Explaining Model Composability using Causal Graphs

Sponsor: OUSD(R&E)

Michael Pennock

11th Annual SERC Sponsor Research Review
November 19, 2019
FHI 360 CONFERENCE CENTER
1825 Connecticut Avenue NW, 8th Floor
Washington, DC 20009

www.sercuarc.org

Acknowledgment

This material is based upon work supported, in whole or in part, by the U.S. Department of Defense through the Systems Engineering Research Center (SERC) under Contract HQ0034-13-D-0004. SERC is a federally funded University Affiliated Research Center managed by Stevens Institute of Technology. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the United States Department of Defense.

An Accelerating Threat Environment Requires Faster Model Development

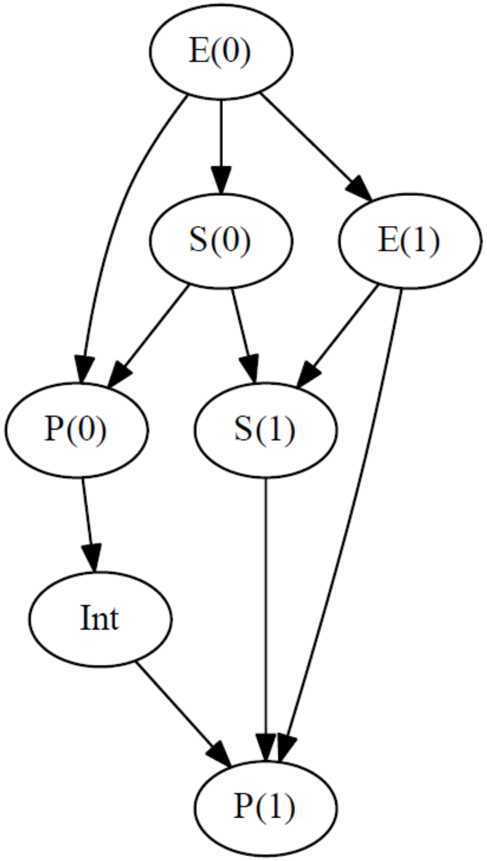
- A capability to rapidly develop and validate computational models would enable to DoD to:
  - —Computationally analyze and respond to rapidly evolving threat scenarios
  - —Accelerate DoD's system design/test cycle
- One approach to accomplish this would be to construct libraries of reusable simulation modules that could be selectively combined to create a new model
- Unfortunately, a 2017 National Science Foundation (NSF) workshop report on modeling and simulation noted that model reuse is "peculiarly fragile"
  - — Fujimoto, R., Bock, C., Chen, W., Page, E., & Panchal, J. H. (Eds.). (2017). *Research challenges in modeling and simulation for engineering complex systems*. Cham: Springer International Publishing

Need to Understand Why Composition Works Sometimes but not Others

- While experienced system modelers may have intuitive explanations for why this is so difficult, an explanatory theory is required to:
  1. Determine when it is and is not appropriate to compose diverse sets of models
  2. Develop methods, processes, and tools to enable composition when it is appropriate
- The specific objective of this incubator project is to investigate the feasibility of adapting an approach from computer science, causal graphs, to determine when sets of computational engineering models are composable.

Causal Networks Enable Isolation of Reusable Mechanisms

- Causal graphs are an evolution of Bayesian Networks
- Techniques have been developed to analyze these graphs to determine when it is possible to assess a causal effect in a data set
- When a causal effect can be isolated, it is called a mechanism, and it can be reused to understand the impact of interventions

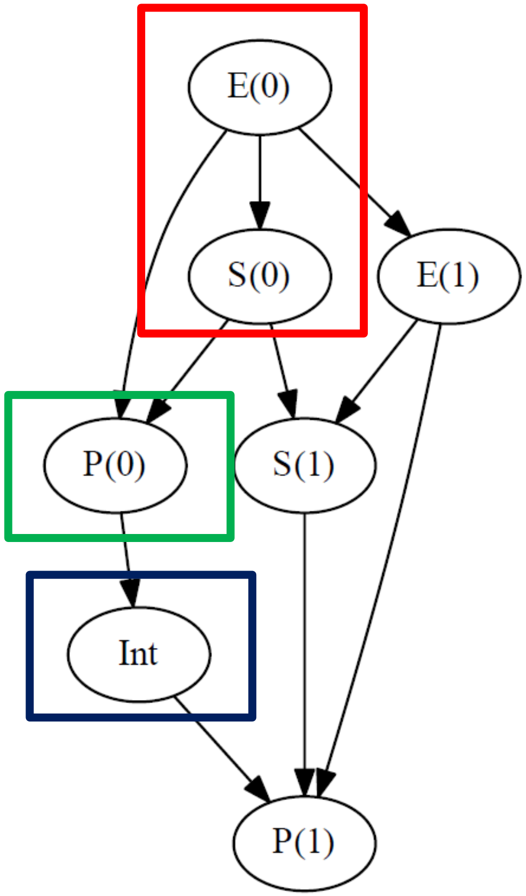


Causal Effects are Isolated by Extending Conditional Independence

- Once the of the advantages of the graph is that we can use a condition called d-separation to identify conditional independence
  - —E.g., given Z, X and Y are independent
- This idea can be extended to identify which sets of nodes must controlled to identify a causal effect
  - —Called the "backdoor" criterion
- Repeated applications of the backdoor criterion can be used determine when a causal effect can be isolated
  - —Called the "frontdoor" criterion

The Frontdoor Adjustment Set Implies Reusability

- If we want to intervene on variable "P(0)" and assess its effect on "P(1)," we can either
  - Control variables "E(0)" an "S(0)" called a backdoor adjustment set or
  - Understand the behavior of "Int" called a frontdoor adjustment set
- The advantage of the frontdoor set is that it allows us to ignore most of the other variables



Models Are Compressed Data

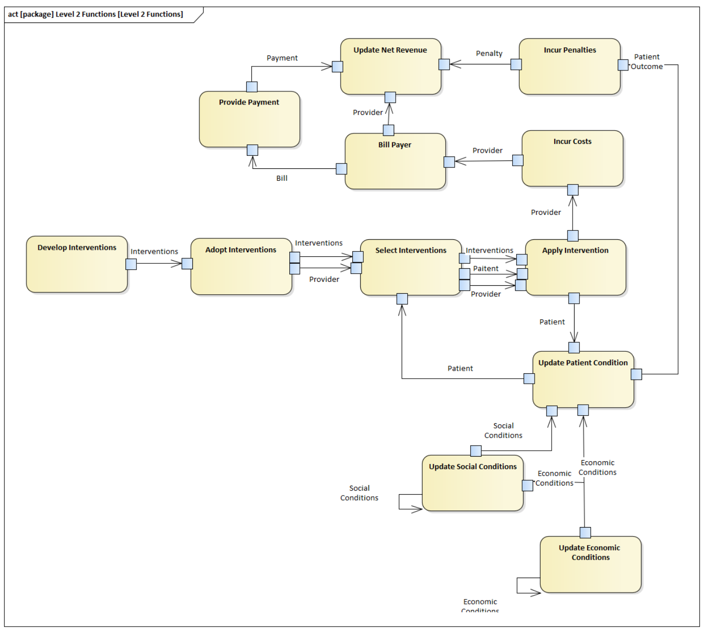
- If we create a causal of graph of the system we wish to model, we can analyze it to determine if it can be decomposed into resuable components based on the frontdoor criterion
  - —If there is no way to decompose a hypothetical data set into the modules we have, then it is impossible to accurately reproduce the data set by composing those modules
  - —This is not to suggest that models that do not satisfy the frontdoor criterion can never be reused. Rather, those that do should be more robust under reuse
- Technically, the approach only requires that the graph nodes represent functions
  - — Computational models can be represented as a composition of functions

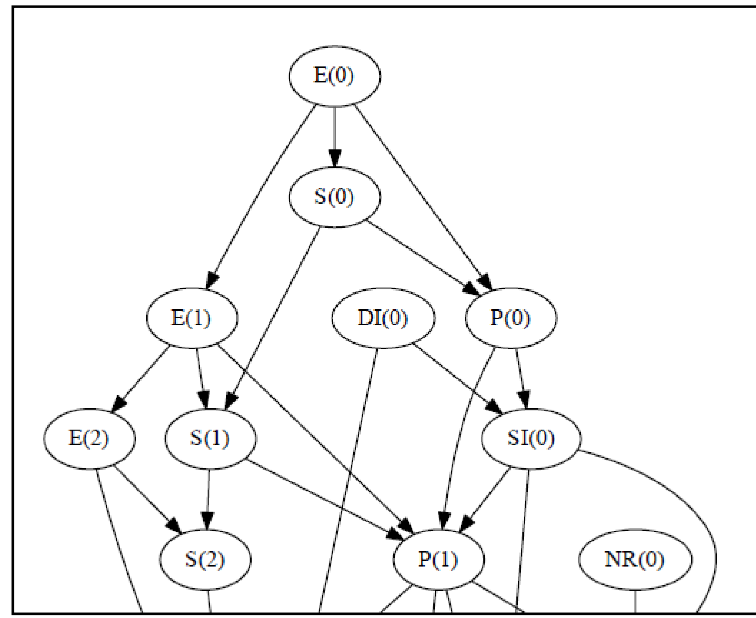
There are Several Potential Applications

- Identify how domains might be decomposed into reusable modules to simulate alternative system configurations
- Build a causal graph of a system, and then determine which if any combination of existing modules can be composed to represent it accurately
- Analyze how a shift in context or behavior might affect existing simulations
- Develop targeted experiments to make inferences about the causal structure of a black-box model

Causal Graphs Can Be Extracted from Standard System Models

- Modeling languages such as SysML and OPM contain the necessary information, but their diagrams need to be unfolded and mapped to functions

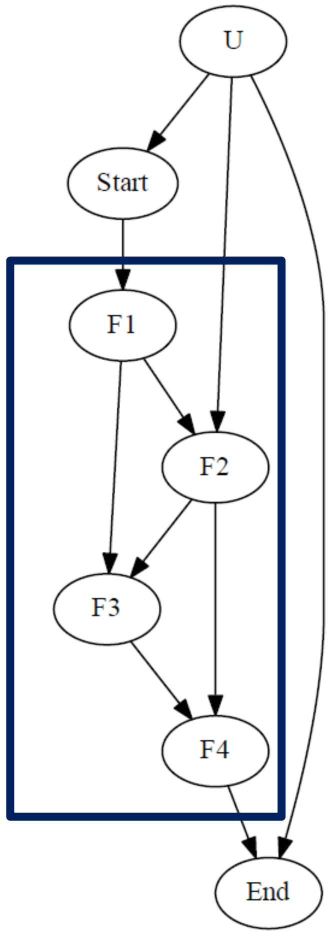




- With a graph of the system, decompositions can be identified and compared

Adapting the Approach to System Models Poses Research Challenges

- As the graphs become more complicated, the potential modules become less obvious
  - Algorithms have been constructed to identify adjustment sets, but they need to be tested for efficacy for this application
- Different contexts may require different graphs, and small changes can impact the viability of modules
- Changing the question asked could change the decomposition
- Integrating different abstractions or levels of resolution should provide interesting insights, but will require more investigation



Next Research Steps

- For the incubator, the concepts and algorithms are being tested on a medium-scale problem from the healthcare domain
  - The intent is to establish the feasibility of the approach
  - The outcome of a prior simulation study enables straightforward comparisons
- If using causal graphs appears feasible, the next step is to test the approach on a larger-scale, more representative system
  - Graph algorithms will need to be adapted
  - Limited simulation development may be required to test predictions generated by from the causal graph models

Questions?