Flow of Control in Networks
Jukka-Pekka "JP" Onnela
Science **343**, 1325 (2014);
DOI: 10.1126/science.1251746

for personalized therapies for treating EBV will continue to grow.

To date, EBV is estimated to be responsible for ~200,000 cancers worldwide (4). The U.S. National Institutes of Health (NIH) recently called for a new initiative to reduce global cancer incidence, with EBV among the top candidates for future advances (14). Further clinical testing of the gp350 vaccine, as well as development of second-generation vaccines and diagnostics to measure vaccine efficacy and cancer risk factors, have been recommended by an NIH-sponsored panel (14, 15). Among the new generation of vaccines will be those that treat latently infected individuals with existing EBV-driven cancers as well as those that are at high risk for developing EBV-associated disease (e.g., solid organ transplant recipients). Any vaccine that stimulates a strong and selective T cell response to EBV-positive tumor cells is likely to provide protection and therapeutic benefit. It is to be hoped that the successes of HPV and hepatitis B virus vaccination programs will encourage new and ongoing efforts to find a suitable immunological or pharmacological treatment for EBV and associated disease. An efficacious antiviral would also provide the final confirmation that EBV is indeed a tumor-causing virus.

References

1. M. A. Epstein, B. G. Achong, Y. M. Barr, *Lancet* **283**, 702 (1964).
2. P. K. Pattengale, R. W. Smith, P. Gerber, *Lancet* **302**, 93 (1973).
3. Y. Chang *et al.*, *Science* **266**, 1865 (1994).
4. D. M. Parkin, *Int. J. Cancer* **118**, 3030 (2006).
5. R. Baer *et al.*, *Nature* **310**, 207 (1984).
6. M. H. Tsai *et al.*, *Cell Rep.* **5**, 458 (2013).
7. L. S. Young, A. B. Rickinson, *Nat. Rev. Cancer* **4**, 757 (2004).
8. N. Raab-Traub, *Curr. Opin. Virol.* **2**, 453 (2012).
9. D. A. Thorley-Lawson, M. J. Allday, *Nat. Rev. Microbiol.* **6**, 913 (2008).
10. J. T. Schiller, P. Davies, *Nat. Rev. Microbiol.* **2**, 343 (2004).
11. E. M. Sokal *et al.*, *J. Infect. Dis.* **196**, 1749 (2007).
12. C. M. Bollard, C. M. Rooney, H. E. Heslop, *Nat. Rev. Clin. Oncol.* **9**, 510 (2012).
13. D. H. Dreyfus, *Autoimmun. Rev.* **11**, 88 (2011).
14. H. Varmus, H. S. Kumar, *Sci. Transl. Med.* **5**, 175cm172 (2013).
15. J. I. Cohen, A. S. Fauci, H. Varmus, G. J. Nabel, *Sci. Transl. Med.* **3**, 107fs107 (2011).

PHYSICS

Flow of Control in Networks

A simple framework allows the classification of complex networks based on the flow of control.

Jukka-Pekka "JP" Onnela

Many complex systems can be viewed as networks, in which nodes represent system elements and edges correspond to interactions between those elements. In such networks, a subset of nodes—the driver nodes—can yield control of the entire network when they are driven by external signals (1–3). However, to control a system, one must know not only what parts need to be controlled but also why these particular parts need to be controlled. On page 1373 of this issue, Ruths and Ruths (4) put forward an elegant framework that elucidates the origin of control in networks. The framework divides nodes into three categories based on how they affect the flow of control in networks. This approach gives rise to control profiles that the authors use to classify a host of empirical and synthetic networks. The results suggest that networks from different domains but in the same category may be more similar to one another than previously thought.

To understand the idea of network control, consider a dynamical process that unfolds on a network, such as an airport transportation network, where the state of a node encodes the number of passengers at that node. If a directed edge leads from node A to node B, then the state of node A influences the state of node B. The system is controllable if it can be driven from an arbitrary initial state to a specified final state in a finite time. In the absence of loops and cycles, each node can control at most one of its neighbors (5). Together, these node-to-neighbor couplings give rise to disjoint directed paths of control in the network; each path, or stem, needs its own independent control. If we can identify the stems, we can control the network.

The standard way to find the set of controllable nodes is to use the maximum matching algorithm (6), where "maximum matching" refers to the maximum set of edges that do not share start or end nodes (1). A node is matched if an edge in the maximum matching points to it; otherwise, it is unmatched and needs its own independent control. If all stems were simple directed paths (see the figure, panel A), the number of independently controlled nodes would equal the number of source nodes, which are nodes with no incoming edges.
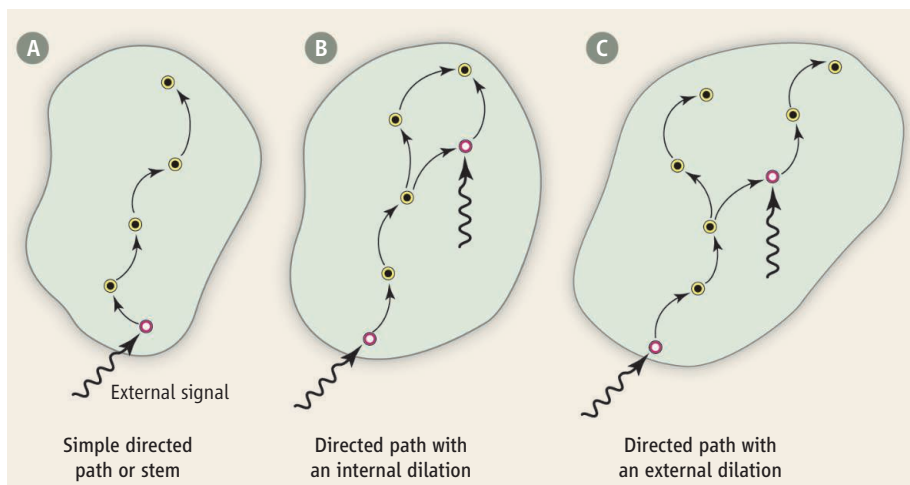
Control-inducing structures in networks. In a simple directed path or stem (**A**), only the source node needs to be controlled. If a stem branches, either internal dilations (**B**) or external dilations (**C**) are introduced, both of which increase the number of nodes that need independent control. This is because each node can control at most one of its neighbors and at a dilation a node has more than one neighbor. In panel C, the number of nodes that need independent control coincides with the number of sink nodes (nodes with no outgoing edges); this is not the case in panel B. Ruths and Ruths show that each control-inducing structure can be classified as a source node, an internal dilation, or an external dilation, and that the control profile of a network is the relative frequency with which each type occurs in the network.

Figure panel labels: **A** Simple directed path or stem — External signal; **B** Directed path with an internal dilation; **C** Directed path with an external dilation

Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA. E-mail: onnela@hsph.harvard.edu

When a stem branches into two or more paths, a dilation occurs, increasing the number of nodes that need to be controlled. Ruths and Ruths distinguish between two types of dilations. In an internal dilation, a path splits into two paths that soon rejoin, much like two lanes of traffic that momentarily split only to merge again soon after (see the figure, panel B). In an external dilation, a path splits without subsequently rejoining, like entering a T-intersection with two dead end streets (see the figure, panel C). An internal dilation does not increase the number of sink nodes, whereas an external one does, by one for each branching.

In the absence of internal dilations, only source nodes and external dilations need to be controlled. When internal dilations are present, the sum of the source nodes and surplus sink nodes gives a lower bound on the number of nodes that need control. In most networks, the number of internal dilations is relatively small, and the bound obtained as the sum of source and surplus sink nodes is therefore fairly tight.

Previous studies have categorized nodes and edges in control structures by considering whether a node always, never, or sometimes needs its independent control (3) and by investigating whether an edge is always, never, or sometimes part of the control structure (1). The degree distribution of a network (that is, the distribution of the number of incoming and outgoing edges of each node) correlates strongly with the number of required independent controls (1). Ruths and Ruths now show that the special role of source and sink nodes in control structures leads to a simple causal link between degree sequence and the size of the driver node set.

Based on the three types of structures that induce the need for independent control (see the figure), real-world and synthetic networks can be profiled by designating all nodes that need to be controlled into source nodes, internal dilations, and external dilations. The control profile of a given network is simply the fraction of nodes in each of these three categories. By investigating a range of directed real-world and model networks, Ruths and Ruths find evidence for three categories of networks, which they call source-dominated, internal-dilation–dominated, and external-dilation–dominated networks.

The authors argue that source-dominated networks, such as neural and social networks, allow relatively uncorrelated behavior across their agents and are thus suitable to distributed processing. Internal-dilation–dominated networks, such as food networks and airport interconnectivity networks, are mostly closed systems and obey some type of conservation laws. Finally, external-dilation–dominated networks with their surplus sink nodes yield correlated behaviors across their agents that are downstream from a common source, as exemplified by trust hierarchies and transcriptional systems.

The framework appears to assume that there are no intrinsic nodal dynamics. If self-loops (directed edges that connect a node to itself) were present at each node, such that the future state of a node depended on its own current state, all nodes could be driven by a single external control, given that loops and cycles are self-regulating structures (7). However, as the authors point out, if the time scale of interactions with network neighbors is much faster than that of the intrinsic dynamics, it is reasonable not to consider intrinsic nodal dynamics, especially for making comparisons across different types of networks.

Some studied networks will contain false positives and false negatives: The network representation of the system as obtained from empirical data will contain some edges that are not present in the real system, and will not contain some other edges that are present in the real system. Ruths and Ruths' framework may be particularly sensitive to false positives and false negatives, because they could change the classification of nodes in a non-trivial manner. Furthermore, the data collection processes that introduce false positives and false negatives may be different across different network types. The impact of false negatives could be investigated by subsampling observed networks to determine how sensitive the control profile of a given network and its subsequent classification are to different forms of sampling.

The framework put forward by Ruths and Ruths is elegant in its simplicity and helps us to understand and interpret previous findings. It also points to interesting avenues for future work, such as the development of more realistic models of directed networks in the context of network control. Going forward, the framework provides a useful tool for moving from passively observing networked systems to pursuing control over them.

References
1. Y.-Y. Liu, J.-J. Slotine, A.-L. Barabási, *Nature* **473**, 167 (2011).
2. M. Pósfai, Y.-Y. Liu, J.-J. Slotine, A.-L. Barabási, *Sci. Rep.* **3**, 1067 (2013).
3. T. Jia *et al.*, *Nat. Commun.* **4**, 1 (2013).
4. J. Ruths, D. Ruths, *Science* **343**, 1373 (2014).
5. C.-T. Lin, *IEEE Trans. Automat. Contr.* **19**, 201 (1974).
6. J. E. Hopcroft, R. M. Karp, *SIAM J. Comput.* **2**, 225 (1973).
7. N. J. Cowan, E. J. Chastain, D. A. Vilhena, J. S. Freudenberg, C. T. Bergstrom, *PLOS ONE* **7**, e38398 (2012).

DEVELOPMENT

Counting the Ways to Decode Dynamic Signals

Roy Wollman

The social amoeba *Dictyostelium discoideum* decodes an oscillatory signal to measure time.

The role of biological signaling networks is to reliably transmit specific information about the extracellular environment to multiple intracellular downstream effectors, allowing the cell to adjust its physiological state to changing conditions. One mechanism that cells use to enhance the performance of signaling networks is the temporal modulation, or dynamics, of the transmitted signals (1–4). The key role that modulating temporal activity of the signal plays in information transmission makes

Cell and Developmental Biology Section, Division of Biological Sciences, and Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093, USA, and San Diego Center for Systems Biology, La Jolla, CA 92093, USA. E-mail: rwollman@ucsd.edu

signaling dynamics an attractive target for therapeutic approaches that interfere with the transmission of specific types of information through the network (5). The diversity of temporal modulation strategies seen in various signaling networks suggests that there is no single optimal strategy for making use of dynamic information. Therefore, to uncover the benefits of temporal modulation strategies, it is important to understand how the suitability of each type of signaling dynamics is matched to the nature of the particular information that is being transmitted. Some types of information are transmitted through frequency-modulated signals, whereas other types are transmitted through modulation of signal amplitude or duration.