network to the first-order self-model. As B is faking his or her base state, observation by A is modeled *not* by a connection $\mathbf{S}_B \rightarrow \mathbf{RS}_{B,A}$ but by connection $\mathbf{FS}_B \rightarrow \mathbf{RS}_{B,A}$.

A second pathway for a person B to get information on person A's state is through communication between persons. For example, if A communicates his or her subjective representation $\mathbf{RS}_{A,A}$ of the own state $\mathbf{S}_A$ to B (e.g., 'I often play this type of music!'), this is modeled by an intralevel connection $\mathbf{RS}_{A,A} \rightarrow \mathbf{RS}_{A,B}$ within the middle plane for the first-order self-model. Also in the communication, B is faking; therefore communication from B to A is *not* modeled by a connection $\mathbf{RS}_{A,B} \rightarrow \mathbf{RS}_{B,A}$, but by connection $\mathbf{RFS}_{B,B} \rightarrow \mathbf{RS}_{B,A}$ (so that B may falsely communicate 'What a coincidence, I also often play that type of music!').

Table 2 Connections in the controlled adaptive social network model and their explanation

Intralevel connections	
$\mathbf{S}_A \rightarrow \mathbf{S}_B$	Social contagion from A to B for state *Z*
$\mathbf{FS}_B \rightarrow \mathbf{S}_A$	Social contagion from B's faked state for *Z* to A
$\mathbf{S}_A \rightarrow \mathbf{FS}_B$	Faking contagion from state *Z* of A to faked state *Z* of B
$\mathbf{RS}_{A,A} \rightarrow \mathbf{RS}_{A,B}$	Communication of state *Z* from A to B
$\mathbf{RFS}_{B,B} \rightarrow \mathbf{RS}_{B,A}$	Communication of faked state *Z* from B to A
$\mathbf{RS}_{A,A} \rightarrow \mathbf{RW}_{A,B}$	Effect of represented state *Z* of A by A on the connection from A to B (bonding by homophily)
$\mathbf{RS}_{B,A} \rightarrow \mathbf{RW}_{A,B}$	Effect of represented state *Z* of B by A on the connection from A to B (bonding by homophily)
$\mathbf{RFS}_{B,B} \rightarrow \mathbf{RW}_{B,A}$	Effect of represented faked state *Z* of B by B on the connection from B to A (bonding by homophily)
$\mathbf{RS}_{A,B} \rightarrow \mathbf{RW}_{B,A}$	Effect of represented state *Z* of A by B on the connection from B to A (bonding by homophily)
Interlevel connections	

$\mathbf{S}_A \rightarrow \mathbf{RS}_{A,A}$	Impact of observation of A's state Z by A on A's representation of A's state Z	Upward from base network to first-order self-model
$\mathbf{S}_B \rightarrow \mathbf{RS}_{B,B}$	Impact of observation of B's state Z by B on B's representation of B's state Z	
$\mathbf{S}_A \rightarrow \mathbf{RS}_{A,B}$	Impact of observation of A's state Z by B on B's representation of A's state Z	
$\mathbf{FS}_B \rightarrow \mathbf{RS}_{B,A}$	Impact of observation of B's faked state Z by A on A's representation of B's state Z	
$\mathbf{RW}_{A,B} \rightarrow \mathbf{S}_B$	Effectuation of base connection weight for social contagion from state Z of A to state Z of B	Downward from first-order self-model to base network
$\mathbf{RW}_{B,A} \rightarrow \mathbf{S}_A$	Effectuation of base connection weight for social contagion from faked state Z of B to state Z of A	
$\mathbf{RS}_{A,A} \rightarrow \mathbf{CC}_{B,A}$	Communication control monitoring connection for A	Upward from first-order self-model to second-order self-model
$\mathbf{RS}_{B,B} \rightarrow \mathbf{CC}_{A,B}$	Communication control monitoring connection for B	
$\mathbf{RS}_{A,A} \rightarrow \mathbf{CO}_{B,A}$	Observation control monitoring connection for A	

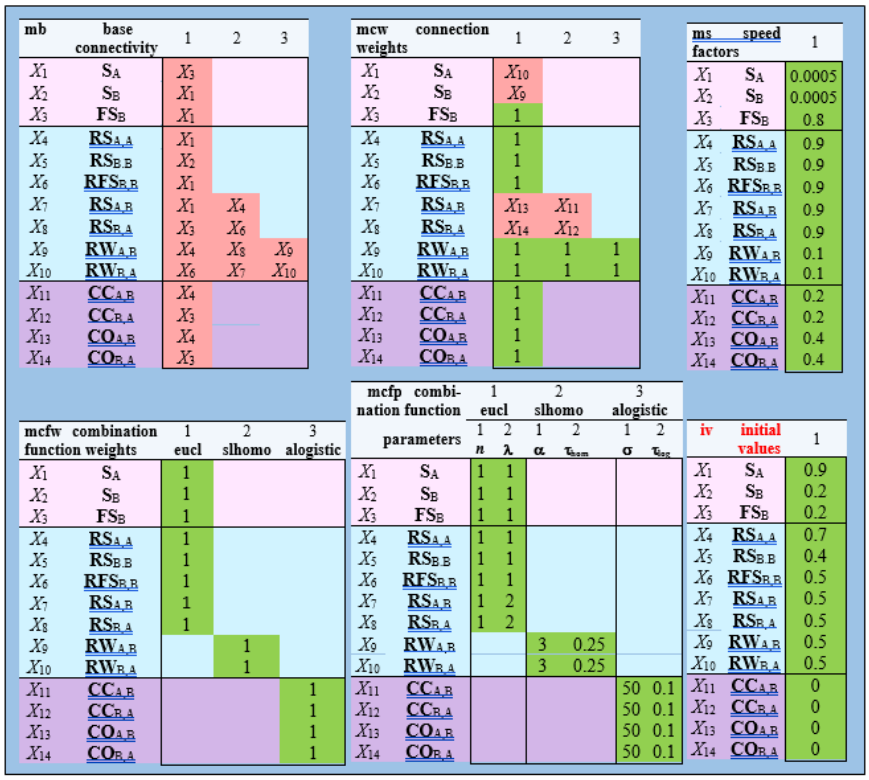
$\mathbf{RW}_{A,B}$ and $\mathbf{RS}_{B,A} \rightarrow \mathbf{RW}_{A,B}$ within the first-order self-model. The connection representations by **RW**-states in turn affect the social contagion within the social network, which is modeled by downward interlevel connections $\mathbf{RW}_{A,B} \rightarrow \mathbf{S}_B$ and $\mathbf{RW}_{B,A} \rightarrow \mathbf{S}_A$ from the first-order self-model in the middle plane to the base network.

To control the social network adaptation processes, two types of control actions are considered in particular:

controlling the observation of state Z from person A by person B is modeled by control state $\mathbf{CO}_{A,B}$ and from person B by person A is modeled by control state $\mathbf{CO}_{B,A}$

controlling the communication about state Z from person A to person B, modeled by control state $\mathbf{CC}_{A,B}$ and the communication about state Z from person B to person A, is modeled by control state $\mathbf{CC}_{B,A}$

Activation of a communication control state makes that the related connection in the first-order self-model in the middle plane gets a high value (1 or close to 1); this is achieved by interlevel connections from control states to **RS**-states in the first-order self-model. For example, activation of communication control state $\mathbf{CC}_{A,B}$ makes that the connection $\mathbf{RS}_{A,A} \rightarrow \mathbf{RS}_{A,B}$ from A's state $\mathbf{RS}_{A,A}$ to B's state $\mathbf{RS}_{A,B}$ gets a high value (1 or close to 1) so that



Box 1

Full specification of the adaptive self-modeling causal network model by role matrices for all (connectivity, aggregation and timing) characteristics causally affecting the network states

As indicated, person A's representation of her or his connection to person B is modeled by $\mathbf{RW}_{A,B}$. It is assumed that for the bonding by homophily adaptation principle, the adaptive change of the represented connection for A to B depends on the internal representation states $\mathbf{RS}_{B,A}$ and $\mathbf{RS}_{A,A}$. Therefore, this adaptation is supported by intralevel connections $\mathbf{RS}_{A,A} \rightarrow$ the transfer of information by communication happens; this is modeled by interlevel connection $\mathbf{CO}_{A,B} \rightarrow \mathbf{RS}_{A,B}$. This can be considered as B asking A for the information about him or herself, upon which A communicates this information. Similarly, activation of an observation control state $\mathbf{CO}_{A,B}$ makes that the connection $\mathbf{S}_A \rightarrow \mathbf{RS}_{A,B}$ from A's state $\mathbf{S}_A$ to B's state $\mathbf{RS}_{A,B}$ gets a high value (1 or close

to 1) so that the transfer of information by observation takes place; this is modeled by connection $\mathbf{CO}_{A,B} \rightarrow \mathbf{RS}_{A,B}$. In the case modeled here, control states such as $\mathbf{CC}_{A,B}$ and $\mathbf{CO}_{A,B}$ themselves may become active depending on B's state $\mathbf{RS}_{B,B}$; this is modeled by connections $\mathbf{RS}_{B,B} \rightarrow \mathbf{CC}_{A,B}$ and $\mathbf{RS}_{B,B} \rightarrow \mathbf{CO}_{A,B}$. But this may be addressed in many other ways as well, including externally determined control, for example, by enabling or allowing observation or communication (only) at specific time slots.

To specify a network model according to the approach described in [7], as discussed in Section 2, three types of network characteristics are to be covered: *connectivity*, *aggregation* and *timing* characteristics. Any state in the network is causally affected by all of such characteristics, each from its own specific role. Following the role matrices specification format defined in [7] (pp. 39-41, 89), they are specified by role matrices as shown in Box 1 which are used as input for the dedicated software environment to automatically obtain the simulation discussed In Section 4.

More specifically, *role matrices* indicate in rows successively for all network states, the factors that causally affect them from the different roles. So in the row for a state *Y*, in each column a causal relation is specified affecting state *Y* for the role described by that role matrix. In this way, role matrices describe the network model by mathematical relations and functions.

In the first place, concerning *connectivity* roles, each state is causally affected by the other states from which it has incoming connections and by the weights of these connections. In role matrix **mb** (see Box 1), for each state it is indicated from which other states it has incoming connections from the same or a lower level. In role matrix **mcw**, it is indicated what are the connection weights for the connected states indicated in **mb**. If these weights are static, their value is indicated, in green shaded cells (here always 1), but if the connection weight is adaptive, instead of a number the self-model state representing this weight is indicated in role matrix **mcw**. This can be seen (cells shaded in a peach-red colour) in **mcw** for the incoming connections for the first two states X_1 and X_2, and for the incoming connections for the states X_7 and X_8. Indicating these adaptive value representations, defines the downward connections of Fig. 1. From the *timing role*, also its speed factor causally affects a state; they are shown in Box 1 (role matrix **ms**, which actually is a vector).

In the lower part of Box 1, showing the *aggregation* roles causally affecting a state, it can be seen which states use which combination functions (role matrix **mcfw**) and which parameter values for them (role matrix **mcfp**). In addition to the five role matrices for the different roles of causal impacts, the initial values for the example simulation are also shown in Box 1, which may be considered as initial causal impacts.

IV. Simulation: Faking Homophily for Bonding

In this section, a simulation of a simulated example scenario will be discussed to illustrate the introduced second-order adaptive causal social network model for faking homophily. In Fig. 2 the simulation for the example scenario is shown. Here the states $\mathbf{S}_X$ are slowly changing whereas the connection representations in the form of the **RW**-states are changing faster. It indeed can be seen that for A and B both directional connection representations $\mathbf{RW}_{A,B}$ and $\mathbf{RW}_{B,A}$ start to gradually increase from time point 5 on to reach values above 0.7 which in the long run eventually reach a value (close to) 1. These changes of the connections are a consequence of the homophily principle, as the values of state $\mathbf{S}_A$ of A and the faked states $\mathbf{FS}_B$ and $\mathbf{RFS}_{B,B}$ for B quickly get close to each other; note that the tipping point for similarity set was 0.25, so a difference between the relevant representation states < 0.25 is strengthening a connection.

In Fig. 2, also the roles that are played by the control states in the form of the **CO**- and **CC**-states and by the **RS**-states for subjective representations can be seen. The two lines that start at 0 and get close to 1 around or soon after time 10 indicate the control states $\mathbf{CO}_{A,B}$ and $\mathbf{CO}_{B,A}$ (light green) for observation and $\mathbf{CC}_{A,B}$ and $\mathbf{CC}_{B,A}$ (light blue) for communication, respectively. This makes that at that time their mutual observation and communication channels $\mathbf{S}_A \rightarrow \mathbf{RS}_{A,B}$ and $\mathbf{FS}_B \rightarrow \mathbf{RS}_{B,A}$, and $\mathbf{RS}_{A,A} \rightarrow \mathbf{RS}_{A,B}$ and $\mathbf{RFS}_{B,B} \rightarrow \mathbf{RS}_{B,A}$ get weights close to 1. This implies that then they indeed both observe and communicate to each other about the type of music they usually listen to. These control states are triggered in this example scenario because each of the persons automatically observes his or herself and therefore they quickly (before time point 4) form representation states $\mathbf{RS}_{A,A}$ and $\mathbf{RS}_{B,B}$ of their own **S**-states concerning music (the red lines, starting at 0.4 for B and at 0.7 for A).

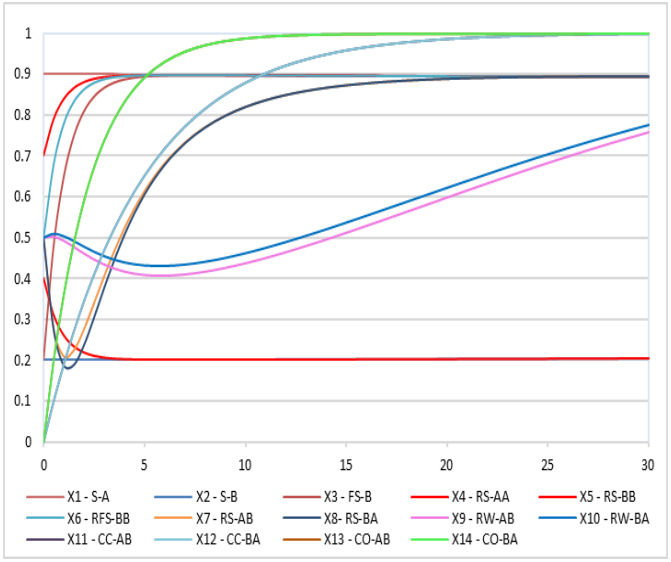


Figure 2 Outcomes for the example scenario simulation

Because of these communication and observation actions, the mutual subjective representations $\mathbf{RS}_{A,B}$ of B about A (the dark green line) and $\mathbf{RS}_{B,A}$ of A about B (the orange line) based on fake information are formed, and around time 20 reach levels close to 0.9. Only now these subjective representations have been formed in a controlled manner, the homophily principle can start to work, as the bonding works through the (subjective) representation **RS**-states, not through the (objective) states $\mathbf{S}_X$ themselves. More specifically, from the moment on that the subjective representations of A about B and A's own subjective representation about her- or himself get closer

than 0.25 (which is just before time point 5), her/his self-model representation $\mathbf{RW}_{A,B}$ of her connection to B (the pink line) starts to gradually increase. Similarly, the effect of the subjective representations of B for A and B's own subjective self-model representation about him or herself, on the subsequent increase of his representation $\mathbf{RW}_{B,A}$ of this connection to A (the blue line) can be noted. Before that point in time their connections were not increasing, but instead go slightly downward; this illustrates the effect of the control via the subjective self-model representation states on the adaptation.

V. Discussion

Causal modeling combines two quite useful properties. In the first place, it is an intuitive, declarative way of modeling supported by often used graphical representations. Secondly, due to the universal character of causality, in principle it should apply to practically all scientific disciplines. However, limitations for dynamics and adaptivity stand in the way of applicability in many domains. In this paper it was discussed how these challenges can be addressed by exploiting the notion of self-modeling network developed from a Network Science perspective [6, 7]. Self-modeling causal networks cover dynamics of the states of the nodes as well as adaptivity of these causal relations. Here adaptivity of a base network is obtained by explicit representations of the characteristics of the causal relations in the form of a self-model added to the base network. These self-modeling causal networks are specified in a declarative manner by mathematical relations and functions, and provide a causal network addressing the adaptation. Therefore, this approach indeed takes causal modeling to a next level so that now dynamics and adaptivity are also covered by a unified causal perspective. By an illustration for a controlled adaptive social causal network model, it has been shown how this widens the scope of applicability of causal modeling.

Another topic that illustrates the applicability of the causal modeling approach based on self-modeling networks well is plasticity and metaplasticity within Cognitive Neuroscience, as described, for example, in empirical literature such as [20-23]. In [7], Ch 4, it is shown how this can be modeled as a self-modeling causal network incorporating a first-order self-model for plasticity and a second-order self-model for metaplasticity.

Such multi-level self-modeling causal networks incorporate different types of causation. In the first place this covers causation between base states, as is a familiar form of causation known from traditional causal models. This is also the form of causation usually focused on (for mental states) within Philosophy of Mind, such as in [14]. However, in self-modeling network models there is also causation from these base states to other types of states representing causal relations, and back. Such forms of causation have to occur as soon as causal relations can change in the world, as such change should be caused by something. In turn, such changes causally affect the future processes.

So, for adaptive cases, from a completeness of causation perspective such less familiar forms of causation cannot be avoided, and have direct relations to what actually happens in the world. Indeed, for example in empirically focussed Cognitive Neuroscience literature such as [20-23], it is described in some detail how states and processes addressing plasticity and metaplasticity are realised by specific (changing) brain configurations and causal relations for them. So, self-models are not just artificial modeling concepts created by some fantasy: they relate to real counterparts of them in the physical world. In that sense, it may be claimed that self-modeling causal networks actually exist in the world, at least for this context of Cognitive Neuroscience. A similar illustration for the biological domain can be found in [7], Ch. 7, addressing a five-level self-modeling causal network model describing different stages in an evolutionary process. Here the different types of states and causation in the self-modeling causal network have counterparts in the physical world in the form of (changing) configurations and processes as described in literature from Biology.

The presented approach allows declarative modeling of dynamic and adaptive behaviour of multiple orders of adaptation from a unified causal perspective. Traditionally, declarative modeling approaches are a strong focus of AI. There are two longstanding themes in AI to which the work presented here relates in particular: causal modeling as already mentioned [1-4] and metalevel architectures and metaprogramming [9-13]. As discussed, a main contribution to the causal modeling area is that this is extended by dynamics and adaptivity of the causal modeling, addressing both the dynamics of the causal effects and the adaptive dynamics of the causal relations themselves. A main contribution to the area of metalevel architectures and metaprogramming is that now network models are covered as well in the form of self-modeling networks, while traditionally the focus in this area is mainly on logical, functional and object-oriented modeling or programming approaches; e.g., [10].

In relation to the area of Neural Networks within AI, the network-oriented modeling approach described here distinguishes itself by a multidisciplinary Network Science focus on causality and adaptation within empirical natural and human-directed sciences. In contrast, the area of Neural Networks has its main focus on artificial neural networks to solve optimisation challenges and on their computational efficiency. Another important distinction is the notion of self-modeling network which is the main focus in the current paper. However, there are also some technical elements in common, for example, the format of the canonical difference equation (1) (see Section 2.1) used here can be considered a form of socalled recurrent network as also used in the Neural Networks area. But a difference here is the use of speed factors per node which enables to model different nodes that are not necessarily synchronous in their dynamics. This asynchrony is usually needed to model real-world processes as these are not often synchronous and can even involve entirely different time scales. This explicit way to model differentiated timing is not a common practice in the Neural Networks area within AI.

From a more theoretical side, following Ashby [26] in [25] Section 3.1 it has been shown that any state-

determined dynamical system (as defined in [26] and also used in [27]) can be described by a set of first-order differential equations, and conversely. Moreover, in [25], Section 3.2 it has also been shown how any set of first-order differential equations can be (re)modeled by a temporal-causal network model. It has been shown in [7], Ch 10 that any self-modeling network obtained by adding a self-model to a temporal-causal network is iself also a temporal-causal network. Therefore, these methods can also be applied to adaptive processes: any description of an adaptation process by a state-determined system or by first-order differential equations can be rewritten as a self-model in temporal-causal network format. This provides evidence from a more theoretical analysis perspective that the approach discussed here has a wide scope of applicability.

There are still some more interesting challenges that can be addressed. A first challenge is to explore other interesting cases of higher-order adaptation and to investigate whether self-modeling causal networks indeed are suitable to model them. Within Cognitive Neuroscience, from an empirical perspective the notions of plasticity and metaplasticity have been introduced [20-23], relating to first- and second-order adaptation. It has been found how these can be modeled by a second-order self-modeling network; see [6] and [7], Ch 4. Similarly, it has been described how second-order adaptive social networks for bonding by homophily can be modeled by self-modeling networks; see [6] and [7], Ch 6. However, in general higher-order adaptation for social networks has not been addressed well in the literature. As an exception, in [28, 29] the notion of inhibiting adaptation for networks has been described, which refers to some form of second-order adaptive social networks. This applies, for example, to terrorist network organisations. It would be interesting to investigate whether and how such second-order social networks can also be described as self-modeling causal networks.

Within Biology, some literature can be found on how evolutionary processes can be described as higher-order adaptation; e.g., [30, 31]. It has been shown in [7], Ch 7, how one case study concerning pregnancy and disgust can be modeled by a fourth-order adaptive self-modeling causal network model. It is interesting to address more case studies in this area. Moreover, in Hofstadter [32] claims that the notion of Strange Loop underlies human intelligence. This is described in [32] informally as a form of self-modeling of multiple levels, where for some n, the n^{th} level is equal to the base level, so that the levels form a cycle. It has been found that this also can be modeled by a self-modeling network; see [7], Ch 8 for an example for a mental network and [33] for an example for a social network. However, the notion of Strange Loop could be explored for more cases.

Finally, as mentioned the research described in the current paper follows the multidisciplinary perspective of Network Science. Therefore, the focus is on adaptation principles known from nature and described in empirical disciplines such as Biology, Neuroscience, Cognitive Science or Social Sciences. In contrast, it may be an interesting challenge to investigate how some wellknown artificial methods for machine learning can be modeled by self-modeling networks. As the self-modeling network approach provides a declarative perspective on modeling adaptation processes, this might provide more declarative descriptions of such artifical methods, which usually are described in a procedural manner by algorithms. As pointed out in one of the paragraphs above, from a theoretical perspective this should be possible. But it would be interesting to see how this would actually look like for some examples. It might provide a more clear modeling separation of the conceptual core of a machine learning method and the procedural optimisation involved. As an example, in this way backpropagation for artificial neural networks could be modeled in a network-oriented manner with gradient descent as conceptual core plus an efficient procedure to do the required calculations; e.g., [34], Ch 7.

REFERENCES

[1] Kuipers, B.J.: Commonsense reasoning about causality: Deriving behavior from structure, Artificial Intelligence 24**,** 169--203 (1984)

[2] Kuipers, B.J., Kassirer, J.P.: How to discover a knowledge representation for causal reasoning by studying an expert physician. In: Proc. of the Eighth International Joint Conference on Artificial Intelligence, IJCAI'83, pp. 49--56. William Kaufman, Los Altos, CA, (1983)

[3] Pearl, J.: Causality. Cambridge University Press (2000)

[4] Wright, S.: Correlation and Causation. Journal of Agricultural Research 20, 557--585 (1921)

[5] Treur, J.: Network-Oriented Modeling: Addressing Complexity of Cognitive, Affective and Social Interactions. Springer Publishers (2016)

[6] Treur, J.: Modeling Higher-Order Adaptivity of a Network by Multilevel Network Reification. Network Science 8, S110--S144 (2020)

[7] Treur, J.: Network-oriented modeling for adaptive networks: Designing Higher-order Adaptive Biological, Mental and Social Network Models. Cham, Switzerland: Springer Nature Publishing (2020)

[8] Treur, J.: Multilevel Network Reification: Representing Higher Order Adaptivity in a Network. In: Aiello, L., Cherifi, C., Cherifi, H., Lambiotte, R., Lió, P., Rocha, L. (eds), Proc. of the 7th Int. Conf. on Complex Networks and their Applications, ComplexNetworks'18, vol. 1. Studies in Computational Intelligence, vol. 812, pp. 635--651, Springer Nature (2018)

[9] Bowen, K.A., and Kowalski, R.: Amalgamating language and meta-language in logic programming. In Logic Programming, K. Clark and S. Tarnlund, Eds. Academic Press, New York, pp. 153--172 (1982)

[10] Demers, F.N., Malenfant, J.: Reflection in logic, functional and objectoriented programming: a Short Comparative Study. In IJCAI'95 Workshop on Reflection and Meta-Level Architecture and their Application in AI, pp. 29--38 (1995)

[11] Sterling, L., Shapiro, E.: The Art of Prolog. MIT Press (Ch 17, pp. 319--356) (1996)

[12] Sterling, L., Beer, R.: Metainterpreters for expert system construction. Journal of Logic Programming 6, 163--178 (1989)

[13] Weyhrauch, R.W.: Prolegomena to a Theory of Mechanized Formal Reasoning. Artificial Intelligence 13, 133–170 (1980)

[14] Kim, J. (1996). Philosophy of Mind. Westview Press (1996)
[15] Mooij, J.M., Janzing, D., Schölkopf, B.: From Differential Equations to Structural Causal Models: the Deterministic Case. In: Nicholson, A., and Smyth, P. (eds.), Proceedings of the 29th Annual Conference on Uncertainty in Artificial Intelligence (UAI-13), pp. 440--448. AUAI Press (2013)
[16] Hebb, D.O.: The organization of behavior: A neuropsychological theory. Wiley (1949)
[20] Abraham, W.C., Bear, M.F.: Metaplasticity: the plasticity of synaptic plasticity. Trends in Neuroscience 19(4), 126--130 (1996)
[21] Garcia, R.: Stress, Metaplasticity, and Antidepressants. Current Molecular Medicine 2, 629--638 (2002)
[22] Magerl, W., Hansen, N., Treede, R.D., Klein, T.: The human pain system exhibits higher-order plasticity (metaplasticity). Neurobiology of Learning and Memory 154, 112--120 (2018)
[23] Robinson, B.L., Harper, N.S., McAlpine, D.: Meta-adaptation in the auditory midbrain under cortical influence. Nat. Commun. 7, 13442 (2016)
[24] Levy, D.A., Nail, P.R.: Contagion: A theoretical and empirical review and reconceptualization. Genetic, social, and general psychology monographs 119(2), 233--284 (1993)
[25] Treur, J.: On the Applicability of Network-Oriented Modeling Based on Temporal-Causal Networks: Why Network Models Do Not Just Model Networks. Journal of Information and Telecommunication 1(1), 23—40 (2017)
[26] Ashby, W.R.: Design for a Brain. Chapman and Hall, London (second extended edition). First edition, 1952 (1960)
[27] Port, R.F., van Gelder, T.: Mind as motion: Explorations in the dynamics of cognition. Cambridge, MA: MIT Press (1995)
[17] McPherson, M., Smith-Lovin, L., and Cook, J.M.: Birds of a feather: homophily in social networks. Annu. Rev. Sociol., 27, 415--444 (2001)
[18] Pearson, M., Steglich, C., Snijders, T.: Homophily and assimilation among sport-active adolescent substance users. Connections 27(1), 47—63 (2006)
[19] Sharpanskykh, A., and Treur, J.: Modelling and Analysis of Social Contagion in Dynamic Networks. Neurocomputing 146, 140--150 (2014)

[28] Carley, K.M.: Inhibiting Adaptation. In Proceedings of the 2002 Command and Control Research and Technology Symposium, pp. 1--10. Monterey, CA: Naval Postgraduate School (2002).
[29] Carley, K.M.: Destabilization of covert networks. Comput Math Organiz Theor 12, 51–66 (2006).
[30] Fessler, D.M.T., Clark, J.A., and Clint, E.K.: Evolutionary Psychology and Evolutionary Anthropology. In: The Handbook of Evolutionary Psychology, (D.M. Buss, ed.), pp. 1029--1046. Wiley and Sons (2015).
[31] Fessler, D.M.T., Eng, S.J., & Navarrete, C.D.: Elevated disgust sensitivity in the first trimester of pregnancy: Evidence supporting the compensatory prophylaxis hypothesis. Evolution & Human Behavior 26(4), 344-351 (2005).
[32] Hofstadter, D.R.: Gödel, Escher, Bach. New York: Basic Books (1979)
[33] Anten, J., Earle, J., Treur, J. : An Adaptive Computational Network Model for Strange Loops in Political Evolution in Society. Proc. of the 20th International Conference on Computational Science, ICCS'20, vol. 2, pp. 604–617. Lecture Notes in Computer Science, vol. 12138. Springer Nature (2020).
[34] Rojas, R.: Neural Networks. Berlin: Springer Verlag (1996).