

Kelk, Steven; Fischer, Mareike; Moulton, Vincent; Wu, Taoyang

Reduction rules for the maximum parsimony distance on phylogenetic trees. (English)

Zbl 1348.68068

Theor. Comput. Sci. 646, 1-15 (2016).

Summary: In phylogenetics, distances are often used to measure the incongruence between a pair of phylogenetic trees that are reconstructed by different methods or using different regions of genome. Motivated by the maximum parsimony principle in tree inference, we recently introduced the maximum parsimony (MP) distance, which enjoys various attractive properties due to its connection with several other well-known tree distances, such as TBR and SPR. Here we show that computing the MP distance between two trees, a NP-hard problem in general, is fixed parameter tractable in terms of the TBR distance between the tree pair. Our approach is based on two reduction rules – the chain reduction and the subtree reduction – that are widely used in computing TBR and SPR distances. More precisely, we show that reducing chains to length 4 (but not shorter) preserves the MP distance. In addition, we describe a generalization of the subtree reduction which allows the pendant subtrees to be rooted in different places, and show that this still preserves the MP distance. On a slightly different note we also show that Monadic Second Order Logic (MSOL), posited over an auxiliary graph structure known as the display graph (obtained by merging the two trees at their leaves), can be used to obtain an alternative proof that computation of MP distance is fixed parameter tractable in terms of TBR-distance. We conclude with an extended discussion in which we focus on similarities and differences between MP distance and TBR distance and present a number of open problems. One particularly intriguing question, emerging from the MSOL formulation, is whether two trees with bounded MP distance induce display graphs of bounded treewidth.

MSC:

68Q25 Analysis of algorithms and problem complexity
05C05 Trees
92D15 Problems related to evolution

Cited in 5 Documents

Keywords:

phylogenetics; parsimony; fixed parameter tractability; chain; incongruence; treewidth

Software:

ComputeTW; SPRSupertrees

Full Text: [DOI](#)

References:

- [1] Allen, B.; Steel, M., Subtree transfer operations and their induced metrics on evolutionary trees, *Ann. Comb.*, 5, 1, 1-15, (2001) · [Zbl 0978.05023](#)
- [2] Arnborg, S.; Lagergren, J.; Seese, D., Easy problems for tree-decomposable graphs, *J. Algorithms*, 12, 308-340, (1991) · [Zbl 0734.68073](#)
- [3] Bodlaender, H. L., A tourist guide through treewidth, *Acta Cybernet.*, 11, 1-2, 1, (1994) · [Zbl 0804.68101](#)
- [4] Bodlaender, H. L.; Koster, A. M.C. A., Treewidth computations I. upper bounds, *Inform. and Comput.*, 208, 3, 259-275, (2010) · [Zbl 1186.68328](#)
- [5] Boes, O.; Fischer, M.; Kelk, S., A linear bound on the number of states in optimal convex characters for maximum parsimony distance, *IEEE/ACM Trans. Comput. Biol. Bioinform.*, (2016), in press, arxiv preprint
- [6] Bordewich, M.; Semple, C., Computing the hybridization number of two phylogenetic trees is fixed-parameter tractable, *IEEE/ACM Trans. Comput. Biol. Bioinform.*, 4, 458-466, (2007)
- [7] Bryant, D.; Lagergren, J., Compatibility of unrooted phylogenetic trees is FPT, *Theoret. Comput. Sci.*, 351, 3, 296-302, (2006) · [Zbl 1086.68097](#)
- [8] Chen, J.; Fan, J-H.; Sze, S-H., Parameterized and approximation algorithms for maximum agreement forest in multifurcating

- trees, *Theoret. Comput. Sci.*, 562, 496-512, (2015) · [Zbl 1303.68154](#)
- [9] Chuzhoy, J., Excluded grid theorem: improved and simplified, (Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, (2015), ACM), 645-654 · [Zbl 1321.05248](#)
- [10] Courcelle, B., The monadic second-order logic of graphs. I. recognizable sets of finite graphs, *Inform. and Comput.*, 85, 1, 12-75, (1990) · [Zbl 0722.03008](#)
- [11] Diestel, R., *Graph theory*, (2010), Springer-Verlag, GmbH & Company KG Berlin and Heidelberg · [Zbl 1204.05001](#)
- [12] Downey, R.; Fellows, M., *Fundamentals of parameterized complexity*, vol. 4, (2013), Springer · [Zbl 1358.68006](#)
- [13] Felsenstein, J., *Inferring phylogenies*, (2004), Sinauer Associates, Incorporated
- [14] Fischer, M.; Kelk, S., On the maximum parsimony distance between phylogenetic trees, *Ann. Comb.*, 20, 1, 87-113, (2016) · [Zbl 1332.05043](#)
- [15] Fitch, W., Toward defining the course of evolution: minimum change for a specific tree topology, *Syst. Zool.*, 20, 4, 406-416, (1971)
- [16] Grigoriev, A.; Kelk, S.; Lekić, N., On low treewidth graphs and supertrees, *J. Graph Algorithms Appl.*, 19, 1, 325, (2016) · [Zbl 1325.05168](#)
- [17] Hickey, G.; Dehne, F.; Rau-Chaplin, A.; Blouin, C., SPR distance computation for unrooted trees, *Evol. Bioinform.*, 4, 17-27, (2008)
- [18] Kelk, S., A note on convex characters and Fibonacci numbers, (2015), submitted for publication
- [19] Kelk, S.; Fischer, M., On the complexity of computing MP distance between binary phylogenetic trees, *Ann. Comb.*, (2016), in press, arxiv preprint
- [20] Kelk, S.; van Iersel, L. J.J.; Scornavacca, C.; Weller, M., Phylogenetic incongruence through the Lens of monadic second order logic, *J. Graph Algorithms Appl.*, 20, 2, 189-215, (2016) · [Zbl 1331.05210](#)
- [21] Moulton, V.; Wu, T., A parsimony-based metric for phylogenetic trees, *Adv. in Appl. Math.*, 66, 22-45, (2015) · [Zbl 1315.05034](#)
- [22] Robertson, N.; Seymour, P.; Thomas, R., Quickly excluding a planar graph, *J. Combin. Theory Ser. B*, 62, 2, 323-348, (1994) · [Zbl 0807.05023](#)
- [23] Robertson, N.; Seymour, P. D., Graph minors. V. excluding a planar graph, *J. Combin. Theory Ser. B*, 41, 1, 92-114, (1986) · [Zbl 0598.05055](#)
- [24] Steel, M.; Penny, D., Distributions of tree comparison metrics-some new results, *Syst. Biol.*, 126-141, (1993)
- [25] Vakati, S.; Fernández-Baca, D., Compatibility, incompatibility, tree-width, and forbidden phylogenetic minors, (LAGOS'15 - VIII Latin-American Algorithms, Graphs and Optimization Symposium, *Electron. Notes Discrete Math.*, vol. 50, (2015)), 337-342 · [Zbl 1347.05234](#)
- [26] Whidden, C.; Zeh, N.; Beiko, R., Supertrees based on the subtree prune-and-regraft distance, *Syst. Biol.*, 63, 4, 566-581, (2014)
- [27] Wu, T.; Moulton, V.; Steel, M., Refining phylogenetic trees given additional data: an algorithm based on parsimony, *IEEE/ACM Trans. Comput. Biol. Bioinform.*, 6, 1, 118-125, (2009)
- [28] Yang, J.; Li, J.; Dong, L.; Grünwald, S., Analysis on the reconstruction accuracy of the Fitch method for inferring ancestral states, *BMC Bioinformatics*, 12, 1, 18, (2011)

This reference list is based on information provided by the publisher or from digital mathematics libraries. Its items are heuristically matched to zbMATH identifiers and may contain data conversion errors. It attempts to reflect the references listed in the original paper as accurately as possible without claiming the completeness or perfect precision of the matching.