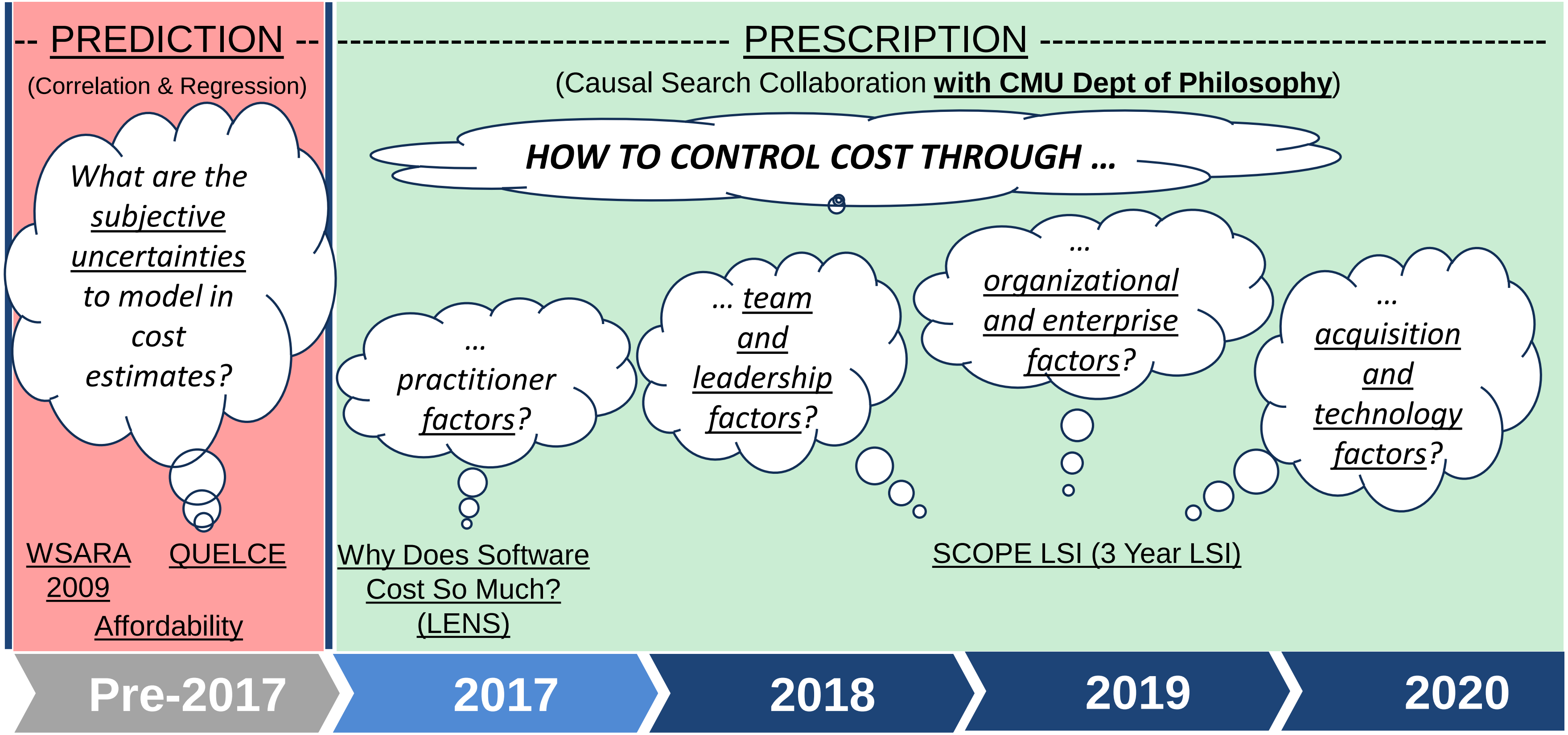
Benefits of SEI-CMU Collaboration regarding Use of Causal Learning

Robert W. Stoddard, Principal Researcher, SEI

Dave Danks, Department Head & L.L. Thurstone Professor of Philosophy and Psychology

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Context of SCOPE Research



Initial SCOPE Causal Search Results

Controlling Size: Only 2 of 4 code size measures appear causal on effort and quality

Controlling Complexity: Only 1 of 3 factors appears causal on performance and quality

Controlling Architecture Violations: Only 1 of 4 violation factors appears causal on quality

Controlling Team Performance: Only 1 of 20+ factors appears causal on quality and cost

Causal search may provide useful feedback:

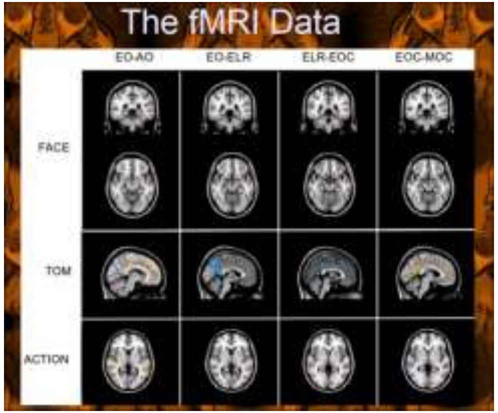
1) *Presence of causal links*
2) *Absence of causal links*

Benefits of CMU Collaboration

1. World class expertise and coaching

2. Search algorithms & Tetrad updates
   (IMAGES, FASK, Multi-FASK, Bootstrapping, Cyclic search)

3. Sharing search approaches from other domains
   (classification approach with fMRI causal results)



4. Students test new algorithms and updates to Tetrad

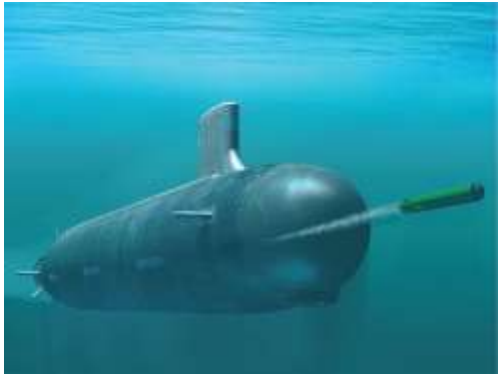
5. Research-to-practice and practice-to-research cycles

Lessons Learned from CMU Collaboration

1. Causal search remains a mix of science and art
2. Causal search strategies are not well understood and routinized
3. Opportunities exist to further integrate machine and causal learning
4. Richer collaboration needed leading up to the research proposal
5. Fundamental research tasks need to be more clearly delineated from restricted research

Moderate Future: Causal Learning for Simulation and Test

Problem
Lack of accredited simulators



Technical Challenge
Experts unsure of the expected result for a given simulated scenario

Research Questions

1. Scale up metamorphic testing to test very complex DoD systems?
2. Machine learning to identify metamorphic relations for testing?
3. Causal learning to drive metamorphic relations testing?

Moderate Future: Causal Learning for Sustainment

Problem
Unscheduled maintenance creates unacceptable costs

Technical Challenge
Traditional statistical approaches helpful, but insufficient

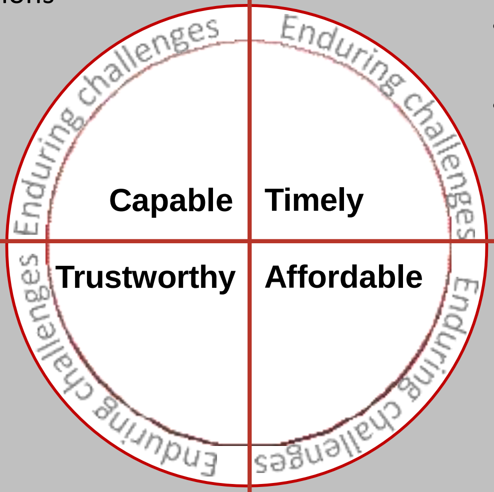
Research Questions

1. Machine learning of engine sensor and control data improve scheduled maintenance?
2. Causal learning integrated with machine learning add value?

Long Term Future: (Causal Learning Examples)

- Causal drivers of workforce performance
- SW architecture strategies and tactics driving system performance
- More efficient experimentation of technical solutions
- Increased realism of complex system simulation
- Autonomous systems controlling consequences
- Machine learning with human-like intelligence (e.g. *"Strong AI"; Pearl, "The Book of Why"*)

- Causal structures from DevOps information stream to control process and lifecycle
- Agile causal systems situationally prescribe practices aligned with goals
- Project risks controlled through causal structures of project parameters



- Causal factors threatening cyber defenses
- Causal factors limiting resilience
- CL combined with ML tools for more affordable and trustworthy SW technologies (e.g. DOD initiative in Digital Engineering)
- Expected behavior from autonomous systems (e.g. "*Explainable AI*"; Jensen, UMass)

- Acquisition practice improved using causal models
- Cost estimates and budget execution using causal models
- Simpler but more effective ROI models based on causal factors (e.g. Model Based Engineering, Architecture practice, Technical Debt)

Contact Information

Presenter / Point(s) of Contact

David Danks

Bob Stoddard

Email:

ddanks@cmu.edu
rws@sei.cmu.edu

Telephone:

+1 412.268.8047 (David)

+1 412.268.1121 (Bob)

Other SEI SCOPE Team Members

Mike Konrad

Bill Nichols

Sarah Sheard

Dave Zubrow

Other CMU SCOPE Contributors

Madelyn Glymour

Joe Ramsey

Kun Zhang

USC SCOPE Contributors

Jim Alstad

Barry Boehm

Anandi Hira