Modeling the emergence of informational content by adaptive networks for temporal factorisation and criterial causation

Jan Treur

Vrije Universiteit Amsterdam, Social AI Group, Netherlands

Received 16 July 2020; accepted 28 October 2020
Available online 17 February 2021

Abstract

Propagated activation of neurons through their network is an important process in the brain. Another crucial part of neural processing concerns adaptation over time of characteristics of this network such as connection strengths or excitability thresholds. This adaptation can be slow, as in learning from multiple experiences, or it can be fast, as in memory formation. These adaptive network characteristics can be considered informational criteria for activation of a neuron. This then is viewed as a form of emergent information formation. Activation of neurons is determined by such information via a process termed criterial causation. In the current paper, the relationship of criterial causation with the principle of temporal factorisation for the dynamics of the world in general is explored. Temporal factorisation describes how the world represents information about its past in its present state, which then in turn determines the world’s future. In the paper, it is shown how these processes are analysed in more detail and modeled by (adaptive) network models.

Keywords: Criterial causation; Temporal factorisation; Adaptive network model; Informational content

1. Introduction

Two main aspects of information processing in the brain involve activation and plasticity (Hebb, 1949; Tse, 2013). This makes a picture of neural processes based on both dynamics within the network of neurons and dynamics of this network. Tse (2013, 2018) considers the settings of the network structure characteristics encoded by a specific brain configuration as serving as criteria for the incoming signals to get a neuron activated; this is called criterial causation. The criteria are viewed as a form of information emerging in the brain, called by Tse (2013), p. 259 ‘physically realised informational criteria’; when in the future these criteria are fulfilled, the neuron will activate.

So, it is not only activation of neurons that changes over time, also network characteristics representing the criteria change over time. These changes are influenced by incoming patterns in the past. This change can be slow or fast in comparison to propagation of activation of neurons: slow as for learning from multiple experiences, or fast as in formation of memories, which is termed ‘rapid resetting of these criteria’ (Tse, 2013; Tse, 2018). This form of network

https://doi.org/10.1016/j.cogsys.2020.10.018
1389-0417/© 2021 The Author(s). Published by Elsevier B.V.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
adaptation creates *emerging information*, and activation of neurons is determined by the created information (Tse, 2013; Tse, 2018).

This paper discusses how the specific criterial causation perspective can be considered to be a special case of temporal factorisation which addresses the world dynamics more in general (Treur, 2007a; Treur, 2007b). Mediating state properties encode information in the world state on the past of the world, and they are determining the future world patterns. The temporal factorisation principle generalises criterial causation to emerging formation of information and usage of this information as a mechanism of the world dynamics in general. Viewed in this way, criterial causation in the brain makes use of that more general principle. The relation between the two is then that the temporal factorisation principle describes how the world dynamics creates and exploits emerging formation of information in its state, and criterial causation addresses particularly how the brain generates and exploits emerging information in its state. In each of these two cases this emergent information determines future patterns.

In this paper, Section 2 makes a more detailed analysis of temporal factorisation and criterial causation and how they relate. Section 3 briefly introduces the adaptive temporal-causal network modeling approach used in the paper. Next, such network models are presented to illustrate both principles. In Section 4, some examples of temporal factorisation are addressed by (nonadaptive) network modeling, including example simulations. Section 5 addresses how for temporal-causal networks criteria can be specified for criterial causation, and Section 6 presents an adaptive network model for criterial causation; in Section 7 results of an example simulation are discussed. In Section 8 it is discussed how the informational content of a mediating state for temporal factorisation and of emerging criteria for criterial causation can be defined, based on the concept of temporal relational specification from Philosophy of Mind. In Section 9, for more technically interested readers some further formalisation details are described in terms of first- and second-order reified temporal predicate logic. Section 10 is a discussion.

2. Temporal factorisation versus criterial causation

In this section the notions of Temporal Factorisation (Section 2.1) and Criterial Causation (Section 2.2) are introduced, and their relationship is discussed (Section 2.3).

2.1. Temporal factorisation

The principle of temporal factorisation (Treur, 2007a; Treur, 2007b) expresses that if in the world a temporal relationship \( a \Rightarrow b \) from past pattern \( a \) to future pattern \( b \) holds, then this can be factorised by a state property \( p \) into temporal relationships \( a \Rightarrow p \) from the past pattern \( a \) to present state \( p \) and \( p \Rightarrow b \) from present state \( p \) to future pattern \( b \). State property \( p \) is termed a *mediating state property*

for \( a \Rightarrow b \). For an illustration, see Fig. 1. Note that the principle expresses that for any such a relationship \( a \Rightarrow b \) a world state property \( p \) exists such that the temporal relationships \( a \Rightarrow p \) and \( p \Rightarrow b \) hold:

\[
[a \Rightarrow b] \Rightarrow \exists p[a \Rightarrow p \land p \Rightarrow b]
\]

(1)

This mediating state property \( p \) describes a configuration of the world state in the present. It can (and often will) be not one simple state property, but a combination of a number of state properties. The notion of pattern is kept a bit informal here (however, see Section 9 for formalisations), but it can be understood as a property of possible temporal traces, where a temporal trace is a sequence of states for all time points. For example, in Fig. 1 pattern \( a \) is illustrated for specific traces \( a_1 \) to \( a_m \) fulfilling that pattern, and pattern \( b \) for traces \( b_1, \ldots, b_n \). Such a property of temporal traces can be expressed as some temporal statement; see Section 9 for more explicit descriptions.

The word factorisation is explained algebraically as follows. Suppose the set of possible past patterns are indicated by the set \( PP \), the set of possible future patterns by set \( FP \), and the set of possible present states by \( PS \). Then the temporal relationship \( a \Rightarrow b \) can be described by an operator or function \( f \), assigning future patterns \( b \) to past patterns \( a \)

\[
f : PP \rightarrow FP
\]

(2)

Temporal factorisation expresses that there are functions \( g \) and \( h \) with

\[
g : PS \rightarrow FP
\]

(3)

\[
h : PP \rightarrow PS
\]

such that function \( f \) is factorised by \( g \) and \( h \) as follows:

\[
f = goh : PP \rightarrow PS \rightarrow FP
\]

(4)

where \( goh \) denotes the functional composition of functions \( g \) and \( h \): first \( h \) is applied then \( g \).

According to the temporal factorisation principle, the present world state configuration \( p \) encodes all relevant information from the past; then the world can ignore the past if it generates a future temporal pattern. So, it essentially claims that world state configurations, from an information-al perspective are rich enough to encode in one (present) state all information on the past which is future-relevant. In (Treur, 2007a; Treur, 2007b), by many examples from Physics and from Cognitive Science it is
illustrated how temporal factorisation works for world dynamics. Note that the converse implication
\[ \exists p [a \Rightarrow p \& p \Rightarrow b] \Rightarrow [a \Rightarrow b] \]  
(5)
of temporal factorisation is a logically valid statement in an almost trivial manner (by transitivity of implication), and as such does not add new information. In contrast, the temporal factorisation principle itself is nontrivial.

Temporal factorisation can also be applied iteratively so that also \( a \Rightarrow p \) and \( p \Rightarrow b \) are factorised further in
\[ a \Rightarrow p_1 \& p \Rightarrow p_2 \& p_2 \Rightarrow b \]  
(6)

In this way, temporal factorisation basically expresses that to describe in more detail global flows of dynamics from patterns in the past to patterns in the future, intermediate states can be found that in a sense break up the global process in smaller causal steps. These intermediate states carry informational content about the past and about the future.

An interesting question is how temporal factorisation relates to perspectives on dynamics from Descartes (1634), Laplace (1825), Ashby (1952), van Gelder and Port (1995). This will be discussed in some detail. From these, Descartes (1634) claims that there are ‘laws of nature’ as relationships between world states for different points in time, in the form that past world states determine future world states. He emphasizes that only the laws of nature determine the dynamics of the world (after a starting time):

‘Know, then, first that by “nature” I do not here mean some deity or other sort of imaginary power. Rather, I use that word to signify matter itself, insofar as I consider it taken together with all the qualities that I have attributed to it, and under the condition that God continues to preserve it in the same way that He created it. For from that alone (i.e., that He continues thus to preserve it) it follows of necessity that there may be many changes in its parts that cannot, it seems to me, be properly attributed to the action of God (because that action does not change) and hence are to be attributed to nature. The rules according to which these changes take place I call the “laws of nature.”’
(Descartes, 1634, Ch 7: On the Laws of Nature of this New World)

This is called the clockwork universe view. It describes how relationships between world states over time (laws of nature) drive world dynamics, in the sense that patterns for past world states \( a \) determine patterns for future world states \( b \), or \( a \Rightarrow b \).

Laplace (1825)’s view is slightly different and a bit more refined; it is summarized by this quote:

‘We may regard the present state of the universe as the effect of its past and the cause of its future.’
(Laplace, 1825)

Both views of Descartes and Laplace describe a deterministic world. Although temporal factorisation somehow relates to the above two perspectives, in contrast it does not claim a fully deterministic world. That principle just applies to world aspects that are deterministic; this is indicated by the condition \( a \Rightarrow b \).

To make the relations to temporal factorisation more precise, Descartes’ view is formulated by
\[ D : \text{‘Dynamics works by temporal past} \rightarrow \text{– future patternrelationships } a \Rightarrow b.\]  
(7)
and Laplace’s view is formulated by
\[ L : \text{‘Dynamics works by temporal past} \rightarrow \text{– present and present} \rightarrow \text{– future relationships } a \Rightarrow p \text{ and } p \Rightarrow b.\]  
(8)

So, \( D \) relates to the antecedent of the principle of Temporal Factorisation \( TF \), and \( L \) to the consequent; temporal factorisation logically connects \( D \) and \( L \) in the following way: Descartes’ view \( D \) and \( TF \) together (by Modus Ponens) entail Laplace’s \( L \):
\[
\begin{align*}
D_r : & \quad a \Rightarrow b \\
TF : & \quad [a \Rightarrow b] \Rightarrow \exists p [a \Rightarrow p \& p \Rightarrow b] \\
\hline
L : & \quad \exists p [a \Rightarrow p \& p \Rightarrow b] \quad \text{-- Modus Ponens}
\end{align*}
\]  
(9)

which may be summarized as
\[ D \& TF \Rightarrow L \]  
(10)
is logically true. Temporal factorisation can be considered a more general assumption that explains the shift from Descartes’s perspective \( D \) to Laplace’s perspective \( L \). While these two perspectives both assume deterministic world processes, temporal factorisation itself does not, due to the conditional. It therefore is more general and also applicable for nondeterministic (parts of the) world(s) for which \( D \) and \( L \) do not hold.

In (van Gelder & Port, 1995), after Ashby (1952) dynamics is based on the concept of state-determined system:

‘...its current state always determines a unique future behaviour ... the future behaviour cannot depend in any way on whatever states the system might have been in before the current state.’
(van Gelder & Port, 1995), p. 6.

This can be described by the implication \( p \Rightarrow b \), which focuses on the step from present state \( p \) to future pattern \( b \), while it does not consider how any past pattern \( a \) leads to the present state \( p \).

One example, in a more or less similar form also used in (Treur, 2007a), illustrating temporal factorisation is the following. Imagine that in a real-world or gaming situation in the current state a door is encountered that in the past was locked by somebody and not unlocked since then. It then can be opened in the future when someone brings the right key. So, the past pattern \( a \) can be described as
\[ 'at some past time point someone locked the door and since then nobody unlocked it' \]  
(11)
and the future pattern \( b \) as
'if at some future time point someone carries the right key, then the door can be unlocked'

Moreover, mediating state property $p$ is here

'\textit{the door is locked by its specific lock}'

(13)

From the temporal relationship $a \Rightarrow b$, here instantiated as (11) $\Rightarrow$ (12), by temporal factorisation it can be derived that some mediating state property $p$ exists for which the temporal past-to-present relationship $a \Rightarrow p$ and the present-to-future relationship $p \Rightarrow b$ hold. In this case, this state property $p$ specifies that the door is locked by its specific lock. Here $a \Rightarrow p$ specifies that if somebody locked the door in the past and nobody unlocked it since then, it is locked in the present, while $p \Rightarrow b$ expresses that if it is locked in the present, when someone brings the appropriate key in the future it can be unlocked. Viewed informationally, within the world in the present state $p$, the specific lock locking the door represents lock information, and when in the future the appropriate key in accordance with that lock information, is brought, the door unlocks. This case illustrates how the mediating state property $p$ encodes information, and this information determines the world’s future by pattern $b$, described as ‘when the right key is applied, the door unlocks’.

Here humans are the actors whose actions encode the information in the physical world, since the lock and also the key were made by humans. So, in this case humans informationise the world, or, in other words, the world becomes more informational because of human intervention. Another example of such human-intervened informationisation of the world can be found in the story of Little Thumb who drops pebbles to mark his route so that he can find back his way home. In contrast to such human-intervened cases, according to the temporal factorisation principle, the world (as an actor itself) is creating and exploiting without human intervention a similar form of encoding information on past patterns in the form of present world state configurations.

\subsection*{2.2. Criterial causation}

Neural processing involves propagation of activation of neurons; e.g., (Hebb, 1949; Tse, 2013). This propagation depends on network characteristics of the network of neurons. For example, it depends on synaptic connection strengths of connections between neurons and on the excitability thresholds of neurons. The configurations in the brain defined by these characteristics serve as criteria for incoming signals for a neuron, determining whether to make it fire: if the pattern of these incoming signals fulfills the criteria, the neuron will fire, otherwise it will not fire. Note that this has some similarity to the door lock and key example used above to illustrate the concept of temporal factorisation: the criteria correspond to the specific lock, and the pattern of incoming signals to the key needed to unlock the door; see Section 2.3 for more details. By Tse (2013, 2018) this process in the brain is termed criterial causation. Here the criteria represent information represented in the specific brain configuration: ‘physically realised informational criteria’ (Tse, 2013), p. 259; in the future, the neuron will fire when these criteria are fulfilled.

'I have proposed a three-stage model of a neuronal mechanism that underlies mental causation and free will, according to which (1) new physical/informational criteria are set in a neuronal circuit on the basis of preceding physical/mental processing at $t_1$, in part via a mechanism of rapid resetting that effectively changes the inputs to a postsynaptic neuron. These changes can be driven volitionally or nonvolitionally, depending on the neural circuitry involved. (2) At $t_2$, inherently variable inputs arrive at the postsynaptic neuron, and (3) at $t_3$ physical/informational criteria are met or not met, leading to postsynaptic neural firing or not.’ (Tse, 2013, p. 14; pp. 148-149)

So, neural dynamics on the one hand concerns the activation levels of neurons, but on the other hand it also concerns how network characteristics representing the above-mentioned criteria change over time: both dynamics within the network and dynamics of the network occur. Changes of network characteristics are created by patterns in the past affecting them. These changes might be slow in comparison to activation propagation of neurons, as in learning from multiple experiences, but might also be very fast, almost instantly, as in creation of memories: rapid resetting of the criteria (Tse, 2013). This form of network adaptation describes emerging informational content; the activation of neurons takes place based on this created informational content (Tse, 2013; Tse, 2018):

'[In addition to activation propagation] . . . neurons can also rapidly and dynamically change the weights, gains, and temporal integration properties of synapses of other neurons without necessarily triggering an action potential in those neurons (§4.54–4.60). This and other physical mechanisms accomplish a recoding of the inputs that will make a neuron fire in the future. Such recoding changes both the physical and the informational criteria that a neuron places on its inputs, even when the threshold for firing at the axon hillock remains constant.’ (Tse, 2013), p. 22

He emphasizes that patterns over time have a causal effect, not just states. The criteria play an important role in decoding these patterns:

'Patterns in input can be genuinely causal only if there are physical detectors, such as neurons, that respond to patterns in input and then change the physical system in which they reside if the criteria for the presence of a pattern have been met' (Tse, 2013), p. 9
‘Criterial decoders allow for the emergence of physically realized informational causal chains that are not just physical causal chains (although they are that), but also, in the case of the brain, mental causal chains. What makes one outcome occur rather than others is that certain informational criteria were met in decoders that then triggered consequences that would have differed had these informational criteria not been met.’ (Tse, 2013), p. 116

During evolution these mechanisms have been developed not in an arbitrary manner but because of their informational content:

‘The whole point of setting up a physical causal system to react criterially to input, as we have in the brain, is to make sure that only the subset of possible physical causal chains that are also informational chains are the ones that will be realized. Far from being an incidental fact about this subset of physical causal chains, natural selection has operated to optimize the efficiency of information processing within possible informational causal chains to fit the needs of an organism within its niche. The fact that information processing is realized in physical causal chains is, in a sense, beside the point. An animal does not get weeded out because it followed such and such a physical causal chain as opposed to another; it gets weeded out because it failed to recognize the lion where there was one. It gets weeded out or survives to the extent that it processes information poorly or well. Evolution of course operates not only to weed out suboptimal bodies, but also suboptimal perceptual, emotional, and cognitive systems realized in neuronal and endocrine system activity.’ (Tse, 2013), p. 131

2.3. How criterial causation relates to temporal factorisation

From Section 1 and the explanations above, it is clear that the notions of temporal factorisation and of criterial causation are in a close relationship; also the pictures shown in Fig. 1 (adopted from (Treur, 2007a)) and Fig. 2 (adopted from (Tse, 2013)) for these notions show similarity. The correspondence is as follows, using the above illustration for temporal factorisation by the lock and key example:

(1) In past pattern $a$, someone locked the door, and since then nobody unlocked the door
(2) This creates in the present state mediating state property $p$ corresponding to the locked door with a specific lock
(3) When in a future pattern $b$ someone attempts to unlock the door with the same key, it is checked whether or not the key actually fits to the lock
(4) When it indeed fits in this future pattern $b$, the door unlocks

This corresponds to criterial causation as follows

(1) In past pattern $a$, someone locking the door (and nobody unlocking it in the meantime) corresponds to setting the criteria for criterial causation
(2) The mediating state property $p$ corresponds to the criteria set
(3) In a future pattern $b$, attempting to fit a key of the right shape in the lock, corresponds to checking of the fulfilment of the criteria.
(4) If after a fit of key and lock in future pattern $b$ (=the criteria are indeed fulfilled), the door unlocks; this corresponds to activation of the neuron in pattern $b$.

More specifically, criterial causation as described in (Tse, 2013) covers the consequent

$$\exists p \ [a \Rightarrow p \ & \ p \Rightarrow b]$$

(14)

of the principle of temporal factorisation, without mentioning a condition $a \Rightarrow b$, similar to Laplace’s description $L(8)$. So, if criterial causation is indicated by $CC$, then the following logical relationship is obtained:

$$\text{D} \ & \ \text{TF} \Rightarrow \text{CC}$$

(15)

But note that criterial causation specifically addresses dynamics within the brain and not of the world in general as TF (1) does, and also D (7) and L (8) do. Moreover, Tse (2013) also mentions that his perspective might have to cover small random nondeterministic fluctuations as well. Note that such nondeterministic fluctuations can fit well to temporal factorisation, as these will concern cases in which the condition $a \Rightarrow b$ is not fulfilled. However, in the current paper this possible way of handling nondeterministic aspects in neural processes will not be considered in more detail.

3. Modeling causation relations by temporal-causal networks

Temporal factorisation describes world dynamics, but also a notion of causality can be used to address world dynamics. In particular, in Tse (2013)’s view concerning criterial causation, causality plays a crucial role; e.g., see the quotes in Section 2.2. Therefore it it has been analysed how both temporal factorisation and criterial causation relate to causation relations describing world dynamics.
A specific modeling approach taking causation relations into account is modeling by temporal-causal networks (Treur, 2016). The current section focuses on briefly describing this general network modeling approach. In Section 4 this modeling approach is applied to the general perspective of temporal factorisation and Sections 5 and 6 in particular to criterial causation.

3.1. Temporal-causal network models

As the notion of causality is used in many scientific disciplines, this is a notion that plays a unifying role. It is found in practically all scientific literature. Indeed, causal relations can be used everywhere, from hitting a ball causing a movement of it (physical disciplines) to how someone’s beliefs cause emotions and behaviour (mental or psychological disciplines) or how in a social context joint forces cause change in society (social disciplines), to mention just a few cases. Temporal-causal networks as described in (Treur, 2016) are a useful modeling tool for causation relations. More specifically, a temporal-causal network is characterised by:

- **connectivity characteristics**
- Connections from a node \( X \) to a node \( Y \) and their weights \( \omega_{X,Y} \)
- **aggregation characteristics**
- For any node \( Y \), some combination function \( c_Y(\cdot,\cdot) \) defines aggregation that is applied to the causal impacts \( \omega_{i,Y} X_i(t) \) on \( Y \) from its incoming connections where \( X_1,\ldots,X_k \) are the states from which \( Y \) gets its incoming connections
- **timing characteristics**
- Each node \( Y \) has a speed factor \( \eta_Y \) defining how fast it changes for given causal impact

The difference (or differential) equations that are useful for simulation purposes and also for analysis of temporal-causal networks incorporate these network characteristics \( \omega_{X,Y}, c_Y(\cdot,\cdot), \eta_Y \): it holds

\[
Y'(t+\Delta t) = Y(t) + \eta_Y [c_Y(\omega_{X_1,Y} X_1(t),\ldots,\omega_{X_k,Y} X_k(t))] - Y(t)\Delta t
\]

for any state \( Y \) and where \( X_1,\ldots,X_k \) are the states from which it gets its incoming connections. Some useful combination functions (where \( \sigma \) is a steepness parameter and \( \tau \) an excitability threshold parameter):

- the advanced logistic sum function \( \text{allogistic}_{x,t}(\cdot) \) defined by:

\[
\text{allogistic}_{x,t}(V_1,\ldots,V_k) = \frac{1}{1 + e^{-\sigma(V_1+\cdots+V_k-\tau)}} - \frac{1}{1 + e^{-\sigma\tau}} (1 + e^{-\sigma\tau})
\]

- the simple logistic sum function \( \text{slogistic}_{x,t}(\cdot) \) defined by:

\[
\text{slogistic}_{x,t}(V_1,\ldots,V_k) = \frac{1}{1 + e^{-\sigma(V_1+\cdots+V_k-\tau)}}
\]

The above concepts enable to design network models and their dynamics in a declarative manner, based on mathematically defined functions and relations; see (Treur, 2016).

3.2. Modeling adaptation of networks by self-modeling or reified network models

Realistic network models are usually adaptive: their network characteristics often change over time. Therefore, their dynamics is usually an interaction (sometimes called co-evolution) of these two sorts of dynamics: dynamics of the nodes in the network (dynamics within the network) and dynamics of the characteristics of the network (dynamics of the network). Dynamics of the network’s nodes are modeled declaratively by declarative mathematical functions and relations. In contrast, the dynamics of the network characteristics traditionally are described in a procedural, algorithmic, nondeclarative manner, which then leads to a hybrid type of model. But by using self-modeling networks or network reification, a network-oriented conceptualisation can also be applied to adaptive networks to obtain a declarative description using mathematically defined functions and relations; see (Treur, 2020a). This works through the addition of new nodes to the network (called self-model states or reification states) which represent (adaptive) network characteristics.

Table 1

| Types of characteristics | Concepts | Notations | Self-model states | Role played by the self-model state |
|--------------------------|----------|-----------|------------------|-----------------------------------|
| Connectivity characteristics | Connections weights | \( \omega_{X,Y} \) | \( W_{X,Y} \) | Connection weight \( W \) |
| Aggregation characteristics | Combination functions and their parameters | \( \eta_Y \) | \( \eta_Y \) | Speed factor \( \eta \) |

Table 1

| Different network characteristics and self-model or reification states for them. | | | | |
|---|---|---|---|
| Types of characteristics | Concepts | Notations | Role played by the self-model state |
| Connectivity characteristics | Connections weights | \( \omega_{X,Y} \) | Connection weight \( W \) |
| Aggregation characteristics | Combination functions and their parameters | \( \eta_Y \) | Speed factor \( \eta \) |

In a graphical 3D-format, such nodes are depicted at a next level (self-model or reification level), where the original network is at a base level. These characteristics with their reification states and their roles are shown in Table 1.
This construction provides an extended network, also called a self-modeling network or reified network, which has been shown to be again a temporal-causal network using an appropriate universal combination function and difference equation; see (Treur, 2020a), Ch 10. Like for all temporal-causal network models, a reified network model is specified in a (network-oriented) declarative mathematical manner based on nodes and connections. These include interlevel connections relating nodes at one level to nodes on another level.

As the outcome is also a temporal-causal network model (Treur, 2020a, Ch 10), this whole construction can easily be applied iteratively to obtain multiple reification levels. This can provide higher-order adaptive networks, and is quite useful to model, for example, plasticity and metaplasticity in the form of a second-order adaptive network with three levels, one base level and a first- and a second-order reification level; e.g., (Abraham & Bear, 1996; Treur, 2020b) or (Treur, 2020a), Ch 4.

To support the design of network models, for any application, from a library predefined basic combination functions \( bcf_i(\ldots) \), \( i = 1, \ldots, m \) are selected by assigning weights \( \gamma_{i,Y} \), where the combination function then becomes the weighted average

\[
e_Y(V_1, \ldots, V_k) = \frac{\gamma_{1,Y} bcf_1(V_1, \ldots, V_k) + \cdots + \gamma_{m,Y} bcf_m(V_1, \ldots, V_k)}{\gamma_{1,Y} + \cdots + \gamma_{m,Y}}
\]

Furthermore, parameters of combination functions are specified, so that \( bcf_i(\ldots) = bcf_i(p,v) \) where \( p \) is a sequence of parameters and \( v \) is a sequence of values.

4. Temporal factorisation modeled by temporal-causal networks

As noted in Section 2.1 (see (6)), temporal factorisation applied iteratively describes how for global causal flows of dynamics from past to future, intermediate states can be found that in a sense break up the global process in smaller causal steps. This idea suggests the use of causal networks of states that together describe the dynamics of the overall process. How this works, will be shown in more detail in the current section for two examples illustrating the temporal factorisation principle.

4.1. An example network model illustrating temporal factorisation

The first example scenario used here to relate temporal factorisation to temporal-causal network modeling is the locked door example described in the last three paragraphs of Section 2.1. It was modeled by the small temporal-causal network with connectivity as shown in Fig. 3. The states can be given the following meaning as shown in Table 2, second column. A second example about animal behaviour will be discussed in Section 4.3.

In this network, \( X_2 \) is the state (locked door, by a specific lock) that is considered in the present. There is one causal pathway from the past, coming from \( X_1 \) (someone locks the door) and a causal pathway for the future involving \( X_3 \) (someone attempts to unlock the door, with a specific key) and \( X_4 \) (unlocked door). The circular arrow for \( X_2 \) models persistence of that state (locked door), assuring that nobody was unlocking the door since it was locked.

In Box 1 the specification of this network model in role matrix format is shown. Any role matrix has rows for all the states of the network. Within each row the impacts on that state from that role are listed. For example, in role matrix \( mb \) addressing base connectivity in the row for \( X_4 \) it

---

Box 1 Role matrices specification and initial values for the network model with connectivity depicted in Fig. 3 illustrating temporal factorization.
is indicated that it gets impact from states $X_2$ and $X_3$. Connectivity role matrix $\text{mcfw}$ addressing connection weights indicates, for example, that $X_4$ gets impact from connection weights 1 from $X_2$ and $X_3$, respectively. Aggregation role matrix $\text{mcfp}$ specifying combination function weights indicates that multiple impacts are aggregated by the advanced logistic sum function $\text{algebraic}_{\sigma_d}(...) \text{ defined by (17). A second combination function is } \text{stepmod}_{\rho_d}(...), \text{ which is used to get independent events happen regularly (at time } \delta) \text{ according to a cyclic pattern with a certain repetitive duration (duration } \rho). \text{ These two combination functions are numbers } 2 \text{ and } 35 \text{ of the Combination Function Library; this selection from the library is indicated by } \text{mcfa} = [235], \text{ after which these combination functions have numbers } 1 \text{ and } 2 \text{ for this model. In aggregation role matrix } \text{mcfp} \text{ for the combination function parameters it is indicated that the parameter values for the function used by } X_4 \text{ are } 40 \text{ (for steepness } \sigma) \text{ and } 1.8 \text{ (for threshold } \tau), \text{ respectively. Finally, in the role matrix } \text{ms} \text{ for timing it is indicated that } X_4 \text{ has speed factor } 0.5, \text{ and in } \text{iv} \text{ it is indicated that all initial values are } 0. \text{ }

4.2. Simulation for the network model illustrating temporal factorisation

In Fig. 4 a simulation for this network model is shown. The considered present is some time point between 41 and 45. As can be seen, someone locks the door around time 41 (blue line), which is considered the past here. As a consequence the door is locked (red line) by a specific type of lock, and this state persists (nobody unlocks it until the present). Therefore this locked door is also there in the present. In the future, from time 47 to 59 someone brings the right key at time 47 and attempts to unlock the door (light green line), given the specific lock fitting to the key (represented by $X_3$), resulting in an unlocked door (darker green line).

4.3. Application of the network model to delayed response behaviour

Another illustration of temporal factorisation uses the notion of ‘delayed response behaviour’ of animals (in the past often considered for experiments with mice), also used in (Treur, 2007a); e.g., (Cromwell, Tremblay, Schultz, 2018; Foster, 1973; Hunter, 1912; Tinklepaugh, 1932). Consider the following concepts for an experiment in which the animal is first kept behind a window and later on released:

- c food at position $P_1$ is visible for a given animal
- d the animal is released
- e the animal is at $P_1$.

In such experiments, after having seen the food, the food is covered by a cup so that it is not visible anymore. A past pattern $a$ can be described by:

'for at least one time point in the past, state $c$ (denoting visible food at $P_1$) occurred and since then it was not visible that it did not occur anymore.'

(20)

A future pattern $b$ can be described by:

'if in future the state $d$ occurs (denoting that the animal is released),
then at a later time the state $e$ will occur (denoting that the animal is at $P_1$).'

(21)

From various experiments it has turned out that always when past trace $a$ occurs, also future trace $b$ occurs: $a \Rightarrow b$. Therefore, temporal factorisation applies to these patterns: since $a \Rightarrow b$ holds, we know that a state property $p$ exists for which the temporal relationships $a \Rightarrow p$ and $p \Rightarrow b$ hold. In this case, this postulated state property $p$ is a cognitive state, which functions as a form of memory. In other words, after observing in many animal experiments past-future relationship $a \Rightarrow b$, the principle of temporal factorisation claims that some mental (memory) state emerges in the present after past pattern $a$. This mental state in the present determines the animal’s future behaviour so that $b$ holds. So, the mediating state property $p$ can be formulated for this case as

'the animal has a mental (memory) state for the food at $P_1$.'

(22)
Such a memory state encodes information about the animal’s environment, and this information determines the future behaviour pattern. For this case the information formation is an human-independent emerging process which takes place without human intervention; even for the mouse it is an automatic process and not intentional. Therefore in this case the world itself is the actor that does information formation, through the brain’s mechanisms.

To model this example, the same network structure as shown above in Fig. 3 can be used, but this time the states are interpreted as shown in Table 3.

So, now the simulation story is as follows (see Fig. 5), while mathematically spoken the graph is essentially the same. The mouse sees food at P1 in the past, around time 20 (blue line). Due to this, it forms a mental (memory) state for the food at P1 (red line). As it was not seen that the food disappeared, this mental state persists and is also there in the present, at some time point between 41 and 45. In the future, from time 47 to 59 the mouse is released after time 47 (light green line) and due to this and the mentioned mental state, it goes to P1 (darker green line).

5. Formalising criteria for criterial causation for temporal-causal networks

In this section to model criterial causation by temporal-causal networks, as a first step the criteria themselves will be analysed in more detail. Based on difference equation (16), the activation criterion for a state $Y$ can be formulated by requiring that the aggregated impact $c_Y(\omega_{X_1,y}X_1(t), \cdots, \omega_{X_k,y}X_k(t))$ on $Y$ has a certain level at least higher than 0.5:

$$c_Y(\omega_{X_1,y}X_1(t), \cdots, \omega_{X_k,y}X_k(t)) > 0.5$$  (23)

Inequality (23) is used here as the general criterion in a temporal-causal network for criterial causation for any arbitrary combination function $c_Y(\cdot)$ for any state $Y$. Below, this general criterion will be instantiated for specific combination functions.

5.1. Criteria using logistic combination functions

Often used examples of combination functions $c_Y(\cdot)$ are the simple logistic $\text{slogistic}_{\sigma,\tau}(\cdots)$ and advanced logistic function $\text{alogistic}_{\sigma,\tau}(\cdots)$, each with steepness parameter $\sigma > 0$ and excitability threshold parameter $\tau$; see (17), (18). For each of the combination functions criterion (23) can be made more specific: for example, for $\text{slogistic}_{\sigma,\tau}(\cdots)$ (18), criterion (23) can be reformulated into the more specific criterion

$$\frac{1}{1 + e^{-\sigma(V_1 + \cdots + V_k - \tau)}} > 0.5$$  (24)

with the $V_i$, the single impacts $\omega_{X_i,y}X_i(t)$ on state $Y$. By algebraic rewriting (shown in Box 2, left), this translates into the following linear inequality in $X_1(t), \cdots, X_k(t)$:

$$\text{slogistic}_{\sigma,\tau}(\cdots)$$

So, (25) is the criterion in particular for $\text{slogistic}_{\sigma,\tau}(\cdots)$. Similarly, for $\text{alogistic}_{\sigma,\tau}(\cdots)$, see (17), the following criterion is found (see Box 2, right), which still is a linear inequality in $X_1(t), \cdots, X_k(t)$:

$$\text{alogistic}_{\sigma,\tau}(\cdots)$$

| Table 3 Explanation of the states for the mouse example. |
|-----------------|------------------|
| State           | Explanation                  |
| $X_1$           | Mouse sees food            |
| $X_2$           | Mental state for food       |
| $X_3$           | Mouse released             |
| $X_4$           | Mouse goes to $P_1$         |

| Nr  | Name                        | Explanation                                      |
|-----|-----------------------------|-------------------------------------------------|
| 1   | Mouse sees food             | Mouse sees food at location $P_1$                |
| 2   | Mental state for food       | Mouse has mental state for food at $P_1$         |
| 3   | Mouse released              | Mouse is released and free to go                 |
| 4   | Mouse goes to $P_1$         | Mouse goes to $P_1$                             |
\[ c_Y(\omega_{X_1, Y}X_1(t), \ldots, \omega_{X_k, Y}X_k(t)) > c \]  

(26)

with \( c = \tau - \log\left(\frac{1}{1 + e^{-\sigma r}}\right) - 1)/\sigma \)

It can be seen that for these two specific combination functions, criteria (25) and (26) are linear inequalities concerning state values \( X_1, \ldots, X_k \) with as constants expressions in terms of network characteristics \( \omega, \sigma, \tau \).

### 5.2. Criteria using other combination functions

In a similar way, the criteria for other combination functions were found: scaled minimum \( \text{sm}_{\text{min}}(\cdot) \) and scaled maximum \( \text{sm}_{\text{max}}(\cdot) \), and scaled sum \( \text{ss}_{\text{sum}}(\cdot) \), Euclidean \( \text{eucl}_{\text{a}}(\cdot) \) and scaled geometric mean \( \text{sg}_{\text{geomean}}(\cdot) \); see Table 4. Note that all of the criteria, except the last two are linear inequalities in \( X_1(t), \ldots, X_k(t) \), whereas the last two are inequalities involving powers and products of \( X_1(t), \ldots, X_k(t) \).

As real-world network models are often adaptive, the network characteristics mentioned can and usually do change over time. In this manner, the criteria and the informational content represented by the criteria emerge dynamically. For example, in criteria (23), (25), (26) the connection weights \( \omega \), the excitability threshold \( \tau \) and the steepness \( \sigma \) may change over time, by which they change the criterion. This can be modeled well in the form of an adaptive network. Using for this a reified temporal-causal network, such adaptive causality needed for criterial causation can easily be modeled. In that setting, for temporal factorisation a mediating state property \( p \) is specified by the values of some states, which can be base level states as in Section 4, or states at any reification level, as will be illustrated in Sections 6 and 7. Also the past patterns \( a \) and the future patterns \( b \) can be specified by state value assignments for a number of states, in this case over time. In relation to criteria (25) and (26), the connection weight coefficients and threshold constants are represented by reification states, whereas the states \( X_1, \ldots, X_k \) in the criteria are base level states.

### 6. An example adaptive network for criterial causation

As explained in Section 3, a temporal-causal network is defined by three main network structure characteristics (connectivity, aggregation, timing), which are modeled by \( \omega_{X, Y}, c_Y(\cdot), \eta_Y \). Difference equations as expressed in (16) above include these network characteristics. For an adaptive network some of these characteristics are explicitly represented by reification states \( \mathbf{W}_{X, Y}, \mathbf{C}_Y, \mathbf{P}_Y \) or \( \mathbf{H}_Y \), so that they can become dynamic. This is applied in the example presented here.

### Table 4

Overview of criteria for criterial causation for a few combination functions.

| Combination function          | Formula | Criterion for criterial causation                                                                 |
|------------------------------|---------|---------------------------------------------------------------------------------------------------|
| \( c_Y(V_1, \ldots, V_k) \)  | \( 1 + e^{-\omega(V_1, \ldots, V_k)} \) | \( c_Y(\omega_{X_1, Y}X_1(t), \ldots, \omega_{X_k, Y}X_k(t)) > 0.5 \)                        |
| \( \text{logistic}_{\text{a}}(V_1, \ldots, V_k) \) | \[ \frac{1}{1 + e^{-\omega(V_1, \ldots, V_k)}} \left(1 + e^{-\sigma r}\right) \] | \( \omega_{X_1, Y}X_1(t) + \cdots + \omega_{X_k, Y}X_k(t) > \tau \)                        |
| \( \text{max}_{\text{a}}(V_1, \ldots, V_k) \) | \( \max(V_1, \ldots, V_k) \) | \( \omega_{X_1, Y}X_1(t) > 0.5 \) for some \( i \)                                            |
| \( \text{min}_{\text{a}}(V_1, \ldots, V_k) \) | \( \min(V_1, \ldots, V_k) \) | \( \omega_{X_1, Y}X_1(t) > 0.5 \) for all \( i \)                                              |
| \( \text{ss}_{\text{sum}}(V_1, \ldots, V_k) \) | \( \sqrt{\sum_{i=1}^{k} V_i} \) | \( \omega_{X_1, Y}X_1(t) + \cdots + \omega_{X_k, Y}X_k(t) > 0.5 \)                        |
| \( \text{eucl}_{\text{a}}(V_1, \ldots, V_k) \) | \( \sqrt{\sum_{i=1}^{k} V_i^2} \) | \( \omega_{X_1, Y}X_1(t) + \cdots + \omega_{X_k, Y}X_k(t) > 0.5 \)                        |
| \( \text{sg}_{\text{geomean}}(V_1, \ldots, V_k) \) | \( \left(\prod_{i=1}^{k} V_i\right)^{1/k} \) | \( \omega_{X_1, Y}X_1(t) + \cdots + \omega_{X_k, Y}X_k(t) > 0.5 \)                        |
6.1. An adaptive example scenario

The scenario used for this section concerns a learning process of a new person in a company who has to learn to recognize colleagues. It goes as follows.

Example Scenario

A new person in a company has to learn to recognize a colleague from only seeing his face; this face is stimulus \( s \). Two colleagues \( a_1 \) and \( a_2 \) are assumed that are options to choose from. Picking one of them is indicated by activation of preparation state \( p_{sa_1} \). A belief \( b_{s1} \) suggests that it is colleague \( a_1 \), and a belief \( b_{s2} \) that it is colleague \( a_2 \). These beliefs are only meant indicative (e.g., based on the location at which the person is encountered), but not sufficient to decide for one of them. As the beliefs and \( s \) are triggered by independent circumstantial factors, for the network model they just happen. Two types of network characteristics are addressed as adaptive: the weights of the connections from sensory representation \( srs \), for \( s \) to \( p_{sa_1} \) and \( p_{sa_2} \), and the excitability thresholds for states \( p_{sa_1} \) and \( p_{sa_2} \). During the scenario these characteristics are learnt so that over time a better decision results. The network characteristics define the criteria for criterial causation of recognizing a face. Given these criteria, in future situations an encounter with \( s \) (also at unexpected locations, such as in a shop or in another town) leads to satisfying the criteria and as a consequence to correct recognition.

Note that for this example scenario the relevant past pattern \( a \) can be described by

\[
\text{in the past a number of times the face is seen (} srs \text{ occurs) co– occurring with the correct belief about which person it was (} b_{s1} \text{)}
\]

and the relevant future pattern \( b \) by

\[
\text{in the future, if the face is seen (} srs \text{ occurs), then the correct recognition takes place (} p_{sa} \text{)}
\]

Moreover, the intermediate state property \( p \) is defined by

\[
\text{the adaptive network characteristics have specific appropriate values} \tag{29}
\]

What such appropriate values are, for example, can be seen in the example simulation in Section 7; see (31).

6.2. Connectivity and aggregation for the adaptive network model

In this section an adaptive network model is introduced that addresses the above example scenario. In a graphical representation of the network’s connectivity in a 3D format, the self-model or reification states are placed in a second (blue) plane, above the (pink) plane for the base network. See Fig. 6 and see Table 5 for explanations of the states. The following connection types are considered: upward and downward connections, and horizontal, leveled connections. Downward connections have a particular effect, as they are effectuating one of the types of adaptive characteristics indicated by their role \( W, C, P \) or \( H \); see also Table 1 in Section 3.

For aggregation, in the example network modeling criterial causation, for most states the logistic function \( \text{allogistic}_a(\cdot \cdot \cdot) \) is used; see (17) above. The exception is that to model Hebbian learning (Hebb, 1949), combination function \( \text{hebb}_p(V_1, V_2, W) \) is applied for reification states \( W_{X,Y} \):

\[
\text{hebb}_p(V_1, V_2, W) = V_1 V_2 (1 - W) + \mu W \tag{30}
\]

where \( V_1, V_2 \) are variables for the single impacts on the reification state \( W_{X,Y} \) (in Fig. 5 in the upper plane) gets from the base states \( X \) and \( Y \) in the addressed adaptive connection (in Fig. 6 in the lower plane) and \( W \) for the connection weight representation \( W_{X,Y} \); moreover, \( \mu \) is a persistence parameter. The following reification states are used in Fig. 6:

- \( W_{X,Y} \) plays the role of connection weight for the adaptive connection from \( X \) to \( Y \) (Hebb, 1949)
- \( T_Y \) plays the combination function parameter role for state \( Y \)'s adaptive excitability threshold \( \tau \) (Chandra & Barkai, 2018)

Fig. 6. 3D representation of the connectivity of the network model to illustrate criterial causation, including: (1) base level for the face recognition (depicted by the lower, pink plane), (2) self-model or reification level for the criteria (depicted by the upper, blue plane) based on the weights \( p_s \) (reified by the two \( W \) states) of the base connections from \( srs \), to \( p_{sa_1} \) and \( p_{sa_2} \) and the excitability thresholds \( \tau \) (reified by the \( T \) states) of these two base states \( p_{sa_1} \) and \( p_{sa_2} \)
6.3. Specification of the adaptive network model by role matrices

The network model was specified by role matrices $mb$ (for the base connection role), $mcw$ (for the connection weight role), $ms$ (for the speed factor role), $mcfw$ (for the combination function weight role), and $mcfp$ (for the combination function parameter role); e.g., (Treur, 2020a, 2020b). In all of these role matrices (see Box 3) the rows are for the different states; at each row for the state $X_j$ indicated in the first column, it is specified in the other columns which other states (via the incoming arrows shown in Fig. 6) or network characteristics have impact on $X_j$ for that role.

These impacts on $X_j$ are distinguished by their role, i.e.: base or non-base, where for the non-base case a distinction is made for respectively: connection weight role, speed factor role, combination function weight role and combination function parameter reification role. For a designed model a list of combination functions used is specified by $mcf = [\ldots]$, for the current example it is $mcf = [2 3 35]$; here the numbers 2, 3, 35 refer to the numbers in the combination function library, where $\text{alogistic}_{\mu}(\cdot)$ has number 2 and $\text{hebb}(\cdot \cdot)$ number 3; number 35 is the stepmod function used to create independent events. Box 3 shows all role matrices for the adaptive network model addressing criterial causation.

Role matrix $mb$ specifying base connectivity indicates at each row for the indicated state $X_j$ from which states it gets
incoming connections from the same or a lower level. For example, the 5th row indicates for state \( X_5 (= \text{ps}_{a1}) \) two incoming base connections, one from state \( X_2 (= \text{sr}_{s}) \), and one from state \( X_3 (= \text{bs}_{1}) \). For another example, row 7 indicates that state \( X_7 (= \text{Wsr}_{s}\text{ps}_{a1}) \) has incoming base connections from \( X_2 (= \text{sr}_{s}) \), \( X_5 (= \text{ps}_{a1}) \) and from \( X_7 \) itself in that order; this ordering is crucial since the Hebbian combination function \( \text{hebb} \) used for this state \( X_7 (= \text{Wsr}_{s}\text{ps}_{a1}) \) is not symmetric in its three arguments, as can be seen in (30).

The four role matrices specifying non-base connectivity describe in each row impacts on the indicated state \( X_j \) by reification states from a higher level (see the downward arrows in Fig. 6); they are as follows: role matrices \( \text{mcw} \) for the connection weight role and \( \text{ms} \) for the speed factor role, and role matrices \( \text{mcfw} \) for the combination function weight role and \( \text{mcfp} \) for the combination function parameter role (see Box 3). Within each of these non-base role matrices cell entries in red cells show the name of a state (at a higher level) that in a reified form represents in an adaptive characteristic; in contrast, entries in green cells indicate static values for nonadaptive characteristics. Therefore, as seen in Box 3 the red cells in \( \text{mcw} \) and \( \text{mcfp} \) indicate the (reification) states \( X_7 \) to \( X_{10} \). For example, in role matrix \( \text{mcw} \) the indication \( X_7 \) in the red cell at row 5 and column 1 specifies that the value of state \( X_7 \) represents the connection weight from \( \text{sr}_{s} \) to \( \text{ps}_{a1} \) (as indicated in \( \text{mb} \)). Unlike this, the 1 in green cell at row 7, column 1 of \( \text{mcw} \) shows the nonadaptive value of weight of the connection from \( X_2 (= \text{sr}_{s}) \) to \( X_7 (= \text{Wsr}_{s}\text{ps}_{a1}) \). In role matrix \( \text{mcfp} \) specifying the combination function parameter role, in the red cell at row 5 and column 2 it is specified that the actual value for the excitability threshold of \( \text{ps}_{a1} \) is represented by the value of reification state \( X_9 (= \text{Tps}_{a1}) \). More explanation of this specification format and how it is used to automatically generate simulations can be found in (Treur, 2020a; Treur, 2020b).

7. Example simulation of criterial causation

This section describes one of the example simulation scenarios that were addressed using the modeling environment developed (Treur, 2020a; Treur, 2020b). In particular, Figs. 7 and 8 show a simulation for the Example Scenario described in Section 6. The settings from Box 3 were used. In accordance with temporal factorisation, past pattern \( a \)
and future pattern \( b \), and the mediating state property \( p \), will be discussed subsequently.

7.1. Past pattern \( a \) (time point 0 to 100)

In the past trace according to pattern \( a \), the stimulus \( s \) (indicating the observed face) is active from time point 25 to time point 50 and from time 75 to time 100. In these time periods, the belief state \( bs_2 \) co-occurs. In Fig. 7, the upper and lower graph show this past trace and how recognition of the stimulus \( s \) as Person 2 improves over time: during the first encounter, preparation state \( ps_{a2} \) (the red line) only increases slowly, but during the second encounter this happens much faster; apparently at that time a better criterion was set for proper recognition already.

The lower graph shows the emerging network characteristics representing this criterion. It can be seen that (under influence of belief state \( bs_2 \)) the reified adaptive connection weight from \( srs \) to \( ps_{a2} \) represented by \( W_{srs,ps_{a2}} \) (purple line) gets stronger, whereas the excitability threshold represented by \( T_{ps_{a2}} \) of \( ps_{a2} \) (pink line) gets lower. Between time points 60 and 70 as an intended irregularity of the history, for a very short period, belief state \( bs_1 \) shows, but this has no serious consequences for the adaptation process.

7.2. Criterion formed at time point 100

Mediating state property \( p \) in general describes the general criterion (15) for criterial causation. In the addressed scenario, the considered mediating state property \( p \) describes the relevant network characteristics: the reified weights \( W_{srs,ps_{a1}} \) and \( W_{srs,ps_{a2}} \) of the connections from \( srs \) to \( ps_{a1} \) and \( ps_{a2} \), and the reification states for the excitability thresholds \( T_{ps_{a1}} \) and \( T_{ps_{a2}} \) for \( ps_{a1} \) and \( ps_{a2} \); so \( p \) is based on the values of reification states \( X_7 \) to \( X_{10} \). Note that all belong to the reification level.

Time point 100 shows the following assignments of values: \( X_7 = 0.136275 \), \( X_8 = 0.93172 \), \( X_9 = 0.78281 \), \( X_{10} = 0.17232 \). It is this configuration (at the reification level) that defines the mediating state property \( p \):

\[
p = X_7 = 0.136275 \text{ and } X_8 = 0.93172 \text{ and } X_9 = 0.78281 \text{ and } X_{10} = 0.17232
\]

(31)

In this way \( p \) describes the constants \( \omega_{X_i,Y} \) and \( \tau \) in general criterion (26)

\[
\omega_{X_i,Y}X_1(t) + \cdots + \omega_{X_i,Y}X_4(t) > c
\]

with \( c = \tau - \log \left( \frac{1}{0.5 e^{-c}} + \frac{1}{e^{c}} - 1 \right) / \sigma \)

for the advanced logistic function for criterial causation. With the above values substituted in it, it works as follows.

At time 100, for preparation state \( ps_{a2} \), the representation \( W_{srs,ps_{a2}} \) for the weight of the concerning connection from stimulus representation \( srs \) to preparation \( ps_{a2} \) has a high value (0.93171758), whereas the excitability threshold \( T_{ps_{a2}} \) for \( ps_{a2} \) a low value (0.17232). Therefore, the criterion for Person 2 can be fulfilled easily just by the impact only coming from \( srs \). Indeed, using (26) after substitution by the values at time point 100 of the relevant \( W^- \) and \( T^- \) states (which reify \( \omega \) and \( \tau \)), and substitution of the static value 5 for \( \sigma \), and 0.5 for the static connection weight \( \omega \) from the belief state, the criterion for activation translates into

\[
0.5 \times bs_2(t) + 0.93172 srs_1(t) > 0.0857
\]

No positive value of \( bs_2(t) \) is needed here, this is fulfilled if \( srs_1(t) > 0.0857/0.93172 = 0.092 \)

This implies that already a low-level sensory face representation signal \( srs_1 \) (as low as 0.1) is enough for the face recognition.

In contrast, for option \( ps_{a1} \), representation \( W_{srs,ps_{a1}} \) for the weight of the considered connection from stimulus
representation $sr_{r}$ to preparation $ps_{al}$ has a low value $(0.136275)$, whereas the excitability threshold $rps_{al}$ for $ps_{al}$ has a high value $(0.78281)$. This makes that for Person 1 the criterion for activation of $ps_{al}$ cannot be fulfilled.

7.3. Future pattern $b$ (time point 100 to 200)

The graph shown in Fig. 8 depicts the future trace (from time point 100 on) according to pattern $b$; it can be seen how now recognition takes place without any activation of the belief state. This is due to the criterion described above in terms of the mediating state property at time 100; the criterion is described by the value assignments for the reification states for connection weights and for the excitability thresholds shown in (31) in Section 7.2 above. In the future trace at times 125 and 175 the face is encountered again ($ss_{r}$), whereby the corresponding belief state $bs_{2}$ is kept 0 (so no additional circumstantial info available), and the criterion is satisfied so that adequate recognition $ps_{s2}$ as Person 2 occurs.

8. Defining informational content by temporal relational specification

This section addresses the question how exactly the mediating state for temporal factorisation and, equivalently for the case of the brain, the state of the criteria for criterial causation, can be interpreted as having informational meaning. In other words: how can the informational content of these states be defined? Here something can be learned from Philosophy of Mind.

8.1. Relational specification of mental content

From Philosophy of Mind, well-known philosopher Kim (1996, pp. 200-202), has put forward the notion of relational specification, in the context of defining mental content of (mental or physical) state properties:

‘The third possibility is to consider beliefs to be wholly internal to the subjects who have them but consider their contents as giving relational specifications of the beliefs. On this view, beliefs may be neural states or other types of physical states of organisms and systems to which they are attributed. Contents, then, are viewed as ways of specifying these inner states; wide contents, then, are specifications in terms of, or under the constraints of, factors and conditions external to the subject, both physical and social, both current and historical.’ (Kim, 1996), pp. 200–201; italics in the original.

For the temporal perspective, in particular, a temporal relational specification is the specification of temporal relationships of the considered state to certain patterns in past and/or future. Kim puts forward that relational specifications are crucial to be able to express laws and explanations.

‘Consider physical magnitudes such as mass and length, which are standardly considered to be paradigm examples of intrinsic properties of material objects. But how do we specify, represent, or measure the mass or length of an object? The answer: relationally. To say that this rod has a mass of 5 kg is to say that it bears a certain relationship to the International Prototype Kilogram (it would balance, on an equal-arm balance, five objects each of which balances the Standard Kilogram). Likewise, to say that the rod has a length of 2 m is to say that it is twice the length of the Standard Meter (or twice the distance travelled by light in a vacuum in a certain specified fraction of a second). These properties are intrinsic, but their specifiations or representations are extrinsic and relational, involving relationships to other things and properties in the world. It may well be that the availability of such extrinsic representations are essential to the utility of these properties in the formulation of scientific laws and explanations.’ (Kim, 1996, p. 201); italics in the original.

In Kim’s approach, a considered (mental) state is distinct from the relationships it has to external other items (Kim, 1996, pp. 200-202). Kim explains as follows that a state property itself is intrinsic, whereas its relational specification indicates how it is linked to items in the world:

‘The approach we have just sketched has much to recommend itself over the other two. It locates beliefs and other intentional states squarely within the subjects; they are internal states of the persons holding them, not something that somehow extrudes from them. This is a more elegant metaphysical picture than its alternatives. What is “wide” about these states is their specifications or descriptions, not the states themselves.’ (Kim, 1996, pp. 201-202).

8.2. Applying temporal relational specification to informational content

The above summarized notion of relational specification for mental states, provides a way to specify the informational content of any mediating state property $p$ postulated by temporal factorisation. For the future direction, the occurrence of $p$ at some time $t$ makes $b$ occur in the future after $t$. A temporal relational specification may express what the causal effect of $p$ on the future is, namely that pattern $b$ will occur, also indicated by $p \Rightarrow b$. In a similar manner a temporal relational specification with respect to the past can be described by $a \Rightarrow p$. Then in fact the basic part of the consequent of the principle of temporal factorisation

$$a \Rightarrow p \land p \Rightarrow b$$

(32)

can be identified as a temporal relational specification of the informational content of state $p$.

For the scenario described by the example network model with connectivity depicted in Fig. 1, the following two temporal statements are used to describe patterns $a$ and $b$ for this example which more informally described by (11) and (12) for the locked door example, and by (20) and (21) for the delayed response animal behaviour example:
• Temporal statement expressing pattern $a$

\[\text{for some point } t_1 < t \text{ it holds } X_1(t_1) > 0.5\]
and for all time points $t_2$
with $t_1 < t_2 < t$  
\[\text{it does not hold } X_1(t_1) \leq 0.5\]  
\[t, \text{such that it holds } X_4(t_4) > 0.5\]  

• Temporal statement expressing pattern $b$

\[\text{if for some point } t_3 > t \text{ it holds } X_3(t_3) > 0.5\]
then there is some time point $t_4 > t_3$
such that it holds $X_4(t_4) > 0.5$

Note that here, to avoid more complex temporal expressions, only an approximation is discussed. To get a more precise treatment, certain minimal durations have to be specified, since a state that has a value $>0.5$ for just one single time point has not much effect. This means that a specific trace over time (which is a sequence of state values for all states for all time points, like in a simulation outcome) satisfies the pattern $a$ and only if the corresponding temporal statement is satisfied by (or valid for) this trace. Note that these two temporal expressions (33) and (34) are indeed valid for the simulation trace shown in Fig. 4 when $t$ is set between 42 and 46. Given the above pattern expressions, the informational content of intermediate state property $X_2$ can be illustrated in a simple version as follows.

• Informational content of $X_2$ for the past:

\[\text{if for some point } t_1 < t \text{ it holds } X_1(t_1) > 0.5\]
and for all time points $t_2$
with $t_1 < t_2 < t$  
\[\text{it does not hold } X_1(t_1) \leq 0.5\]  
then $X_2(t) > 0.5$

• Informational content of $X_2$ for the future:

\[\text{if } X_2(t) > 0.5 \text{ then}\]
\[\text{if for some point } t_3 > t \text{ it holds } X_3(t_3) > 0.5\]  
then there is some time point $t_4 > t_3$
such that it holds $X_4(t_4) > 0.5$

Note that these temporal statements (35) and (36) are satisfied (are valid) for the simulation trace shown in Fig. 4 when $t$ is set between 42 and 46. This notion of informational content by temporal relational specification can loosely be formulated as: the informational content of $p$ is the occurrence of pattern $a$ in the past and of pattern $b$ in the future. Note that here it might be more satisfactory if it is more explicitly formulated that the occurrence of past pattern $a$ is also implied by $p$, and the occurrence of $p$ is implied by future pattern $b$, which actually is a kind of converse of (32). Then (32) can be sharpened to

\[a \iff p \iff b\]  

This can be formulated as: the informational content of $p$ is on the one hand the occurrence of pattern $a$ in the past and on the other hand of pattern $b$ in the future. Then for the above example it becomes:

• Informational content of $X_2$ for the past:

\[\text{if for some point } t_1 < t \text{ it holds } X_1(t_1) > 0.5\]
and for all time points $t_2$
with $t_1 < t_2 < t$  
\[\text{it does not hold } X_1(t_1) \leq 0.5\]  
\[\text{if and only if } X_2(t) > 0.5\]

• Informational content of $X_2$ for the future:

\[X_2(t) > 0.5 \text{ if and only if}\]
\[\text{if for some point } t_3 > t \text{ it holds } X_3(t_3) > 0.5\]  
then there is some time point $t_4 > t_3$
such that it holds $X_4(t_4) > 0.5$

Note that again these temporal statements (38) and (39) are satisfied (are valid) for the simulation trace shown in Fig. 4 when $t$ is set between 42 and 46. The latter if-and-only-if approach expressed in (38) and (39) might fit better to an intuition of representational content. In fact, the difference with (35) and (36) lies in an additional temporal completion assumption. Such an assumption indicates that the description of pattern $a$ covers all possibilities in the past to get $p$, and similarly, state description $p$ covers all possibilities for $b$ to occur in the future. Adopting this additional temporal completion assumption, the stronger (38) and (39) can be taken as describing the informational content of $p$. Note that under this temporal completion assumption, also the temporal factorisation principle can get a different form, where in addition to (37), also by temporal completion $a \Rightarrow b$ is replaced by $a \iff b$:

\[[a \iff b] \Rightarrow \exists p \ [a \iff p \& p \iff b]\]  

There are pros and cons for this temporal completion assumption. An interesting pro is that more elegant equivalences are obtained for the informational content of $p$. A con may be that it might lead to disjunctive expressions for the past and future pattern specifications and for $p$ itself to cover all alternative possibilities; this may complicate things and may not always be applicable well. Note that (28) can also be reformulated only in terms of equivalences as
expresses that in trace $t$ relativised by the indicated trace for temporal factorisation is obtained.

ticates the state properties used. RTTPL is shown. Here the predicate Stateproperty indicates the state properties used for time points $t$, state properties $p$, and traces $tr$. For somewhat similar or comparable uses of Predicate Logic see (Bosse, Jonker, van der Meij, Sharpanskykh, & Treur, 2009; Bosse & Treur, 2011; Sharpanskykh & Treur, 2010; Treur, 2009) or (Treur, 2016), Ch. 13. Here it is discussed for Temporal Factorisation, but as Criterial Causation is a special case of Temporal Factorisation it applies to Criterial Causation as well.

9. Formalisation of temporal factorisation and criterial causation in a reified temporal predicate logic language

In this section it is shown how the notions discussed in the paper can be formalised in terms of the formal language of a form of Predicate Logic called Reified Temporal Trace Predicate Logic RTTPL. This is a use of Predicate Logic for which numbers can be used and for which constants, terms and variables can be used for time points $t$, state properties $p$, and traces $tr$. A past trace property is a property $P(t, tr)$ for time point $t$ and trace $tr$ for which every quantor for a variable $t'$ over time is relativised by $t' < t$.

A future trace property is a property $F(t, tr)$ for time point $t$ and trace $tr$ for which every variable $t' > t$ over time is relativised by $t' > t$.

A current trace property $C(p, t, tr)$ about the present expresses that in trace $tr$ at state property $p$ holds.

These past trace properties and future trace properties are used to describe past patterns and future patterns for the indicated trace $tr$. Then the following formalisation for temporal factorisation is obtained.

In Box 4 the formalisation of Temporal Factorisation in RTTPL is shown. Here the predicate Stateproperty indicates the state properties used.

9.1. Formalisation of temporal factorisation in reified temporal trace predicate logic

First the notions of past pattern and future pattern have to be formalised:

**Past and future trace properties and current properties**

A past trace property is a property $P(t, tr)$ for time point $t$ and trace $tr$ for which every quantor for a variable $t'$ over time is relativised by $t' < t$.

A future trace property is a property $F(t, tr)$ for time point $t$ and trace $tr$ for which every variable $t'$ over time is relativised by $t' > t$.

A current trace property $C(p, t, tr)$ about the present expresses that in trace $tr$ at state property $p$ holds.

These past trace properties and future trace properties are used to describe past patterns and future patterns for the indicated trace $tr$. Then the following formalisation for temporal factorisation is obtained.

In Box 4 the formalisation of Temporal Factorisation in RTTPL is shown. Here the predicate Stateproperty indicates the state properties used.

An illustration for the lock and mouse example of Section 4 is shown in Box 5; see also (21) and (22) in Section 8. Note that if $X$ is a state, then $X(t, tr)$ denotes the value of $X$ at time $t$ in trace $tr$.

**Box 5** Temporal Factorisation for the example of Section 4 formalised in RTTPL.

$$P(t, tr) \equiv \exists r1 \leq tX_1(t1, tr) > 0.5 \& \forall r2 > t$$

$$1 | r2 < t \Rightarrow \text{not} X_1(t1, tr) \leq 0.5 |$$

$$F(t, tr) \equiv \forall r3 > tX_3(r3, tr) > 0.5 \Rightarrow$$

$$\exists r4 > tX_4(\text{r4}, tr) > 0.5 \& C(p, t, tr)$$

$$\equiv X_2(t, tr) > 0.5$$

Temporal factorisation for this case:

$$\forall t, tr[\exists r1 < tX_1(t1, tr) > 0.5 \& \forall r2 > t$$

$$1 | r2 < t \Rightarrow \text{not} X_1(t1, tr) \leq 0.5 |] \Rightarrow$$

$$\forall r3 > tX_3(r3, tr) > 0.5 \Rightarrow$$

$$\exists r4 > tX_4(\text{r4}, tr) > 0.5] \Rightarrow$$

$$\forall t, tr[\exists r1 <$$

$$tX_1(t1, tr) > 0.5 \& \forall r2 > t1 | r2 < t \Rightarrow$$

$$\text{not} X_1(t1, tr) \leq 0.5 |) \Rightarrow$$

$$X_3(t, tr) > 0.5 \&$$

$$[X_2(t, tr) > 0.5 \Rightarrow \forall r3 > tX_3(r3, tr) >$$

$$0.5 \Rightarrow \exists r4 > tX_4(\text{r4}, tr) > 0.5]$$

**Box 6** Informational content formalised in RTTPL by implications and illustrated for the example of Section 4.

$$\forall t, tr[P(t, tr) \Rightarrow C(p, t, tr)] \forall t, tr[C(p, t, tr) \Rightarrow F(t, tr)]$$

For the example of Section 4:

$$\forall t, tr[\exists r1 < tX_1(t1, tr) > 0.5 \& \forall r2 > t1 | r2 < t \Rightarrow$$

$$\text{not} X_1(t1, tr) \leq 0.5 |] \Rightarrow$$

$$X_2(t, tr) > 0.5]$$

$$\forall t, tr[X_2(t, tr) > 0.5 \Rightarrow$$

$$\forall r3 > tX_3(r3, tr) > 0.5 \Rightarrow$$

$$\exists r4 > tX_4(\text{r4}, tr) > 0.5]$$

The statements (38) and (39) from Section 7 defining Informational Content by bi-implications can be formalised as shown in Box 7.
9.2. Formalisation of temporal factorisation in Second-Order reified temporal trace predicate logic

It is in principle also possible to use a reified predicate logic $R^2$TTP in which the temporal properties expressed as statements of RTTPL are reified by names for which constants, terms and variables can be used. For a similar approach, see MetaTTL in (Bosse & Treur, 2011). Then temporal factorisation can be formalised as shown in Box 8.

| Box 8 Temporal Factorisation formalised in $R^2$TTL. |
|----------------------------------------------------|
| $\forall P,F[[\text{Pasttraceproperty}(P,t) \& \text{Futuretraceproperty}(F,t) \&$ |
| $\forall t,tr[[\text{Holdsfor}(P,t,tr) \Rightarrow \text{Holdsfor}(F,t,tr)]]] \Rightarrow$ |
| $\exists p[[\text{Stateproperty}(p) \& \forall t,$ |
| $tr[[\text{Holdsfor}(P,t,tr) \Rightarrow \text{Holdsfor}(p,t,tr)]] \&$ |
| $[\text{Holdsfor}(p,t,tr) \Rightarrow \text{Holdsfor}(F,t,tr)]]]]$ |

Here predicates Pasttraceproperty and Futuretraceproperty are used to indicate these properties, and Holdsfor($X$, $t$, $tr$) expresses that temporal property $X$ is satisfied for trace $tr$ with respect to time point $t$, and Holdsfor($Y$, $t$, $tr$) that state property $Y$ is valid for time $t$ in trace $tr$. Although in principle this works, it is technically much more complicated compared to RTTPL, as the coding of temporal properties requires some technical notations.

10. Discussion

In (Tse, 2013) the notion of criterial causation was introduced. This notion describes how in the brain by plasticity specific configurations can emerge that represent informational criteria for future processing. The current paper addressed how this notion relates to temporal factorisation for world dynamics, which was introduced in (Treur, 2007a). The core of each of the two notions is that according to some adaptive process, emerging (past) brain patterns or world patterns occur, which as patterns lead to brain configurations or world configurations in the present; these in turn drive the (future) brain pattern or world pattern. Such a configuration, in one state encodes information about the past which is relevant for the future. Future processes are driven by this information (for a person, or for the world). It also was shown how temporal-causal network models may be used to model these processes.

In conclusion and making it more specific, by addressing all what was covered, the following main points have been achieved:

- Explanation of how informational content in the world emerges by the world’s own dynamics and causation (Sections 2, 4, 6)
- Explanation of the notions temporal factorisation (Treur, 2007a) and criterial causation (Tse, 2013) and how they relate to each other and to the emergence of informational content (Section 2)
- Illustration of how reified adaptive causal network modeling can be related to temporal factorisation, criterial causation and emerging informational content (Sections 3, 4, 6, and 7)
- Formalising criteria for causal causation for reified adaptive temporal-causal network models (Section 5)
- Explanation of how informational content in the world can be described by Kim’s (1996)’s notion of relational specication originally meant for mental content (Section 8)
- Formalisation of temporal factorisation and criterial causation by a first- and second-order reified temporal logical language (Section 9)

For future work, it can be explored in how far the notion of Extended Mind (e.g., Bosse, Jonker, Schut, & Treur, 2005; Bosse, Jonker, Schut, & Treur, 2006; Clark & Chalmers, 1998; Tollefsen, 2006) can be covered in a similar manner, where both the brain and the world play their role. This goes beyond the notion of criterial causation (which was only described for brain states), but still fits in the more general temporal factorisation principle. Moreover, it can be explored how also metaplasticity (Abraham & Bear, 1996; Treur, 2020a) can be addressed in adaptive processes concerning criterial causation, that will become second-order adaptive then.

In the current paper, links to the notions free will or mental causation (as discussed by Tse) from Philosophy of Mind have been left aside; this was on purpose. The notion of criterial causation of Tse (2013) has much value by itself, independent of such philosophical links, as is also emphasized by Levy (2013). This value has been the focus here.

A preliminary presentation of part of the work described here can be found in (Treur, 2019b). The current paper, meant for a journal publication, has improved and extended this by $\geq100\%$ more content.

Availability of data and materials

The software used for the simulations is available at.
The full specifications for the example network models are in the paper in Box 1 and Box 3.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

None.

Funding

None.

Authors' contributions

All.

References

Abraham, W. C., & Bear, M. F. (1996). Metaplasticity: The plasticity of synaptic plasticity. Trends in Neuroscience, 19(4), 126–130.

Ashby, W. R. (1952). Design for a Brain: the Origin of Adaptive Behaviour. Revised edition 1960. London: Chapman & Hall.

Bosse, T., Jonker, C. M., van der Meij, L., Sharpskykh, O., & Treur, J. (2019). Specification and Verification of Dynamics in Agent Models. International Journal of Cooperative Information Systems, 18, 167–193.

Tinklepaugh, O. L. (1932). Multiple delayed reaction with chimpanzees and monkeys. Journal of Comparative Psychology, 13, 207–243.

Tollefsen, D. P. (2010). From extended mind to collective mind. Cognitive Systems Research, 7, 140–150.

Treur, J. (2007). Temporal Factorisation: A Unifying Principle for Dynamics of the World and of Mental States. Cognitive Systems Research, 8(2), 57–74.

Treur, J. (2009). Temporal Factorisation: Realisation of mediating state properties for dynamics. Cognitive Systems Research, 8(2), 75–88.

Treur, J. (2016). Network-Oriented Modeling: Addressing Complexity of Cognitive Affective and Social Interactions. Cham, Switzerland: Springer Nature Publishers.

Treur, J. (2019). Adaptive Network Modeling for Criterial Causation In: H. Cheriﬁ, S. Gaito, J.F. Mendes, E. Moro, L.M. Rocha (Eds.), Complex Networks and Their Applications VIII: Proc of the Eighth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2019. Studies in Computational Intelligence (Vol. 2, pp. 827–841). Springer Nature Publishers (vol 882).

Tse, P. U. (2018). Two Types of Libertarian Free Will are Realized in the Human Brain. In G. D. Caruso & O. J. Flanagan (Eds.), Neuroexistentialism: Meaning Morals and Purpose in the Age of Neuroscience (pp. 248–290). Oxford University Press.

van Gelder, T. J., & Port, R. F. (1995). It’s About Time: An Overview of the Dynamical Approach to Cognition In R.F. Port, T. van Gelder (Eds.) (1995) Mind as Motion: Explorations in the Dynamics of Cognition (pp 1–43). Cambridge Mass: MIT Press.