



Physarum Learner: A bio-inspired way of learning structure from data



T. Schön^{a,b}, M. Stetter^b, A.M. Tomé^c, C.G. Puntonet^d, E.W. Lang^{a,*}

^a CIML Lab, Dept. Biophysics, University of Regensburg, Germany

^b Dept. Bioinformatics, University of Applied Science Weihenstephan-Triesdorf, Germany

^c DETI – IEETA, Universidade de Aveiro, Portugal

^d Dept. Architecture and Technology of Computers, Universidad Granada, Spain

ARTICLE INFO

Keywords:

Physarum Learner
Structure learning
Bayesian Network

ABSTRACT

A novel *Score-based Physarum Learner* algorithm for learning *Bayesian Network* structure from data is introduced and shown to outperform common score based structure learning algorithms for some benchmark data sets. The *Score-based Physarum Learner* first initializes a fully connected *Physarum-Maze* with random conductances. In each *Physarum Solver* iteration, the source and sink nodes are changed randomly, and the conductances are updated. Connections exceeding a predefined conductance threshold are considered as *Bayesian Network* edges, and the score of the connected nodes are examined in both directions. A positive or negative feedback is given to the edge conductance based on the calculated scores. Due to randomness in selecting connections for evaluation, an ensemble of *Score-based Physarum Learner* is used to build the final *Bayesian Network* structure.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Bioinformatics intends to help solving biological problems with the help of computer science. In favorable cases, however, it is just the other way round and nature offers solutions for algorithmic problems as is the case with the slime mold *Physarum polycephalum*. The latter recently has emerged as a fascinating example of biological computation through morphogenesis. Through its growth process, this single cell organism was shown to be able to solve various minimum cost flow problems (Nakagaki, Yamada, & Toth, 2000). Recently, a number of bio-inspired optimization methods based on the Physarum Solver model proposed by Tero, Kobayashi, and Nakagaki (2006) have been proposed and applied to various optimization problems. Following we provide a short survey of the current literature about the subject and motivate our own research in this field.

1.1. Finding shortest paths

Nakagaki et al. (2000) convincingly demonstrated how the slime mold *P. polycephalum* can indeed find the shortest path between two food sources in a given maze. In an environment with distributed food sources, the plasmodium forms a network of protoplasmic tubes connecting these food sources. This network thus

represents a planar graph with food sources as nodes and tubes as edges. Nakagaki, Yamada, and Tóth (2001, 2007) demonstrated that the topology of the network optimizes harvesting on scattered food sources and renders the flow and transport of intracellular components more efficient. Consequently, research efforts in the following years focused on examining the strategy used by Physarum to understand network formation (Niizatoa, Shirakawab, & Gunjia, 2010) and the adaptive dynamics of its transport network (Nakagaki et al., 2001; Nakagaki, 2001; Nakagaki, Yamada, & Hara, 2004). Indeed, learning how biological systems solve problems (Nakagaki, Tero, Kobayashi, Ohnishi, & Miyaji, 2009) could help designing new methods of computation. Information processing in simple cellular organisms thus is interesting to learn about solving combinatorial optimization problems, such as solving a maze or a shortest paths network problem.

Tero et al. (2006, 2007) were the first to propose a mathematical model for the underlying transport network based on hydrodynamics. They showed that their model, (called *Physarum Solver*) (PhySol), can solve the maze introduced by Nakagaki et al. (2000) in the same manner as the biological slime mold does. Further, Tero, Yumiki, Kobayashi, Saigusa, and Nakagaki (2008) adapted the mathematical model to problems with more than two food sources by randomly changing the source and sink node within the network at each iteration of the solving algorithm. Last but not least, Nakagaki et al. (2007) described a mathematical model for an adaptive – tube network which was in good agreement with a risk minimizing behavior of a real slime mold whose

* Corresponding author. Tel.: +49 941 943 2599; fax: +49 941 943 2479.

E-mail address: elmar.lang@ur.de (E.W. Lang).

embedding maze was exposed locally to light (see also Fricker, Boddy, Nakagaki, & Bebbler (2009)). Roughly at the same time, Miyaji and Ohnishi (2007) proved that Tero's model with linear dynamics converges to an optimal solution in any instance of a planar graph. Miyaji and Ohnishi (2008) later proved with mathematical rigor the shortest path-finding ability of the Physarum Solver model on a graph embedded in a two-dimensional manifold with Riemannian metric, i. e. a Riemannian surface. They showed that the equilibrium point corresponding to the shortest path is globally asymptotically stable for the Physarum Solver on a Riemannian surface. They also extended their proof to a surface composed of a finite number of Riemannian surfaces pasted with one another continuously. A detailed mathematical analysis and convergence study by Miyaji and Ohnishi (2008) and Brummitt et al. (2010) further showed that in ring-shaped and Wheatstone bridge-shaped networks the global asymptotically stable equilibrium point of the model corresponds to the shortest path connecting two special points on the network in cases where the shortest paths are determined uniquely. Miyaji, Ohnishi, Tero, and Nakagaki (2008) also explained why in laboratory experiments with networks encompassing a double-edge, *P. polycephalum* sometimes yields erroneous results. Similarly, Ito, Johansson, Nakagaki, and Tero (2011) discussed shortest path finding of the Physarum Solver on a di-graph $G = (V, E)$, where E denotes edges and V vertices or nodes of the graph. They proved that the volumetric flow rate $Q(t)$, following Fick's law for stationary flows, and the related conductance $D(t)$, considered a time-dependent state variable, converge with an exponential rate to a positive flow minimizing the total length or cost $L(Q) = \sum_{ij} L_{ij} Q_{ij}$. The limit flow has full support on an optimal set \mathcal{H} which often forms a spanning subgraph.

Bonifaci, Mehlhorn, and Varma (2012) proved that, considering the Physarum Solver model, the mass of the mold will converge to the shortest path between two specified nodes of the network that the mold lies on, independently of the initial mass distribution or the structure of the network, thus relieving the restriction to planar graphs. In addition, Bonifaci (2013) recently presented a short proof of a standard model for the *P. polycephalum* slime mold which correctly computes the shortest path in an undirected weighted graph. Furthermore, Bechetti, Bonifaci, Dirnberger, Karrenbauer, and Mehlhorn (2013) proved that a discretization (Euler integration) of the coupled system of differential equations, as proposed by Tero et al. (2007), computes a $(1 + \epsilon)$ -approximation of the shortest path in $O(mL(\log(n) + \log(L))\epsilon^{-3})$ iterations. Arithmetics were performed on numbers of $O(\log(nL\epsilon^{-1}))$ bits, where n and m represent the number of nodes and edges of the graph, respectively, and L measures the largest length of an edge. In addition, the authors also obtained two results for a directed Physarum model, proposed by Ito et al. (2011), concerning convergence in the general, nonuniform case and convergence and complexity bounds for the discretization of the uniform case (see also an enlightening discussion by Richard Lipton at <http://rjlipton.wordpress.com>).

In an attempt to understand locomotion and morphogenesis of the slime mold, Gunjia, Shirakawab, Niizatoa, Yamachiyoa, and Tania (2011) recently presented an adaptive and robust vacant-particle transportation model. They could emulate the network formed by *P. polycephalum* and reproduce morphogenic patterns quantitatively. This corroborates that both, the slime mold as well as the model, balance the exploration-exploitation trade-off inherent to many bio-inspired optimization algorithms. Similarly, Johannson and Zou (2012) analyzed a mathematical model of the amoeboid growth dynamics. They showed how to encode general linear programming (LP) problems as instances of PhysSol and proved that under the growth dynamics considered, PhysSol is guaranteed to converge to the optimal solution of the LP.

1.2. Network optimization problems

Tero et al. (2010) recently discussed rules for biologically inspired adaptive network design. Robust network performance involves a complex trade-off between cost, transport efficiency and fault tolerance. In addition, they develop without centralized control and may represent a readily scalable solution for growing networks in general. Zhang, Zhang, Hu, Deng, and Mahadevan (2013) discussed network optimization problems, more specifically constrained shortest path (CSP) problems. The authors combine an adaptive amoeba algorithm with a Lagrangian relaxation algorithm to solve the CSP problem. Within a two-step process they first solve the shortest path problem (SPP) in a directed network, and in a second step the adaptive amoeba model is combined with a Lagrangian relaxation method to solve the CSP problem. Recently, Zhang and Liu et al. (2013) extended their adaptive amoeba algorithm to address shortest path tree (SPT) problems in dynamic graphs. In the latter edge weight updates can result in either edge weight increases or decreases or a mix of both. Existing work analyzes the nodes influenced by edge weight updates and recompute these affected vertices. In large networks, this approach becomes very complex and computationally costly, however. The adaptive amoeba algorithm can recognize edge weight updates and the affected vertices and reconstruct them spontaneously. To evaluate the proposed adaptive amoeba algorithm, the authors compared it with the Label Setting algorithm and Bellman-Ford algorithm.

Besides solving shortest path finding problems, recent work also considers characteristics of the underlying networks. Zhang, Zhang, Wei, and Deng (2012) studied centrality measures in weighted networks based on an amoeboid algorithm. For weighted networks, the authors proposed a new centrality measure based on an amoeboid algorithm, which they called *Physarum centrality*. In addition to shortest paths, the measure also includes contributions from competing paths. With PhysSol, each edge is endowed with a flux and edges on shorter paths are of greater flux. By defining flux as criticality of edges, Physarum centrality of a node is simply calculated by summing up the criticality of edges attached to it.

Biological systems, like the slime mold *P. polycephalum*, are composed of unreliable components which self-organize into balanced systems with respect to efficiency and robustness. Becker (2011) therefore studied slime mold inspired algorithms for constructing fault tolerant connection networks. Given that transport network solutions learnt with PhysSol suffer from a strong dependence on a single parameter, the author presents an agent-based approach for constructing fault tolerant connection networks using PhysSol. The agent-based simulation reproduced the variance in the behavior of the natural slime mold much better. However, analysis of the cost-benefit ratio of bio-inspired network construction lead the author to conclude that considering classical efficient computational algorithms for the problem of constructing minimal fault tolerant networks might be worthwhile. Following similar ideas, Li, Torres, Thomas, Rossi, and Shen (2011) considered slime mold inspired routing protocols for wireless sensor networks. Taking both efficiency and robustness into account, the authors exploit two different mechanisms to design localized routing protocols for wireless sensor networks (WSNs). Inspired by the slime mold's mechanism of path growth, the authors construct path growth routing protocols by treating data sources and sinks as singular potentials thereby establishing routes from the sinks to all the data sources. Similarly, in the plasmodium, path evolution is driven by a non-linear feedback mechanism of flux intensity onto tube diameter. The authors thus adapted tube dynamics for the purpose of a path evolution routing protocol and identify one key adaptation parameter to adjust the trade-off between efficiency and robustness of network routes.

Download English Version:

<https://daneshyari.com/en/article/383752>

Download Persian Version:

<https://daneshyari.com/article/383752>

[Daneshyari.com](https://daneshyari.com)