Towards automating reasoning over cybersecurity data

Jonathan M. Spring¹ and David Pym¹,²

¹ University College London
² Alan Turing Institute

Data Science for Cybersecurity
September 26, 2017
Goal

Discuss results of Pym et al. [2017]

- Use of Separation Logic for finding memory management errors

Motivate application of similar approach to reasoning in cybersecurity

- Hope to avoid some shortcomings of machine learning

- Sketch application to incident response (IR) and investigation

This contributes to a broader discussion involving evidence evaluation [Spring and Illari, 2017] and philosophy of security [Spring et al., 2017, Hatleback and Spring, 2014].
J. Pearl

Hierarchy of causal importance [Pearl, 2016]:

► What is?
  ▶ “can be computed efficiently using Bayesian Networks, or any of the graphical models that support deep-learning systems”

► What if?
  ▶ Requires reasoning about interventions, see Woodward [2003], Halpern [2015]

► Why?
  ▶ Requires manipulable models with structure
  ▶ Pearl suggests Counterfactuals

Pearl’s is not the only answer to “what if” & why
Separation Logic

But the answer is not traditional ML

SL introduces $\ast$ for “and, separately”
  ▶ Distinct from $\land$ for the usual sense of “and”

We take lessons from the history of SL implementation in Infer, file systems, OS scheduling, etc.

SL works
  ▶ Infer is open source, used by FB, others
  ▶ What we wanted to know was why it works
Enabling Innovations

- Overlap of two (types of) models:
  - Useful engineering model of RAM
  - Resource-based logical model and language

- Scalable proof theory
  - Compositional local reasoning via a modified Frame Rule
  - Automation of abductive inference

- Hard work adapting to actual software development practices (see O’Hearn [2015] for this)
Abduction

Third reasoning type, vice induction and deduction

Initially introduced by Peirce around 1900:

‘Abduction is the process of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea’

[Bergman and Paavola, 2016, CP 5.171]
Abduction

In Separation Logic, this looks like:

1. Attempt a proof of a code segment
   1.1 Fail

2. Using purpose-built proof rules check:
   2.1 If the heuristic matches the failed proof state
   2.2 What additional condition would be necessary for the proof to work
   2.3 Guess the program contains this condition somewhere
      2.3.1 And look for it

The proof-rules match engineer’s heuristics for fixing bugs
Abduction

Instead of a software developer’s bug-fixing heuristics, we want to re-make this for:

*Heuristics to figure out how an attacker broke in during computer security incident response*
IR Primer

Incident management at a high level
(Figure 3 from Alberts et al. [2004])
IR Primer

The following legal statements apply to this image of the 5 high-level steps in incident management anywhere it occurs (though to nothing else in the presentation):
Copyright 2004 Carnegie Mellon University. NO WARRANTY THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT. Permission to reproduce this document and to prepare derivative works from this document for internal use is granted, provided the copyright and "No Warranty" statements are included with all reproductions and derivative works.
Intrusions

IR is a response to intrusions, several models:

- A “Common Language” [Howard and Longstaff, 1998]
- 5 steps [Bejtlich, 2004]
- 7-step “kill chain” [Hutchins et al., 2011]
- “Diamond model” [Caltagirone et al., 2013]

Investigators ask “what-if” questions based on models like these [Spring and Hatleback, 2017]

- Within “response” we have parts, too
- Investigation is: evidence collection, analysis, and reporting
Intrusions

IR is a response to intrusions, several models:

- A “Common Language” [Howard and Longstaff, 1998]
- 5 steps [Bejtlich, 2004]
- 7-step “kill chain” [Hutchins et al., 2011]
- “Diamond model” [Caltagirone et al., 2013]

Investigators ask “what-if” questions based on models like these [Spring and Hatleback, 2017]

- Within “response” we have parts, too
- Investigation is: evidence collection, analysis, and reporting
Investigation is:

A process carried out by an agent to build an explanation of the mechanism by which security policy was violated.\(^1\) Forensic investigation is historical.\(^2\) The process includes data collection, in some jargon, via manipulation of available resources. The agent’s modelling decisions are resource-constrained, and follow a methodology. The output of the process is that the agent reports results, namely, their new beliefs relative to goals.

\(^1\) More generally, the phenomenon of interest

\(^2\) As opposed to engineering or design which are future-oriented.
Investigation is:

A process carried out by an agent to build an explanation of the mechanism by which security policy was violated.\textsuperscript{1} Forensic investigation is historical.\textsuperscript{2} The process includes data collection, in some jargon, via manipulation of available resources. The agent’s modelling decisions are resource-constrained, and follow a methodology. The output of the process is that the agent reports results, namely, their new beliefs relative to goals.

\textsuperscript{1} More generally, the phenomenon of interest
\textsuperscript{2} As opposed to engineering or design which are future-oriented
Investigation is:

A process carried out by an agent to build an explanation of the mechanism by which security policy was violated.\(^1\) Forensic investigation is historical.\(^2\) The process includes data collection, in some jargon, via manipulation of available resources. The agent’s modelling decisions are resource-constrained, and follow a methodology. The output of the process is that the agent reports results, namely, their new beliefs relative to goals.

\(^1\) More generally, the phenomenon of interest

\(^2\) As opposed to engineering or design which are future-oriented
In a logic this would be...

We need a concept of time

- Temporal logic tends to be about the future
- Existing SL-TL mixtures keep a syntactic separation [Espinosa and Brotherston, 2017]

Plan to use histories of change as a resource

- Allows interleaving of different operators

Start with just one investigator for simplicity

- Other work talks about coordination between CSIRTs [Osorno et al., 2011]
In a logic this would be...

We need to understand what heuristics incident responders use

- Science can give generalization & discovery heuristics [Spring and Illari, 2017]
- Extract some from the process of mathematical modelling
- Draw on standards [Cichonski et al., 2012]
- Draw on case studies [Stoll, 1989, Mandiant, 2013]

The heuristics will not be perfect at first, but at least the language should allow for testing and evaluation
Summary

Reasoning about cybersecurity data cannot rely on black-box algorithms

- Need to know ‘What if’ and Why
- ‘What is’ questions have a place

There are languages that enable such questions
History of Infer shows Separation Logic is one such language

- Also teaches there are no shortcuts
- Heuristics and tools must be carefully tuned to the problem of interest
Questions?

“Why Separation Logic Works” draft available: http://www0.cs.ucl.ac.uk/staff/D.Pym/recent.htm
Contact: jonathan.spring.15 (AT) ucl.ac.uk
References I


References II


References III


Copyright

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International License. To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/4.0/.
(CC BY-SA 4.0)
Acknowledgements

The Separation Logic paper, from which these slides are partly derived, is joint work with Peter O’Hearn.
Spring is supported by University College London’s Overseas Research Scholarship and Graduate Research Scholarship. orcid.org/0000-0001-9356-219X