

Motivation: Hierarchy is everywhere. How can we model it using symbolic regression?

Given a hidden hierarchical system with N state variables (v_1, v_2, \dots, v_N) and K inputs (s_1, s_2, \dots, s_K)

Symbolic regression will return a set of flat expressions:

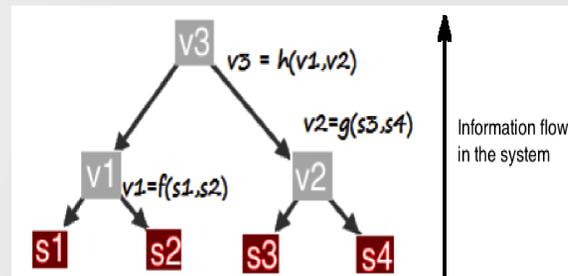
$$v_3 = v_2 * (s_1 - s_2)$$

$$v_2 = s_3 * s_4$$

$$v_1 = s_1 - s_2$$

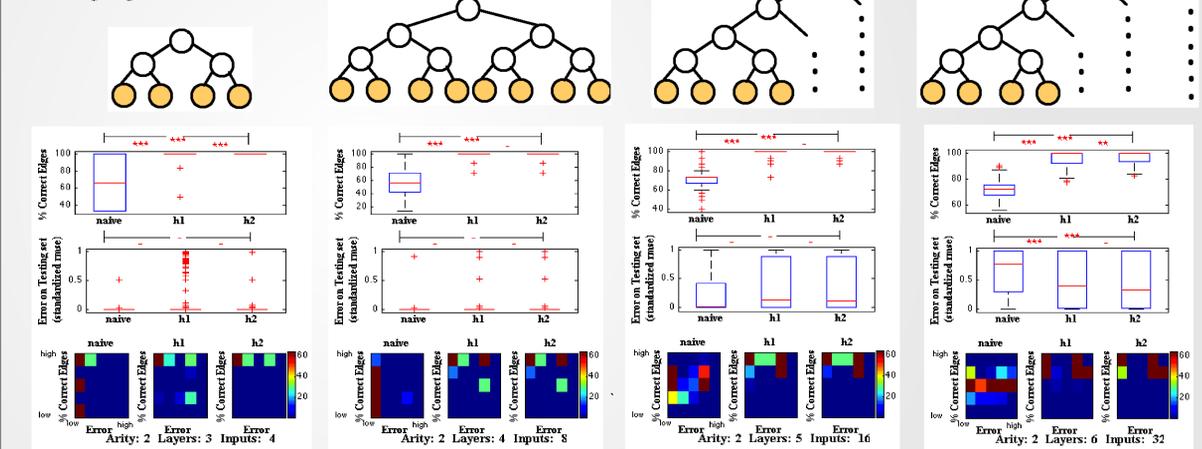
More variables \rightarrow Less intelligible models

Our method: Build variable (functional) dependency graphs based on symbolic regression

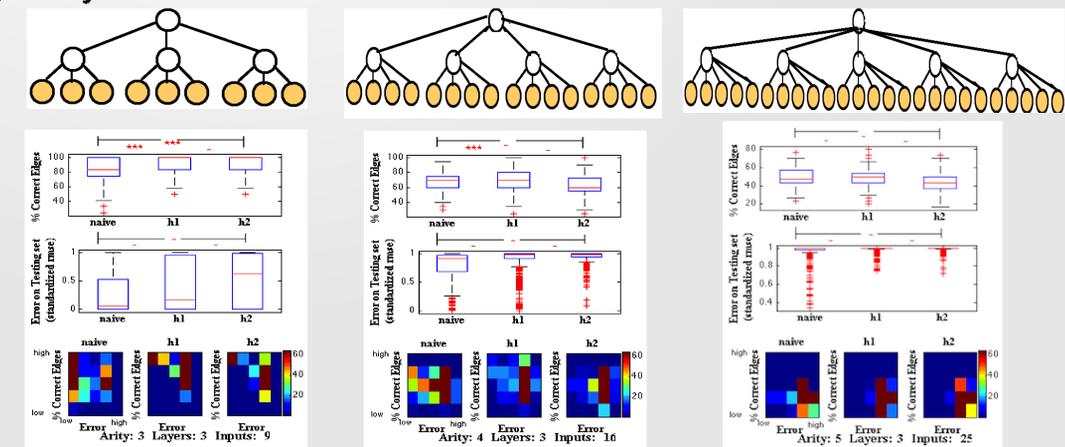


Experimental results (10 min. run-time budget)

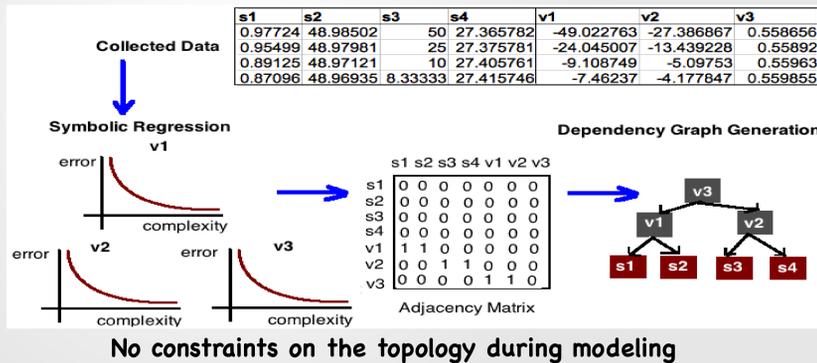
Binary Systems



Three-layered Systems



From data to hierarchy: the naive approach



Our hypotheses:

- the naive approach does not guarantee discovery of hierarchy
- low prediction error does not necessarily mean that the correct topology is captured
- If you want to capture hierarchy, you need to explicitly seek hierarchy

The hierarchical approach

- bottom-up approach of extracting functional relationships in the data in an iterative manner

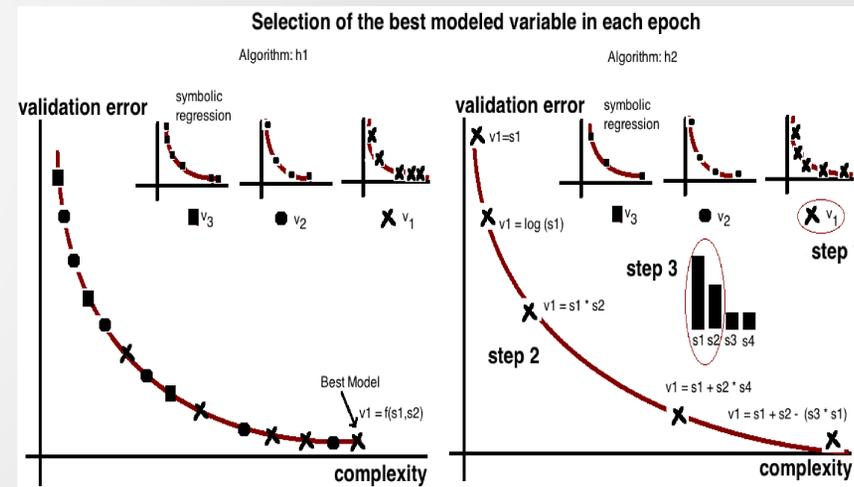
- at the beginning, only the stimuli (input) variables are included in the set of independent variables

Until all state variables are modeled:

- at each epoch, the state variable that is best explained by a subset of the current independent variables is selected

- these independent variables are removed from the set of independent variables and the selected state variable is added instead

- the identified functional relationship is marked on the adjacency matrix



Conclusions

- for binary systems, the hierarchical approaches continue to outperform the naive approach as the layers of hierarchy in the hidden target system increases
- for three-layer hierarchical hidden target systems, the hierarchical approach outperforms the naive approach until arity=5
- the difficulty of scaling (all three approaches) to hierarchical systems with higher arity is due to the inherent issue of feature selection in symbolic regression (being studied separately)

References

- [1] S.V. Buldyrev, N.V. Dokholyan, S. Erramilli, M. Hong, J.Y. Kim, G. Malescio, and H.E. Stanley. *Hierarchy in social organization*. Physica A: Statistical Mechanics and its Applications, 330(3-4):653-659, 2003
- [2] Hironori Hirata and Robert E. Ulanowicz. *Information theoretical analysis of the aggregation and hierarchical structure of ecological networks*. Journal of Theoretical Biology, 116(3):321- 341, 1985.
- [3] Erzsébet Ravasz and Albert-László Barabási. *Hierarchical organization in complex networks*. Phys. Rev. E, 67:026112, Feb 2003.
- [4] Ralf J. Sommer. *Homology and the hierarchy of biological systems*. Bioessays, 30(7):653-8, 2008.

Acknowledgements

This research was supported by a DARPA grant #FA8650-11-1-7155.

The authors acknowledge the Vermont Advanced Computing Core which is supported by NASA (NNX 06AC88G), at the University of Vermont for providing High Performance Computing resources that have contributed to the research results reported within this paper