WRT-1011: Explaining Model Composability using Causal Graphs (Incubator)

Michael Pennock
School of Systems and Enterprises
Stevens Institute of Technology

Research Task / Overview

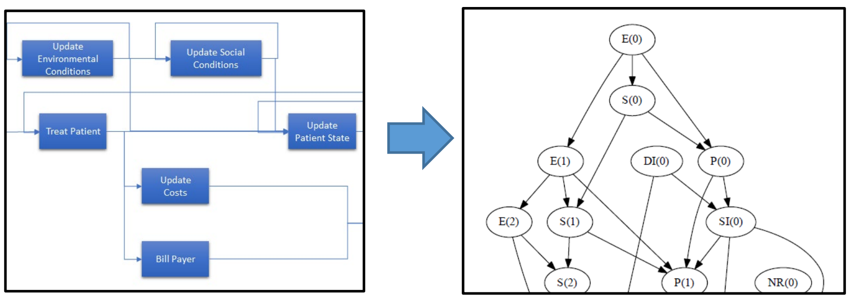
A mechanism for assessing model composability would aid those aiming to accelerate DoD's system design/test cycle through Model Based Systems Engineering (MBSE) and those aiming to construct computational frameworks for analyzing and responding to rapidly evolving threat scenarios.

The long-term goal of this line of research is to determine if causal graphs can explain why combing diverse computational engineering models works in some circumstances but not others.

Causal graphs are an evolution of Bayesian networks that enable the analysis of interventions and counterfactuals. Causal diagrams have already shown success in fusing heterogeneous statistical models and data sets. Since computational models can be viewed compressed data sets, it stands to reason that the same approach can also determine composability for computational models.

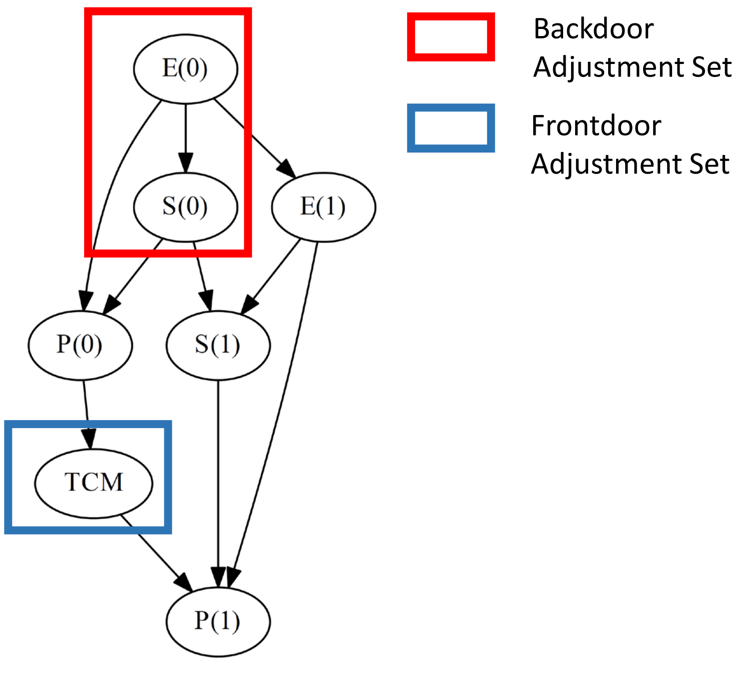
Data & Analysis

The focus of this study is on understanding semantic composability as opposed to syntactic composability. Thus, we need to understand the semantic structure of the system to be modeled. This structure is not typically represented using causal graphs. Instead, it may be described using techniques such as ontology languages, causal loop diagrams, influence diagrams, IDEF0, SysML, etc. These need to be converted to a causal diagram where each node on the diagram represents a function performed in the system. An excerpt of a conversion of a model of a healthcare system to a causal graph is shown below:



Ideally, we could partition the causal graph into sub-graphs to identify candidate modules. Unfortunately, not all partitions result in useful modules without duplication of some graph components among modules. Also, changing the question asked can change the structure of the graph. These types of issues can lead to inconsistencies when reusing modules.

The concept of d-separation can be used to find adjustment sets that may correspond to stable modules. If we want to find how the medical intervention TCM impacts the transition of patient state from P(0) to P(1) in the model below, we see that TCM is a frontdoor adjustment set. This means it can be causally isolated and captured using a separate module. This condition may be testable by controlling the indicated backdoor adjustment set during validation. If, on the other hand, the social context of the patient S(1) affects the implementation of TCM (which has been asserted), TCM could no longer be isolated and even passing the value of S(1) to the TCM module may not be sufficient.



Goals & Objectives

The specific objective of this incubator project is to investigate the feasibility of adapting approaches from computer science usedto analyze causal graphs to determine if the resulting partitions correspond with sets of computational engineering models that would be composable.

Specific questions of interest include:

- Is a condition called d-separation sufficient to establish independence among models?
- Can causal graphs be traversed to find conflicts among existing models?
- Is a set called the frontdoor adjustment set sufficient to identify a minimal module?
- Is a set called the backdoor adjustment set sufficient to establish test conditions to evaluate the validity of candidate models?

Based on the results of the feasibility assessment, the output of the incubator project will be a plan to test the approach on a larger-scale, real-world modeling problem.

Methodology

Assess the feasibility of adapting causal graph methods by applying them to a real modeling problem as a small-scale test. The test will use an existing simulation developed for a previous research project in the healthcare domain. This simulation models the adoption of a healthcare intervention called the Transitional Care Model (TCM). The steps are:

1. Construct an ontology and process model for the problem domain.
2. Convert the ontology and process model into a causal graph.
3. Test the application of d-separation, frontdoor adjustment, and backdoor adjustment to identify candidate graph decompositions.
4. Analyze how various applications of the "do()" operator alter the decompositions.
5. Compare the candidate decompositions against known the known outcomes of the prior healthcare simulation study.
6. Based on the findings, develop a plan to test the approach on a large-scale systems modeling problem.

The causal graphs will be built and analyzed using the open source Python package CausalGraphicalModels

Future Research

- Investigate the impact of aggregation relationships extracted from the system ontology on the analysis. It is expected that certain levels of abstractions will be more amenable to modularization than others.
- Investigate the impact of shifting graph structures under different contexts on the approach.
- Investigate how alternative abstractions of a system can be integrated into the analysis.
- Develop algorithms to convert models in standard languages (UML, SysML, OPM, etc.) to causal graphs

Contacts/References

Michael Pennock, PhD – Principal Investigator
Assistant Professor – Stevens Institute of Technology
mpennock@stevens.edu

References:

Bareinboim, E., & Pearl, J. (2016). Causal inference and the data-fusion problem. *Proceedings of the National Academy of Sciences*, 113(27), 7345-7352.

Pearl, J. (2000). *Causality: models, reasoning and inference* (Vol. 29). Cambridge: MIT press.