PNAS
www.pnas.org

Supplementary Information for

Optimizing the Human Learnability of Abstract Network Representations

Qian, W., Lynn, C.W., Klishin, A.A., Stiso, J., Christianson, N.H., Bassett, D.S.

To whom correspondence should be addressed: Dani S. Bassett, Department of Bioengineering University of Pennsylvania, 210 S. 33rd Street, 240 Skirkanich Hall, Philadelphia, PA 19104-6321. E-mail: dsb@seas.upenn.edu, Phone: 215-746-1754.

This PDF file includes:

Supplementary text
Figs. S1 to S7 (not allowed for Brief Reports)
Table S1 (not allowed for Brief Reports)
SI References

Supporting Information Text

Supplementary Results.

The analytic approach. One way of gaining insight into how to optimally modulate emphasis of edges in a network presented to a human learner is by directly solving for the optimal transition matrix from Equation 1 (main text). In particular, given some target transition matrix A we can exactly solve the relation $f(A^*) = A$ to find the proper A^* to present to the learner, finding

$$A^* = ((1 - e^{-\beta})I + e^{-\beta}A)^{-1}A. \quad [1]$$

In the limit $\beta \to \infty$, we recover the result that $A^* \to A$, agreeing with the approach for finding the optimally learnable network explored in the main text.

However, for lower values of β, the two approaches differ drastically, as there are a few limitations that make the direct solution of A^* difficult to use. Firstly, the analytic A^* matrix nearly always includes non-negative elements on the diagonal, making it impractical for usage in situations where presenting networks with self-loops to a human learner is not considered.

Next, to understand the properties of the analytic A^* further, it is useful to consider the eigendecomposition of A. Specifically, we can write

$$A_{ij} = \sum_{k=1} v_i^{(k)} \lambda_k v_j^{(k)}, \quad [2]$$

where λ_k represents the k^{th} eigenvalue and $v_i^{(k)}$ represents the i'th entry of the corresponding eigenvector.

It is important to note that A^* shares the same eigenvectors as A, because it can be expressed as a combination of linear transformations and inversions of A. Additionally, the eigenvalues of A^* are transformed via analogous linear transformations and inversions as follows:

$$A^*_{ij} = \sum_k v_i^{(k)} \frac{\lambda_k}{(1 - e^{-\beta}) + e^{-\beta}\lambda_k} v_j^{(k)}. \quad [3]$$

The denominator of the expression in the summation shown in Equation S3 suggests that it is possible for A^* to be singular. In particular, when $\lambda_k = 1 - e^{\beta}$ holds for any eigenvalue, the expression diverges. Furthermore, if we restrict our analyses to symmetric, connected target networks A (as done in the main text), we can apply the Perron-Frobenius theorem to conclude that $-1 \leq \lambda_k \leq 1$ for all eigenvalues. from this, we can conclude that for $\beta < \ln(2) \approx 0.693$, it is possible for Equation S3 to diverge.

To investigate whether this poses a problem in practice, we computationally analyzed the properties of the analytic A^* for various networks over a range of β values (Fig. S1). In particular, we applied the approach to the the modular graph (Fig. 1A, main text), a random graph with the same number of nodes and edges, and a complete graph corresponding to a uniform transition matrix. We found that for all three networks, there were some values of β below ln(2) that resulted in the argument to the matrix inverse in Equation S1 to become singular. While it may be surprising that even the A^* obtained for the transition network corresponding to the complete graph suffers from this issue, we note that this is because we do not consider networks with self-loops. In particular, if A is a completely uniform transition matrix that also includes self-loops (nonzero elements along the diagonal), then the only singularity occurs at $\beta = 0$.

Finally, as shown in Fig. S1A, unphysical values appearing in the A^* transition matrices occur across a wide range of β values (transition probabilites that are negative or greater than 1).

From these findings, we can conclude three main difficulties with the practical usage of the analytic solution to A^*: 1) the presence of self-loops, 2) the singular expressions that appear at lower values of β relevant to the human learning regime, and 3) the presence of unphysical values in the analytic A^* solutions.

The lattice graph exemplar. To understand how optimizing network learnability varies with the topology of the target network, we also consider the lattice graph presented in Fig. S2A. Unlike the modular graph, the lattice graph has only two structurally unique edges: edges within triangles and edges between triangles. Thus, an input network A_{in} can be fully described by only one free parameter, the weight λ_l of edges between triangles (orange), relative to the weight of edges within triangles (grey). Due to the reduction in the number of parameters, we are able to characterize gains in learnability in the lattice network while continuously varying both λ_l and β (Fig. S2B). Notably, learnability of the lattice graph increases markedly in the regime of low β and low λ_l. Moreover, across all values of β, we find that optimizing learnability requires de-emphasizing the edges between triangles (Fig. S2C).

We note that there are considerable differences between the optimal emphasis modulation strategies of modular and lattice graphs as a function of β. First, the profiles of the curves of optimal edge weight values are significantly different between the two networks (compare main text Fig. 1E and Fig. S2C). Specifically, the optimal edge weight curve of the lattice network shows an inflection point, whereas both optimal edge weight curves for the modular graph do not. Similar qualitative differences also appear between the optimal Kullback-Leibler divergence curves (compare main text Fig. 1F and Fig. S2D). These differences arise despite the fact that both networks were chosen to share the same local properties (all nodes have 4 neighbors), and thus have corresponding transition matrices with the same stationary distribution. This observation suggests that different networks require different approaches to maximize learnability, and that the efficacy of these approaches will differ by topology.

A Sierpiński graph exemplar. To assess whether the strategy of over-emphasizing edges within clusters and de-emphasizing those between clusters extends to larger networks with more complex community organization, we consider a modified version of the Sierpiński network with 3 hierarchical levels and 5 communities at each level. Specifically, the network was modified to include a sixth community at the highest level (Fig. S3*A*), allowing the graph to become 5-regular, and therefore allowing its transition network to be uniform. This network was chosen to assess how edges at various levels ought to be weighted to maximize learnability in networks with hierarchically modular organization. Despite containing 150 nodes, the network possesses a large degree of structural symmetry, and has only four unique edges: level-2 cross-cluster edges (λ_{cc}^2, orange), level-3 cross cluster edges (λ_{cc}^3, blue), boundary edges adjacent to level-2 cross cluster edges (λ_b^2, green), and boundary edges adjacent to level-3 cross-cluster edges (λ_b^3, grey). As before, we reduce the number of free parameters by 1 and fix $\lambda_b^3 = 1$.

Overall, we find that de-emphasizing both classes of cross-cluster edge weights is an effective strategy for optimizing the learnability of the Sierpiński network (Fig. S3*B,C*). However, there are slight differences in optimal edge weights between the level-2 and level-3 edges (Fig. S3*D*). In particular, we find that edges at the highest level of organization (level-3 edges) ought to be de-emphasized slightly more than level-2 edges. The efficacy of these optimization strategies scales similarly with β as in the case of the 15-node modular graph (compare main text Fig. 1*F* and Fig. S3*F*). The learned representations of the Sierpiński network with and without edge weight optimization are shown for $\beta = 0.05$ in Figs. S3*C* and S3*E*, respectively. These findings further suggest that the learnability of both modular and hierarchically modular networks can be substantially enhanced through the de-emphasis of cross-cluster edges, and the reinforcement of within-cluster edges. Moreover, by optimizing learnability, the learned representation of the hierarchically modular network maintains the fine-scale community structure (Fig. S3*E*). Interestingly, at low β values, the learned representation resulting from optimal edge weights strikes a trade-off between local and global features: it strongly captures the features of each of the small 5-node cliques, but poorly captures the hierarchical structure of the network. This pattern is likely a consequence of the fact that, at low β values, near-perfect learning is impossible, and thus an optimal weighting strategy for minimizing the Kullback-Leibler divergence would place emphasis on accurately learning the most commonly occurring substructure.

Watts-Strogatz networks. Our analysis of the lattice network (Fig. S2*A*) demonstrated that edges that do not contribute to the formation of small clusters or triangles should be de-emphasized in order to optimize network learnability. To assess this conclusion in a more general class of networks, we consider the optimization of learnability for Watts-Strogatz networks. Prior to any rewiring ($p = 0$), such networks begin as a ring-like lattice of nodes, with each node only having connections to its nearest neighbors in the ring (Fig. S4*C*). Given the density of local connections, these ring-like networks are highly clustered. When a small fraction of edges are then rewired, Watts-Strogatz networks maintain similar levels of clustering, but display markedly lower average path lengths (1, 2). In this regime, Watts-Strogatz networks can often be characterized by small-worldness, a concept relevant to a number of real-world networks including brain networks, language networks, and metabolic networks (3–7). Finally, in the limit of high rewiring $p = 1$, the structure of Watts-Strogatz networks is very similar to that of Erdős–Rényi networks. Motivated by previous work reporting that networks with high clustering coefficients are more learnable (8), we investigate the optimization of learnability in Watts-Strogatz networks at different rewiring probabilities p. In particular, we seek to determine whether the rewired edges, which deviate from the original highly-clustered ring network, ought to be de-emphasized when presented to human learners. In addition, we aim to identify whether the efficacy of strategies for optimizing network learnability depend on the emergence of small-world structure, which tends to appear for rewiring probabilities of $10^{-2} \leq p \leq 10^{-1}$ (9).

To investigate how rewired edges in Watts-Strogatz networks should be weighted to maximize network learnability, we consider the optimal weight λ_{nr} of rewired edges relative to non-rewired edges on the ring. For low rewiring probabilities ($p < 10^{-0.5}$), we find that network learnability is optimized by de-emphasizing rewired edges and over-emphasizing edges on the ring (Fig. S4*A*). Considering that the original lattice-like ring is highly clustered, and is therefore naturally easier to learn (8), these findings suggest that de-emphasizing areas of a network that do not contribute to clustering may be an effective general strategy for enhancing network learnability. This finding is consistent over the range $10^{-3} \leq \beta \leq 0.2$ of β values analyzed. In particular, in the limit $\beta \to 0$, the optimal non-ring edge weight approaches $\lambda_{nr} \to 0$ for nearly all rewiring probabilities p, whereas higher β values (more accurate learning) do not require such stark de-emphasis of non-ring edges. In addition, as the rewiring probability p approaches 1, the weight given to non-ring edges approaches 1. Given that highly-rewired Watts-Strogatz networks are equivalent to random Erdős–Rényi networks, it is reasonable that for high values of p, there is no distinction between ring and non-ring edges.

Interestingly, we also find that improvements in learnability resulting from tuning non-ring edge weights are most prominent at intermediate rewiring probabilities near $p \sim 10^{-1}$ (Fig. S4*B*). This finding suggests that the learnability of networks with small-world structure is significantly more optimizable when compared to highly ordered lattice-like networks or to highly disordered Erdős–Rényi networks.

Relaxing the symmetry constraint during optimization. To address the possibility that the symmetry constraints imposed during the network optimization process applied in the main text might explain observed results, we optimized the modular graph (Fig. 1*A*, main text) and the lattice graph (Fig. S2*A*) using the same scheme that was applied for the semantic networks. Specifically, the weights of structurally symmetric edges were not reduced to one parameter, and were instead allowed to be optimized independently as free parameters. As shown in Fig. S6, relaxation of the symmetry constraint of the network optimization process for the modular graph does not substantially change the efficacy of network optimization strategies as a function of β. In particular, the qualitative findings that distinguish modular graph optimizability from that of the semantic networks still remain; we still observe high optimizability near $\beta = 0$, which drops drastically at higher β. Nonetheless, there is a very small but nonzero gap between the two curves at low β, confirming that it is in principle, possible for a network with less symmetry than the target network A to be a more optimal choice for A^*.

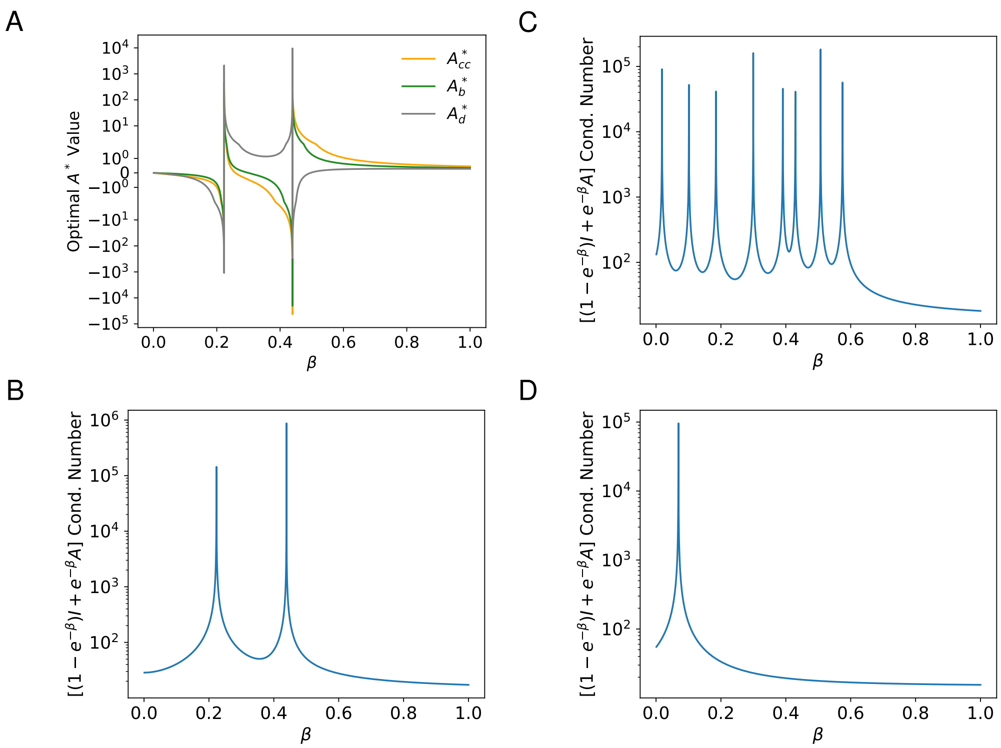

Fig. S1. Properties of the analytic solution to A^*. Here we show how properties of the analytic solutions to A^* depend on the learning accuracy β for different target networks A. *(A)* Values of entries of the analytic A^* transition matrix for the modular graph (Fig. 1*A*, main text)—corresponding to the cross-cluster, boundary, and deep edges—are shown for different values of β. The condition number of the argument to the matrix inverse versus β for *(B)* the modular graph, *(C)* a random, undirected graph with 15 nodes and 30 edges, and *(D)* the complete graph with 15 nodes.

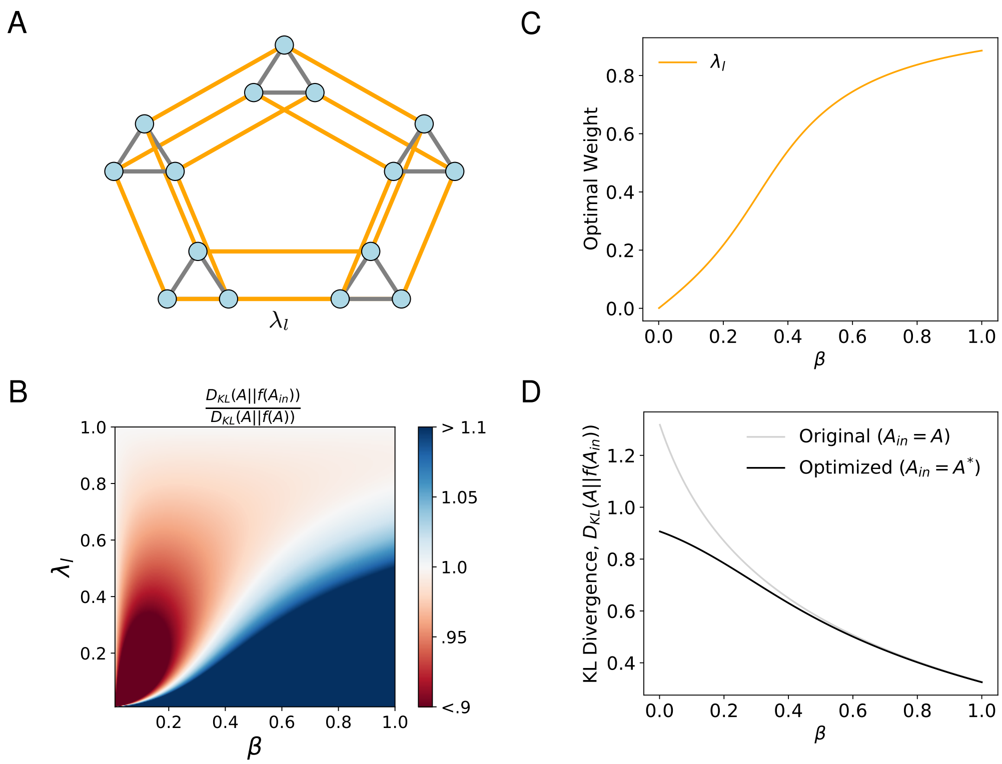

Fig. S2. Optimizing the learnability of a lattice network. *(A)* A lattice network with 15 nodes, each with degree $k_i = 4$, resulting in 30 edges. *(B)* Here we show the Kullback-Leibler divergence ratio (less than 1 indicates enhanced learnability) across a section of the λ_l, β parameter space. For increased contrast, the ratios have been truncated to the range $[0.9, 1.1]$. *(C)* The optimal edge weight λ_l for $0 < \beta < 1$. *(D)* The Kullback-Leibler divergence between the learned network and the true network for different values of β, with and without input network optimization.

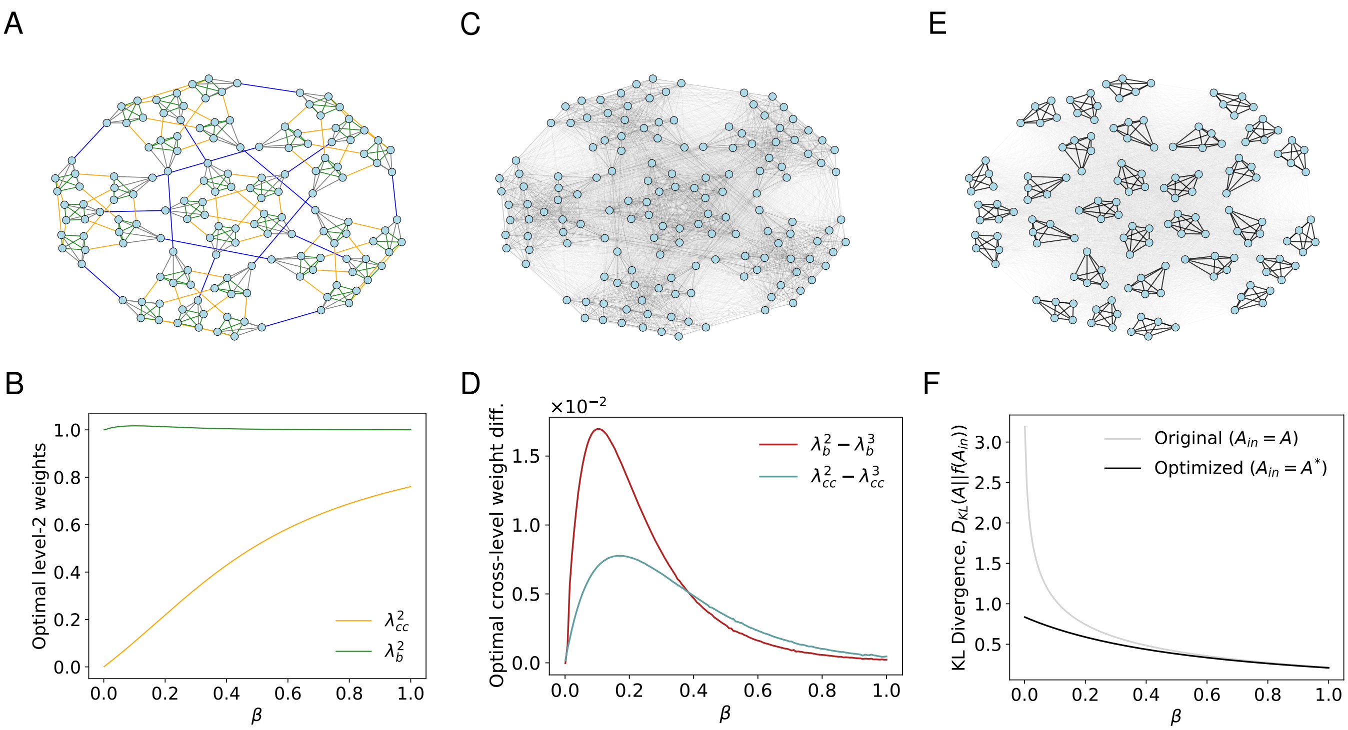

Fig. S3. Optimizing the learnability of a Sierpiński network. *(A)* The Sierpiński network S^3_5 with 3 levels, modified to have 6 communities at the final level. *(C, E)* The learned representations of the Sierpiński network at $\beta = 0.05$, both with *(E)* and without *(C)* input network optimization. *(B)* The optimal level-2 edge weights λ^2_{cc} and λ^2_b for $0 < \beta < 1$. *(D)* The differences $\lambda^2_b - \lambda^3_b$ and $\lambda^2_{cc} - \lambda^3_{cc}$ between optimal edge weights of levels 2 and 3, for $0 < \beta < 1$. *(F)* The Kullback-Leibler divergence between the learned network and the true network for different values of β, with and without input network optimization.

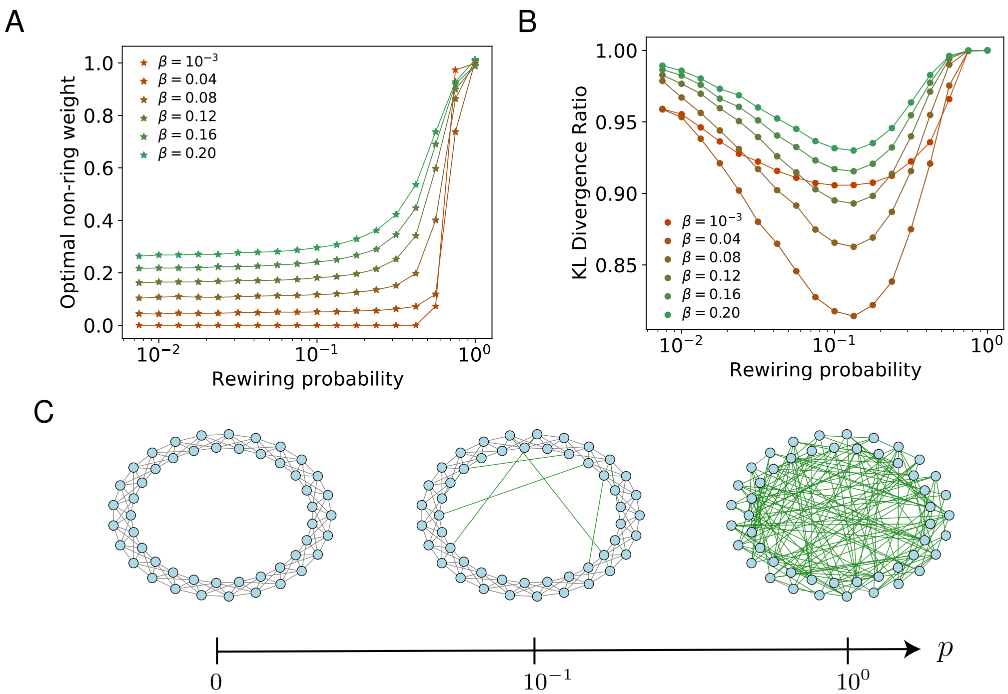

Fig. S4. Optimizing the learnability of small world networks. *(A)* The optimal non-ring edge weight λ_{nr} for enhancing learnability versus the rewiring probability p of a Watts-Strogatz network at different values of β. *(B)* The Kullback-Leibler divergence ratio $\frac{D_{KL}(A||f(A_{\text{in}}))}{D_{KL}(A||f(A))}$ achieved with optimal non-ring edge weights at different values of β. The findings reported in panels *(A,B)* represent results obtained for networks with $N = 200$ nodes and an average degree of $\langle k \rangle = 10$. Each curve is an average over the results from 25 generated networks. *(C)* A schematic demonstrating how the structure of Watts-Strogatz networks changes as the rewiring probability p increases. Non-ring edges are shown in green.

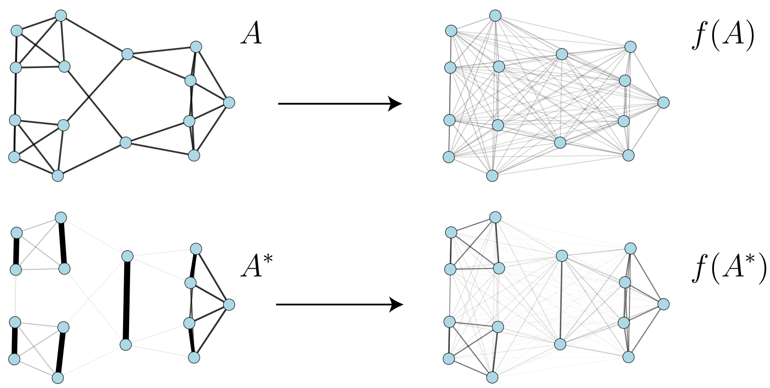

Fig. S5. Optimal emphasis modulation of an example network when considering nonexistent edges. Here we show the learned networks resulting from human learning of an example network (*top*), as well as from the example network optimized for learnability (*bottom*). The optimized network was determined with the addition of nonexistent edges as free parameters. Optimized and learned networks were both computed at $\beta = 0.05$.

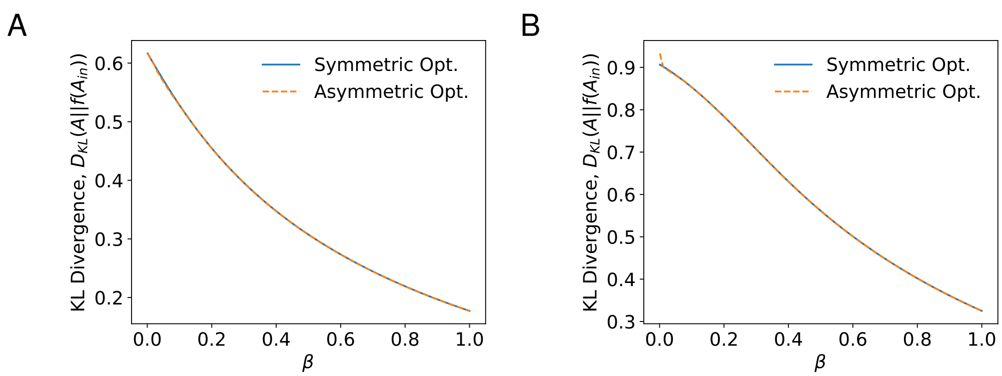

Fig. S6. Optimization of highly symmetric networks with relaxed symmetry constraints. Here we show the efficacy of network optimization strategies, with and without the symmetry constraint used in the main text. This is shown for *(A)* the modular graph (Fig. 1*A*, main text), and *(B)* the lattice graph (Fig. S2*A*.)

Axler	
Concept 1	Concept 2
real number	complex number
absolute value	complex conjugate
vector space	domain
real number	nonnegative
singular value	positive operator
nonconstant polynomial	complex coefficient
diagonal entry	arbitrary basis
nonconstant polynomial	factorization
division algorithm	polynomials

Peterson	
Concept 1	Concept 2
unique solution	initial value problem
subspace	trivial
inverse	generalize
matrix product	matrix multiplication
linear map	linear function
Gauss elimination	upper triangular form
Frobenius canonical form	similarity invariant
diagonal matrix	unitarily equivalent
diagonal entry	upper triangular

Bretscher	
Concept 1	Concept 2
linear	multiplicative operation
linear transformation	isomorphic
linear system	inconsistent
ellipse	unit circle
invertible	noninvertible matrix
dot product	orthogonality
transformation	partition
diagonalization	diagonalizable
rotation	sin cos

Greub	
Concept 1	Concept 2
injective	linear map
positive basis	induced orientation
factor space	differential operator
subalgebra	extension field
commutative	subalgebra
linear mapping	surjective
induced transformation	minimum polynomial
algebras	subalgebra
isomorphic	differential operator

Table S1. A sample of edges found to improve network learnability when strengthened. Listed are ten different concept pairs found in the semantic networks extracted from the linear algebra textbooks authored by Axler, Peterson, Bretscher, and Greub. These edges were selected from networks that were optimized at $\beta = 0.2$.

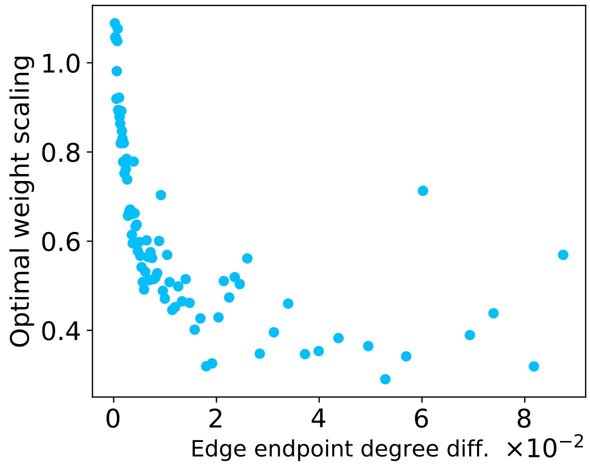

Fig. S7. Optimal edge scaling versus edge endpoint degree difference. The optimal edge weight scaling versus the absolute difference of the degrees of edges' endpoints, aggregated over all semantic networks for $\beta = 0.2$. Each datapoint represents an average over 500 edges binned by edgepoint degree difference.

References

1. MD Humphries, K Gurney, Network 'small-world-ness': A quantitative method for determining canonical network equivalence. *PLoS ONE* **3**, e0002051 (2008).
2. V Latora, M Marchiori, Efficient behavior of small-world networks. *Phys. Rev. Lett.* **87**, 198701 (2001).
3. DS Bassett, E Bullmore, Small-world brain networks. *The Neurosci.* **12**, 512–523 (2006).
4. RFi Cancho, RV Solé, The small world of human language. *Proc. Royal Soc. London. Ser. B: Biol. Sci.* **268**, 2261–2265 (2001).
5. A Wagner, DA Fell, The small world inside large metabolic networks. *Proc. Royal Soc. London. Ser. B: Biol. Sci.* **268**, 1803–1810 (2001).
6. QK Telesford, KE Joyce, S Hayasaka, JH Burdette, PJ Laurienti, The ubiquity of small-world networks. *Brain Connect.* **1**, 367–375 (2011).
7. L de Arcangelis, HJ Herrmann, Self-organized criticality on small world networks. *Phys. A: Stat. Mech. its Appl.* **308**, 545–549 (2002).
8. CW Lynn, L Papadopoulos, AE Kahn, DS Bassett, Human information processing in complex networks. *Nat. Phys.* **16**, 965–973 (2020).
9. DJ Watts, SH Strogatz, Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).