

Letter

Systemic States of Spreading Activation in Describing Associative Knowledge Networks: From Key Items to Relative Entropy Based Comparisons

Ismo T. Koponen 

Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland; ismo.koponen@helsinki.fi

Abstract: Associative knowledge networks are central in many areas of learning and teaching. One key problem in evaluating and exploring such networks is to find out its key items (nodes), sub-structures (connected set of nodes), and how the roles of sub-structures can be compared. In this study, we suggest an approach for analyzing associative networks, so that analysis is based on spreading activation and systemic states that correspond to the state of spreading. The method is based on the construction of diffusion-propagators as generalized systemic states of the network, for an exploration of the connectivity of a network and, subsequently, on generalized Jensen–Shannon–Tsallis relative entropy (based on Tsallis-entropy) in order to compare the states. It is shown that the constructed systemic states provide a robust way to compare roles of sub-networks in spreading activation. The viability of the method is demonstrated by applying it to recently published network representations of students’ associative knowledge regarding the history of science.

Keywords: systemic states; associative knowledge; networks; diffusion propagator



Citation: Koponen, I.T. Systemic States of Spreading Activation in Describing Associative Knowledge Networks: From Key Items to Relative Entropy Based Comparisons. *Systems* **2021**, *9*, 1. <https://dx.doi.org/10.3390/systems9010001>

Received: 8 December 2020

Accepted: 18 December 2020

Published: 23 December 2020

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2020 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Knowledge that is central in learning and teaching often starts with factual knowledge of agents, objects, and events, and how they are related. This is equally true in such different areas of learning and instruction as, for example, in physics [1–3] and history [4,5]. The familiarization of new key concepts involves making connections with what is already known and how new items and terms can be integrated into more extensive knowledge structures [6–8]. In that process, associative recall and memorization of thematic connections and relationships between knowledge items (e.g., words, terms or concepts) create an interlinked system of knowledge items, which is referred to as associative network [9–12]. The knowledge processing may then continue with more organized strategies, eventually leading to more integrated knowledge systems [6–8].

The formation of associative connections between knowledge items can be thought to occur as pairwise (dyadic) thematic word or term associations, which are established on the basis of thematic resemblance or kinds of taxonomic family resemblance. Consequently, associative knowledge is often modelled as a complex network of interlinked words and terms, where the associative connection becomes activated through the spreading of activation between words and terms that are related, for example, either thematically or taxonomically [9–12]. The idea of spreading activation [11,12] within semantic and associative networks has been utilized in several areas of cognitive psychology in modeling retrieval and memorization of knowledge or words (see [12,13] and the references therein). According to spreading activation, when one word or term in a semantic or associative network is activated, that activation spread to other words, through connections within the semantic network and so it affects the state of the network [11–13]. Thus, spreading activation has provided a dynamic and systemic perspective on associative knowledge networks, guiding attention on how different parts of the networks are interacting or communicating.

Spreading activation has been modelled as a diffusion or random walk process in a network, where a random walker (e.g., information that is mediated and causes activation) that starts from a given node propagates to nodes that are accessible through links connecting the nodes [12,13]. In this study, we suggest a model, where the focus is shifted from spreading activation as a process to holistic states of the network, where pairwise activations between nodes in course of spreading activation are equated with state of the system. To this state, we refer in what follows as the systemic state. The basic picture of spreading process is similar as in random walk based models, but modelling through systemic states quantifies the dynamically changing state of the network and brings it in focus. In constructing the systemic states, we follow recent advances to model diffusion-like dynamics on complex networks [14–19]. In that approach, the connectivities within the networks (i.e., nodes and links between nodes) are taken as starting point to describe how a network becomes activated when diffusion spreads, and the normalized probabilities that describe the activations of links at a given instant are associated with a systemic state.

One important question and open problem in spreading activation is to quantify how it is affected by different types of sub-structures within the network, most typically parts of the network that belong to taxonomically or thematically different categories (i.e., different types of words or terms). This is essentially a problem that is related to a comparison of networks [17–19]. Based on systemic states of a network, it is possible to construct information theoretic entropic measures in order to quantify the differences between networks, one of the most robust and widely used being Jensen–Shannon divergence, which is a symmetric relative entropy between two different states [14–19]. The relative entropy measures the information lost, when two networks are joined, in comparison to information that is contained in original state. Here, we generalize the Jensen–Shannon relative entropy for the model of nonlinear spreading activation and the systemic states corresponding it.

In this study, we show that the generalization provided in the present study have good resolving power and sensitivity in exploring the role of sub-structures in spreading activation in associative networks. As a concrete example, in order to demonstrate the viability of analysis that is based on systemic states, we use empirical data from a recently reported case of university students' associative knowledge regarding the history of science [3]. Such a network has an interestingly complex structure and a rich internal community structure organized around key items (often persons) in the network. However, it proved to be difficult to quantify the role of relationships between persons, scientific ideas, and inventions, and general historical events in the formation of the whole network [3]. Here, attention is on roles of the different type of relations between key items in the network and how they constitute the sub-structures within the network as a whole. The systemic states constructed to describe spreading activation are used as a starting point to quantify the importance of the sub-networks for spreading activation in a complete network.

2. Methods: Systemic States and Their Comparison

The model for spreading activation, which allows for exploring associative networks and the role of its substructure in spreading activation, is developed in three steps. First, we introduce the quantities that describe the connectivity of the network, the so-called adjacency matrix \mathbf{A} and diagonal matrix \mathbf{D} for node degrees, and on basis of them, graph Laplacian \mathbf{L} describing diffusion-like spreading is introduced. Second, based on graph Laplacian \mathbf{L} , generalized systemic states ρ_q are constructed. Third, and, finally, relative information theoretic entropy (generalized Jensen–Shannon–Tsallis divergence) is introduced to compare systemic states.

2.1. Graph Laplacian and Diffusive Spreading

The role of a given node in spreading activation is related to its capacity (or activity) to channel information to other nodes within a network. This depends on how a given node is connected within a network and how it can communicate through the connections; how activation spreads throughout the network. Spreading activation can be modeled

using a diffusion-like propagator, which is based on a graph Laplacian [18,19]. The graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{A}$ of a network is defined in terms of adjacency matrix \mathbf{A} with elements $[\mathbf{A}]_{ij} = a_{ij}$, where $a_{ij} = 1$ if nodes i and j are connected and, otherwise, $a_{ij} = 0$. In what follows, symmetric adjacency matrices are assumed. The diagonal matrix \mathbf{D} is composed of elements $[\mathbf{D}]_{ii} = d_i$, where $d_i = \sum_j a_{ij}$ is the degree of node i . The matrix elements of the so defined graph Laplacian are thus given by $[\mathbf{L}]_{ij} = d_i \delta_{ij} - a_{ij}$, where δ_{ij} is Kronecker's delta as usual. The graph Laplacian \mathbf{L} can be interpreted as a discrete Laplacian operator describing random walk in a network, analogous to ordinary Laplacian operator ∇^2 in a normal diffusion equation [20,21].

Next, we represent the property of node that describes its role in diffusive spreading as a vector $\bar{p} = (p_1, p_2, \dots, p_N)$ in a network of N nodes. The property \bar{p} can be associated with a probability that a random walker in a network stays in a node or, alternatively, as a capacity of node to retain the information (i.e., to block the diffusion) [18,19]. Note that it is not necessary to define the property \bar{p} in more detail, because the focus of interest will be the diffusion propagator, which can be used in order to describe the diffusion (spreading) dynamics of the network.

The diffusion process in a network, by using graph Laplacian \mathbf{L} , can now be described by discrete difference equation [20,21] (see also refs. [18,19])

$$\frac{d\bar{p}}{d\beta} = -\mathbf{L}\bar{p}, \quad (1)$$

This equation has a solution $\bar{p} = \exp[-\beta\mathbf{L}]\bar{p}_0$, where β has a role analogous to time and \bar{p}_0 is the initial state. According to this picture, the propagator $\exp[-\beta\mathbf{L}]$ regulates how property that is described by \bar{p} spreads out in that diffusion process. In what follows, the propagator and its generalization are in focus.

2.2. Systemic States and Activity

We now generalize the diffusion equation in Equation (1), so that the property \bar{p} of nodes to affecting spreading activation is taken into account non-linearly, so that high values of \bar{p} become reduced (i.e., their role in retardation is diminished) in relation to values that are low. This means that nodes that facilitate spreading are allowed to become even more important for it. In that case, the evolution equation in Equation (1) for the property \bar{p} of interest generalizes to (as compared with e.g., [22])

$$\frac{d\bar{p}}{d\beta} = -\mathbf{L}\bar{p}^q, \quad \text{where } q \in [1, 2] \quad (2)$$

The value $q = 1$ corresponds to the spreading activation that parallels with normal diffusion. It is straightforward to show that a (normalized) solution of Equation (2) is $\bar{p} = \rho_q(\beta)\bar{p}_0$ (for initial state \bar{p}_0), where

$$\rho_q(\beta) = Z_q^{-1} [\mathbf{I} - (q-1)\beta\mathbf{L}]^{-1/(q-1)}, \quad (3)$$

with $Z_q = \text{Tr} [\mathbf{I} - (q-1)\beta\mathbf{L}]^{-1/(q-1)}$ as the normalization factor. In Equation (3), the expression $\exp_q[\mathbf{x}] = [\mathbf{I} - (q-1)\mathbf{x}]^{-1/(q-1)}$ is the q -generalized matrix exponential, which, in the limit $q \rightarrow 1$, approaches the normal matrix exponential [23,24]. The propagator ρ_q has the central role in regulating the spreading activation when modeled as a diffusion or random walk process. Consequently, we now take the solution ρ_q in Equation (3) to represent the systemic state of the network in the stage of spreading activation that corresponds to given value of time-like parameter β and refer to it as q -generalized systemic state in what follows. The q -generalized systemic state is mathematically well-defined positive semidefinite matrix possessing eigenvalues that are equal or larger than zero (and, thus, a proper density matrix) [23,24]. A similar description of networks through a state corresponding to a normal matrix exponential in the limit $q \rightarrow 1$ is provided in

reference [19]. The model for spreading activation based on Equations (2) and (3) can be taken as a generalization of such a description.

In order to define the changing, dynamic activity α_i of a node i in spreading activation, we note that the diagonal values $[\rho_q]_{ii} \in [0, 1[$ of systemic state are in a diffusion-picture proportional to average probability of return to a node [18,19]; the lower the diagonal value of the systemic state for node i , the more important it is that the node is in activation spreading. This notion allows for us to define the dynamic activity α_i of a node i as

$$\alpha_i = -\log[\rho_q]_{ii}, \text{ where } [\rho_q]_{ii} \in [0, 1] \quad (4)$$

The dynamic activity α_i is then the basic quantity for monitoring the role of individual nodes in activation spreading, with a high value of α_i (i.e., low value of $[\rho_q]_{ii}$) corresponding to high activity and a low value a low activity.

2.3. Divergence for Comparisons

We next ask how two different systemic states ρ_q and σ_q of the network can be compared. An obvious approach is to use relative information theoretic entropy as a measure of difference (see, e.g., [14,15,18,19]). Relative information theoretic entropy quantifies the difference between information that is contained in initial states in comparison to the state obtained by completely mixing the initial states. Thus, it is a measure of maximal loss of information in mixing. Relative entropies are widely used in a comparison of networks, because they provide a well-defined macrolevel description of difference and take the networks as a whole into account in quantifying the differences [14,15,18,19]. Here, a method of comparison based on relative information theoretic entropy is generalized for the q -generalized systemic states defined in Equation (3).

The q -generalized states in Equation (3) are closely related to a q -generalized (non-extensive) Tsallis-entropy [23,24], and they emerge as density matrices that extremize Tsallis-entropy. In that picture, parameter q can be related to the degree of non-extensivity, often originating, e.g., from fractality or compartmentalization of the system [24]. However, deeper discussion is not necessary here, and we can treat q simply as parameter controlling the non-linearity in Equation (2). However, to connect the q -generalized state to Tsallis-entropy, we allow the q -index q^* for Tsallis-entropy to differ from value of q in Equations (2) and (3), and only afterwards, fix the connection between the two indices. The Tsallis-entropy with index q^* is

$$H_{q^*} = \frac{1 - \text{Tr} \rho^{q^*}}{q^* - 1} \text{ where } q^* \geq 0 \quad (5)$$

Depending on constraints posed on maximization, the Tsallis entropy, which is maximized by ρ_q , may have index q^* different from q [23–25]. Here, we choose to pose the linear constraints on extremization of the Tsallis-entropy, by requiring: (1) the conservation of probability $\text{Tr} \rho_q = 1$ and (2) constancy of expectation values of the Laplacian $\text{Tr} \mathbf{L} \rho_q$ [23–25]. An alternative would be to adopt the so-called escort-probabilities [23–25], but we choose to avoid that choice. With the choice of linear constraints, the index q^* is [23–25]

$$q^* = 2 - q \quad (6)$$

The relation in Equation (5) between indices q and q^* is typical to the so-called dual forms of Tsallis-distributions, encountered when so-called escort probabilities are excluded in favor for the ordinary probabilities and expectation values related to them [23–25]. In the limit $q^* \rightarrow 1$ reduces to von Neumann entropy $H = -\text{Tr} \rho \ln \rho$, where \ln is the natural logarithm.

We can now proceed to define a symmetric version of relative information theoretic entropy for a comparison of states ρ_q and σ_q , and that is compatible with the Tsallis-entropy with index q^* . This is a q -generalization J_{q^*} of Jensen–Shannon–Tsallis relative entropy [14,15,18,19]

$$J_{q^*}[\rho_q|\sigma_q] = H_{q^*}\left[\frac{\rho_q + \sigma_q}{2}\right] - \frac{1}{2}(H_{q^*}[\rho_q] + H_{q^*}[\sigma_q]) \quad (7)$$

where the first term at right is the entropy of the mixed state $(\rho_q + \sigma_q)/2$ and the last term an average of entropies that correspond to initial states. In what follows, J_{q^*} defined in Equation (7) is referred as q-generalized Jensen–Shannon–Tsallis divergence (q-JST divergence in short) and is the basis for comparison of different systemic states of the network. Note that, in defining J_{q^*} , the index $q^* = 2 - q \in [0, 1]$ is different from q-index for states ρ_q and σ_q , thus weighting low-probability states of interest more than the ordinary Jensen–Shannon divergence ($q^* = 1$) (compare with refs. [14,15], where q^* is treated as a free parameter).

The different mathematical quantities that are used in the step-by-step derivation above are summarized in Table 1 for reference. Of these, the activity α_i and q-JST divergence J_{q^*} are used in exploring an example of an associative network.

Table 1. Key quantities and their mathematical symbols used in definitions. A brief description is provided for each quantity and the range of variation of parameters are given.

Quantity and Symbol	Parameters	Description	
Adjacency matrix	A	-	Matrix of links (entries 0 or 1) between nodes
Degree matrix	D	-	Matrix (diagonal) containing degrees of nodes
Laplacian matrix	L	-	Matrix (diagonal) for diffusion (Laplacian) operator
Systemic state	ρ_q	$q \in]1, 2], \beta > 0$	q-generalized systemic state of spreading activation
Activity of node	α_i		Activity of a node in spreading activation
Tsallis-entropy	H_{q^*}	$q^* = 2 - q \in [0, 1[$	Non-extensive entropy with index q^*
q-JST divergence	J_{q^*}	$q^* = 2 - q \in [0, 1[$	q-generalized Jensen-Shannon-Tsallis divergence

3. Results of Application: Associative Knowledge Network

The associative networks that are explored in this study from a systemic perspective are based on a recent study, which explored university students' (associative) knowledge networks related to history of science [3]. Here, detailed construction of the networks and results of their analysis for key terms and their rankings are not repeated. Instead, having the previous results [3] available for the present study, the focus is here on global properties of the networks and role of sub-networks within it. First, however, essential aspects of network construction are briefly summarized. Second, the activity of nodes in spreading activation is obtained, based on q-generalized systemic states. Third, the q-JST divergence is calculated in order to explore the role of sub-structures of the associative network in spreading activation.

3.1. Networks and Key Items

The empirical data of associative network are based on a sample of data that were collected from a science (physics) history course for a third and fourth year students (pre-service teachers) [3]. The course discussed science history roughly from year 1550 to 1850 from a viewpoint placing history of physics and science as part of general history. The data set provides information regarding the students' judgments of the relevance of given persons, ideas, inventions, and events during those three centuries. The agglomerated associative network based on that data was presented in ref. [3] and its key items were analyzed. Here, the same data, made available for the present study, is analyzed from the perspective of spreading activation and systemic states that are based on it. Note that, here, only the two-core (nodes having at least two links) of the whole networks is used in analysis, because auxiliary nodes with single connections only are of low relevance here. The two-core consists of about 600 nodes and 1500 links (the original network has 1300 nodes and 2500 links).

The agglomerated associative knowledge network in focus here consists of items in three groups: (a) historical characters and persons (persons), (b) scientific ideas and inven-

tions (science), and (c) events and phenomena in general history (history). Consequently, the network contains six sub-networks: (a) persons-to-persons, (b) persons-to-science, (c) persons-to-history, (d) science-to-science, (e) science-to-history, and (f) history-to-history, which are schematically shown in Figure 1. The groups that are shown in Figure 1 show that person-to-science connections are the most dominant ones and, after that, person-to-history connections. The connections that are the fewest and supposedly of the least importance for the overall structure of the networks are science-to-history and history-to-history connections. Although the relative importance of different sub-networks are visually seen in Figure 1, it is difficult to make more detailed comparisons of the relative importance. For that, we will introduce later on the q-JST divergence, which quantifies the differences that are already visible in Figure 1.

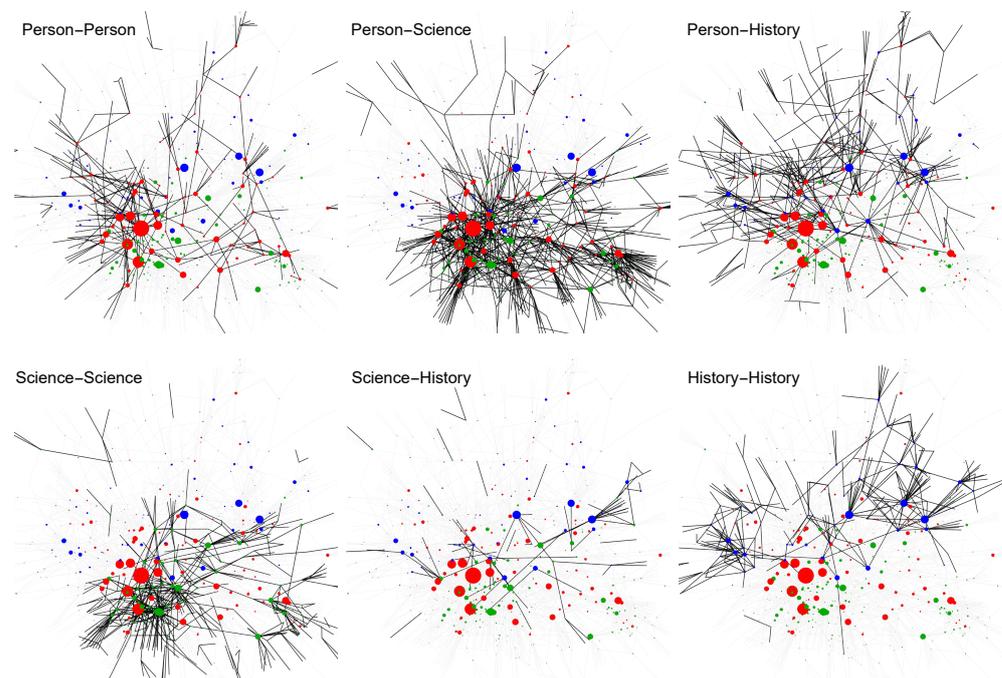


Figure 1. Sub-networks (see legends) of complete associative knowledge network. The nodes corresponding to persons are shown in red, science in green and history in blue. The size of the node is proportional to a number of links (degree) attached to it. In each of the six cases, the links forming the sub-networks are shown in bold black lines.

In what follows, for purposes of demonstrating the viability of method of analysis, it is not necessary to know, in more detail, what different nodes represent. It is enough to mention that nodes fall in three categories, as follows: persons (red nodes in Figure 1 and nodes 1–6, 11, 12, 14 in Figure 2); science (green nodes and 7, 9, 10, 13, 15, 16, 18 in Figure 2); and, history (blue nodes and 8, 17–20 in Figure 2). Although the six leading nodes are persons, the nodes are distributed relatively evenly in all three categories, which is also the case for other remaining 43 out of 63 that were included in analysis (see Figure 2 later on).

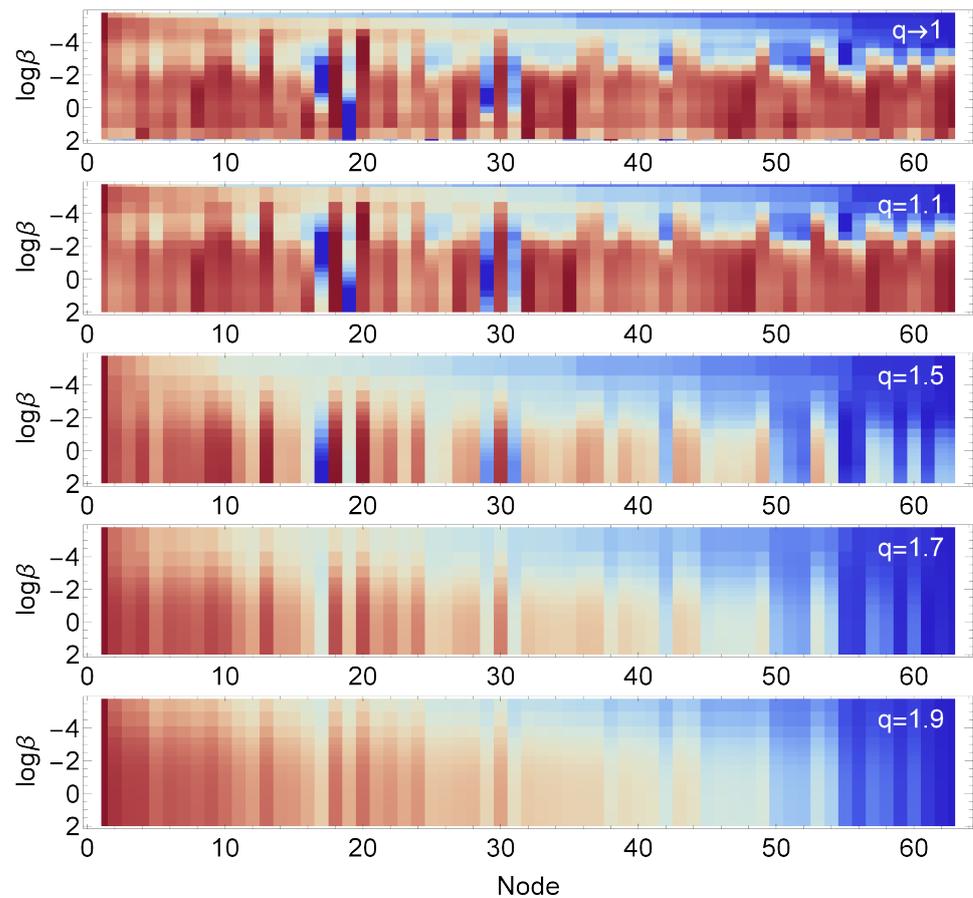


Figure 2. Activity $\alpha_i = -\log[\rho_q]_{ii}$ of nodes i for systemic states ρ_q for $q \rightarrow 1$ and $q = 1.1, 1.5, 1.7,$ and 1.9 (from top to bottom) for 63 most important key items (nodes). With red tones are shown the high activity nodes, while blue indicates nodes of lower activity. The first 20 nodes fall in three categories, as follows: persons (1–6, 11, 12, 14); science (7,9,10, 13, 15, 16, 18); and, history (8, 17–20).

3.2. Activity of a Node in Spreading Activation

The activities α_i calculated from Equation (4) of the most active 63 nodes are shown in Figure 2, as a heat-maps for values of β scaled to value 1 for the maximum of the q -divergence J_q^* (to be discussed later on). Figure 2 is a heat map showing how the activity (high activity in red, lower activity with blue) of the most important 63 nodes changes when the parameters q and β are increased. The relevant information to be paid attention on in Figure 2 is contained in color coding, showing how parts of nodes gain higher activity, while, for a part of the nodes, activity is reduced. It is not intended that the exact values of activities (which are not relevant here) can be read from the Figure 2, since the exact values are of no relevance here.

For $q \rightarrow 1$ and low values of β (the topmost panel in Figure 2), the nodes with large degrees are the most active, but, when β is increased (i.e., activity spreads increasingly through the network), the rankings of the nodes change, eventually leading to a situation that differs from the initial stage. However, then, nearly all nodes gain similar activity, indicating that state with $q \rightarrow 1$ weights all nodes nearly equally, and thus becomes of low value for differentiating the key nodes based on their activity.

When higher values of q are used to define the systemic state, thus giving more weight on nodes of high activity, it is seen that nodes are grouped more orderly in high activity (in red tones) and lower activity (in blue tones) clusters. Visually, the largest diversity of values is activities appears to be reached for $q \approx 1.5$. For larger values of q , the breakdown to different clusters of high and low values of activities does not change appreciably, but the relative differences (diversity) appear to be reduced. In all these cases, irrespective of value

of q , the breakdown in the clusters does not change much with increasing values of β . Consequently, a robust classification of key items (nodes) based on their dynamic activity is obtained for $q \approx 1$, with good stability with increasing values of β , as seen in Figure 2.

3.3. Role of Sub-Networks: Divergences

The role of sub-networks (as shown in Figure 1) for spreading activation can be quantified through systemic states, where a given sub-network is removed from the complete network (by setting links contained in it to zero) and then, the q -divergence between the so-modified and complete network are calculated. In what follows, the modified networks are denoted by G_{X-Y} referring to network obtained by removing sub-network consisting of links X-to-Y (where, X and Y are either, person, science, or history) from a complete network G. The corresponding systemic states are calculated in order to obtain q -divergence, which is simply denoted as $J_{q^*}[X, Y]$, with results shown in Figure 3.

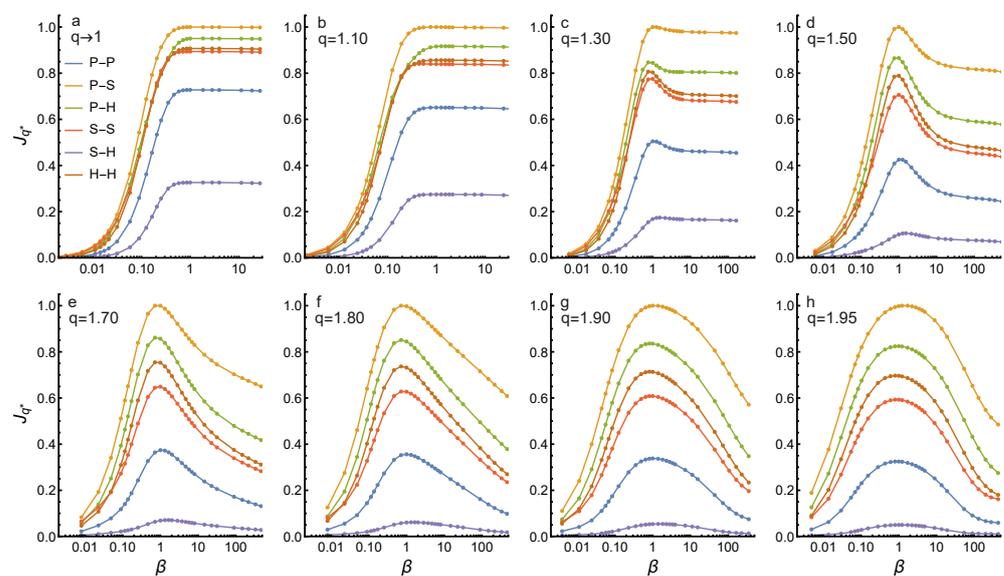


Figure 3. Divergences $J_{q^*}[X, Y]$ where X and Y persons (P), science (S) or history (H). The results are shown for $q^* = 2 - q$ for $q \rightarrow 1$ (a), $q = 1.10$ (b), $q = 1.30$ (c), $q = 1.50$ (d), $q = 1.70$ (e), $q = 1.80$ (f), $q = 1.90$ (g), and $q = 1.95$ (h). Divergences are scaled to maximum value 1 of the largest divergence (P-S) and they are shown as a function of β also scaled value of 1 corresponding to the maximum of the largest divergence (but, for $q < 1.3$, corresponding to maximum of $q = 1.3$).

The lowest value $q \rightarrow 1$ corresponds to the exponential systemic state, leading to von Neumann entropy and normal Jensen–Shannon relative entropy. For low values $\beta \ll 1$ q -divergences are always close to zero, because all nodes contribute nearly equally (state is essentially an identity matrix, see Equation (3)). Only with increasing values of β , divergences start to increase. The q -divergence J_1 with $q \rightarrow 1$ provides large values of divergences when sub-networks P-S, P-H, S-S, and H-H are removed and a clearly smallest divergence when S-H is removed, as seen in Figure 2. This is, of course, in line with the situation that is shown in Figure 1, where it is visually clear that most of the links are found in sub-networks whose removal leads to largest divergences. However, this means that J_1 provides no more information than is already contained in link counting. Thus, the conventional exponential state with $q \rightarrow 1$ and the corresponding divergence J_1 only pay attention to local connectivity of nodes and they are insensitive to global connectivities. This interpretation is also in accordance with the behavior of activities α_i for $q \rightarrow 1$ that is shown in Figure 2.

With increasing values of $q > 1$ the q -divergence begins to be more sensitive on differences between the sub-networks. The optimal case is reached for $1.3 < q < 1.5$, where the differences between sub-networks are relatively large and the maximum clear, but where

q-divergence for large values of β levels off to stable, constant values. This stability is also observed in activity of nodes that were obtained for values $q \approx 1.5$, as shown in Figure 2. For values $q > 1.5$, the divergences begin steadily to decrease with increasing values of β , which indicates that spreading activation as described by systemic states covers all nodes. In that region, the resolving power of q-divergence is still retained, but the relative differences between activity of nodes does not change. This is also obvious in heat maps showing the activities for $q = 1.7$ and $q = 1.9$, which are quite similar. This behavior means that q-divergences with $1.3 < q < 1.5$ are optimal in providing information regarding the spreading activation and roles of sub-networks. The results for divergences in these cases can be interpreted to contain robust information of the role of sub-networks, when global connectivities within the sub-networks are taken into account. In this case, the role of sub-networks appear quite differently, the sub-network P-P having the largest effect, and clearly larger than the next most important sub-network P-H. Furthermore, now, the sub-network S-H has a nearly negligible role for spreading activation in the complete network.

4. Discussion

A comparison of networks is an important problem in many areas of network science [14–19]. In this study, we suggested a method of analysis of networks and their comparison based on systemic states of network, which are constructed to model spreading activation [9–12]. Spreading activation picture can be modeled as diffusion-like process, where the extent of spreading is described in terms of diffusion propagator, in the form of a density matrix that takes into account all connections between the terms included in the system (network) [18,19]. Here, the q-generalized systemic states are constructed in the form of a diffusion-propagator, which allows for weighting the role nodes in spreading activation differently, with values of $q = 1$ corresponding to normal diffusion and $q > 1$ cases where a role of high activity nodes is emphasized in spreading. The diffusion propagator is a holistic and complete description of the whole network describing its changing state in spreading activation. Consequently, the propagator can be taken as a systemic state describing the network.

On basis of the q-generalized systemic states, it becomes possible construct an associated q-generalization of the usual Jensen–Shannon relative entropy (divergence) measuring the difference between the states [18,19]. The q-generalized divergence is based on Tsallis-entropy with index $q^* = 2 - q$ that results from a specific choice of (linear) constraints, related to so-called dual-forms in Tsallis-statistic. A practical motivation for introducing the relation $q^* = 2 - q$ is to achieve an optimal resolution for quantifying the differences between q-generalized states, the dual index $q^* = 2 - q$ in q-generalized divergence weights more the important states occurring with low probability, when $q > 1$. The best resolving power of differences is achieved when $1.3 < q < 1.7$ and, correspondingly, $0.3 < q^* = 2 - q < 0.7$.

The viability and advantages of the approach based on q-generalized systemic states and divergences is demonstrated here by applying it in the recently studied case of university students' associative knowledge of history of science, represented as an agglomerated network [3]. It was shown that an analysis that is based on systemic states for spreading activation has the potential to be better than conventional approaches based on use of exponential states ($q \rightarrow 1$) in revealing and quantifying the role of different sub-networks in spreading activation in associative networks (and, consequently, in similar kinds of complex networks in general). The optimal resolving power to detect global differences was achieved when $1.3 < q < 1.5$ ($0.7 < q^* < 0.5$). The results of the analysis were fully compatible with the more traditional analysis [3], but it was shown that systemic states provide better insight on roles the sub-networks have in a complete network, through their activity in spreading activation. Moreover, the method provides robust and reliable way to quantify such differences, with the possibility of much potential in quantitative analysis and comparisons of complex systems that can be represented in the form of networks.

5. Conclusions

The systemic approach for analyzing complex networks, as suggested here, is based on the construction of systemic, holistic states to describe the network, from a perspective of diffusion-like spreading activation. The method quantifies the notion of spreading activation, central for many applications related to associative knowledge networks. The new results contained in the present study are: (1) q -generalized systemic states that are constructed as diffusion propagator solutions of a non-linear discrete diffusion equation describing spreading activation, and (2) q -generalized relative information theoretic entropy measure for quantifying differences between systemic states, which is constructed as a Jensen–Shannon–Tsallis divergence for the q -generalized states. The q -generalized systemic states contain two parameters q and β , which can be used to tune the states sensitive to desired scales, from local to global connectivity, corresponding to extent of spreading activation, thus providing new tools for robust quantitative characterization of complex networks from a systemic perspective.

Finally, the method was applied in the case of a recently empirically studied associative knowledge network, with the results demonstrating the viability of the method and its sensitivity on global, systemic level properties of the network in the course of spreading activation. In summary, the approach that is suggested here embodies the spreading activation view on networks and provides a quantification in terms of systemic states. The results of application demonstrate not only the viability of the proposed method, but also its better sensitivity on the details of the networks not easily captured by more conventional analysis methods.

Funding: This research was funded by Academy of Finland grant number 311449.

Data Availability Statement: Data sharing not applicable. No new data were created or analyzed in this study.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Kubsch, M.; Touitou, I.; Nordine, J.; Fortus, D.; Neumann, K.; Krajcik, J. Transferring Knowledge in a Knowledge-in-Use Task—Investigating the Role of Knowledge Organization. *Educ. Sci.* **2020**, *10*, 20. [[CrossRef](#)]
2. Thurn, M.; Hänger, B.; Kokkonen, T. Concept Mapping in Magnetism and Electrostatics: Core Concepts and Development over Time. *Educ. Sci.* **2020**, *10*, 129. [[CrossRef](#)]
3. Koponen, I.; Nousiainen, M. University students' associative knowledge of history of science: Matthew effect in action? *Eur. J. Sci. Math. Educ.* **2012**, *6*, 69–81.
4. van Boxtel, C.; van Drie, J. "That's in the Time of the Romans!" Knowledge and Strategies Students Use to Contextualize Historical Images and Documents. *Cogn. Instr.* **2012**, *30*, 113–145. [[CrossRef](#)]
5. Van Drie, J.; Van Boxtel, C. Historical reasoning: Towards a framework for analyzing students' reasoning about the past. *Educ. Psych. Rev.* **2008**, *20*, 87–110. [[CrossRef](#)]
6. Ifenthaler, D.; Hanewald, R. *Digital Knowledge Maps in Education: Technology-Enhanced Support for Teachers and Learners*; Springer: New York, NY, USA, 2014.
7. Lachner, A.; Nückles, M. Bothered by Abstractness or Engaged by Cohesion? Experts' Explanations Enhance Novices' Deep Learning. *J. Exp. Psych. Appl.* **2015**, *21*, 101–115. [[CrossRef](#)]
8. Amadiou, F.; van Gog, T.; Paas, F.; Tricot, A.; Marine, C. Effects of Prior Knowledge and Concept-map Structure on Disorientation, Cognitive Load and Learning. *Learn. Instr.* **2009**, *19*, 376–386. [[CrossRef](#)]
9. Siew, C.S.Q. Applications of Network Science to Education Research: Quantifying Knowledge and the Development of Expertise through Network Analysis. *Educ. Sci.* **2020**, *10*, 101. [[CrossRef](#)]
10. Steyvers, M.; Tenenbaum, J.B. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cogn. Sci.* **2005**, *29*, 41–78. [[CrossRef](#)]
11. Collins, A.M.; Loftus, E.F. A spreading-activation theory of semantic processing. *Psychol. Rev.* **1975**, *82*, 407–428. [[CrossRef](#)]
12. Lerner, I.; Bentin, S.; Shriki, O. Spreading Activation in an Attractor Network With Latching Dynamics: Automatic Semantic Priming Revisited. *Cogn. Sci.* **2012**, *36*, 1339–1382. [[CrossRef](#)]
13. Hills, T.T.; Jones, M.N.; Todd, P.M. Optimal Foraging in Semantic Memory. *Psychol. Rev.* **2012**, *119*, 431–440. [[CrossRef](#)]
14. Gerlach, M.; Font-Clos, F.; Altmann, E.G. Similarity of Symbol Frequency Distributions with Heavy Tails. *Phys. Rev. X* **2016**, *6*, 021009. [[CrossRef](#)]

15. Ré, M.A.; Azad, R.K. Generalization of Entropy Based Divergence Measures for Symbolic Sequence Analysis. *PLoS ONE* **2014**, *9*, e93532. [[CrossRef](#)]
16. Dias, L.; Gerlach, M.; Scharloth, J.; Altmann, E.G. Using text analysis to quantify the similarity and evolution of scientific disciplines. *R. Soc. Open Sci.* **2018**, *5*, 171545. [[CrossRef](#)]
17. Fu, S.; Shi, G.; Petersen, I.R.; James, M.R. Entropy Evolution in Consensus Networks. *Sci. Rep.* **2017**, *7*, 1520. [[CrossRef](#)]
18. Biamonte, J.; Faccin, M.; De Domenico, M. Complex networks from classic to quantum. *Commun. Phys.* **2019**, *2*, 53. [[CrossRef](#)]
19. De Domenico, M.; Biamonte, J. Spectral Entropies as Information-Theoretic Tools for Complex Network Comparison. *Phys. Rev. X* **2016**, *6*, 041062. [[CrossRef](#)]
20. Estrada, E. *The Structure of Complex Networks: Theory and Applications*; Oxford University Press: Oxford, UK, 2012.
21. Newman, M.E.J. *Networks: An Introduction*; Oxford University Press: Oxford, UK, 2010.
22. Tsallis, C. Economics and Finance: q-Statistical Stylized Features Galore. *Entropy* **2017**, *19*, 457. [[CrossRef](#)]
23. Guo, R.; Duo, J. The adiabatic static linear response function in nonextensive statistical mechanics. *Physica A* **2014**, *414*, 414–420. [[CrossRef](#)]
24. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*; Springer: New York, NY, USA, 2009.
25. Ferri, G.L.; Martinez, S.; Plastino, A. Equivalence of the four versions of Tsallis's statistics. *J. Stat. Mech.* **2005**, *2005*, P04009. [[CrossRef](#)]