

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                |                                                       |                                                           |                                                   |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|--------------------------------|-------------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------|
| REPORT DOCUMENTATION PAGE                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                   |                                | Form Approved OMB NO. 0704-0188                       |                                                           |                                                   |
| <p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p> |                   |                                |                                                       |                                                           |                                                   |
| 1. REPORT DATE (DD-MM-YYYY)<br>15-08-2018                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         |                   | 2. REPORT TYPE<br>Final Report |                                                       | 3. DATES COVERED (From - To)<br>18-Dec-2013 - 17-Dec-2016 |                                                   |
| 4. TITLE AND SUBTITLE<br>Final Report: Learning, Awareness, Optimism, and Confidence                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                   |                                | 5a. CONTRACT NUMBER<br>W911NF-14-1-0017               |                                                           |                                                   |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                | 5b. GRANT NUMBER                                      |                                                           |                                                   |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                | 5c. PROGRAM ELEMENT NUMBER<br>611102                  |                                                           |                                                   |
| 6. AUTHORS<br>Joseph Y. Halpern                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                | 5d. PROJECT NUMBER                                    |                                                           |                                                   |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                | 5e. TASK NUMBER                                       |                                                           |                                                   |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                | 5f. WORK UNIT NUMBER                                  |                                                           |                                                   |
| 7. PERFORMING ORGANIZATION NAMES AND ADDRESSES<br>Cornell University<br>Office of Sponsored Programs<br>373 Pine Tree Road<br>Ithaca, NY 14850 -2820                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                   |                                | 8. PERFORMING ORGANIZATION REPORT NUMBER              |                                                           |                                                   |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)<br>U.S. Army Research Office<br>P.O. Box 12211<br>Research Triangle Park, NC 27709-2211                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                   |                                | 10. SPONSOR/MONITOR'S ACRONYM(S)<br>ARO               |                                                           |                                                   |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                | 11. SPONSOR/MONITOR'S REPORT NUMBER(S)<br>64452-NS.46 |                                                           |                                                   |
| 12. DISTRIBUTION AVAILABILITY STATEMENT<br>Approved for public release; distribution is unlimited.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |                   |                                |                                                       |                                                           |                                                   |
| 13. SUPPLEMENTARY NOTES<br>The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                |                                                       |                                                           |                                                   |
| 14. ABSTRACT<br>In decision theory, it is typically assumed that a decision maker knows the state space, the outcome space, and the set of feasible acts. But this is far from clear in practice. In a complex decision problem, agents may be unaware of many relevant features, and thus unaware of possible states, outcomes, and feasible acts.<br><del>This project has investigated how to proceed when these assumptions</del>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |                   |                                |                                                       |                                                           |                                                   |
| 15. SUBJECT TERMS<br>learning, awareness, decision theory, game theory, MDPs                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |                   |                                |                                                       |                                                           |                                                   |
| 16. SECURITY CLASSIFICATION OF:                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                   |                                | 17. LIMITATION OF ABSTRACT<br>UU                      | 15. NUMBER OF PAGES                                       | 19a. NAME OF RESPONSIBLE PERSON<br>Joseph Halpern |
| a. REPORT<br>UU                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | b. ABSTRACT<br>UU | c. THIS PAGE<br>UU             |                                                       |                                                           | 19b. TELEPHONE NUMBER<br>607-255-9562             |

# RPPR Final Report

## as of 20-Sep-2018

Agency Code:

Proposal Number: 64452NS

Agreement Number: W911NF-14-1-0017

### INVESTIGATOR(S):

**Name:** Joseph Halpern  
**Email:** halpern@cs.cornell.edu  
**Phone Number:** 6072559562  
**Principal:** Y

Organization: **Cornell University**

Address: Office of Sponsored Programs, Ithaca, NY 148502820

Country: USA

DUNS Number: 872612445

EIN: 150532082

**Report Date:** 17-Mar-2017

Date Received: 15-Aug-2018

**Final Report** for Period Beginning 18-Dec-2013 and Ending 17-Dec-2016

**Title:** Learning, Awareness, Optimism, and Confidence

**Begin Performance Period:** 18-Dec-2013

**End Performance Period:** 17-Dec-2016

**Report Term:** 0-Other

Submitted By: Joseph Halpern

Email: halpern@cs.cornell.edu

Phone: (607) 255-9562

**Distribution Statement:** 1-Approved for public release; distribution is unlimited.

**STEM Degrees:** 0

**STEM Participants:** 3

**Major Goals:** The goal of this project is to investigate operations on causal models that will be critical making best use of causal information. The initial focus will be on two operations: combining models and abstracting models.

With regard to combining models, the assumption is that experts will provide causal models to a policymaker, who wants to combine (the information in) these models to reach her decision. Bradley, Dietrich, and List have provided arguably reasonable desiderata for combining causal models, and showed that there is no way of combining causal models so as to satisfy all their desiderata. The hope is that we can provide conditions under which models are compatible. The general approach to combining experts' judgements would then be to combine the experts' models when they are compatible, and if not, just place a probability on each model being the right model, using relatively standard techniques based on the perceived reliability of the experts who proposed them. Intuitively, if the experts' models are not compatible, then the experts are disagreeing, so we should not try to combine their models; rather, we should just assign a likelihood to each model being right. On the other hand, if the models are compatible, then we should consider a model that takes into account both experts information.

With regard to abstraction, suppose that an expert provides a detailed causal model, involving many variables. Such a detailed model may be far too complicated for a policymaker to understand and work with. What the policymaker wants is a high-level "macro" model of the situation. Such a high-level model can result from combining many variables at the "micro" level into one "macro" variable. There have been recent arguments made that such a high-level model actually can, in a precise sense, provide more information than a detailed model. A high-level model can also be helpful when it comes to deciding appropriate interventions to perform. For example, by introducing a

## RPPR Final Report as of 20-Sep-2018

new concept such as “soccer mom”, we may be able to make more sense out of voting data, and think of more useful interventions that can be performed (e.g., run political ads to influence soccer moms). But what precisely does it mean that a high-level causal model is a faithful abstraction of a low-level model? Intuitively, this should mean that the high-level model captures all the information in the low-level model, at the level of variables used in the high-level model. I hope to make this notion precise, and come up with algorithms for testing if a high-level model is a faithful abstraction of a low-level model. In practice, we also will want a notion of approximate abstraction; that is, a high-level model may not exactly capture all the information in a low-level model, but it may come close. A good notion of approximation will let us quantify how far off one causal model is from another. This will enable us to quantify how much we lose by working with a simpler, high-level model.

**Accomplishments:** We have made significant progress on the goal of combining models. In a paper that appeared in AAAI 2018, Dalal Alrajeh, Hana Chockler, and I provided an approach for doing just what was hoped for in the proposal: we defined a notion of compatibility for causal models, and showed how compatible causal models can be combined. In recent follow-on work with my Ph.D. student Meir Friedenberg, which will KR 2018, we showed how causal models can be combined in cases where the experts might disagree on the causal structure for variables that appear in both models due to having different focus areas. We provided a new formal definition of compatibility of models in this setting and show how compatible models can be combined.

With regard to abstraction, I am working actively with Sander Becker (Utrecht) on Frederick Eberhardt (Cal Tech) on finding good notions of abstraction and approximate abstraction.

**Training Opportunities:** Nothing to Report

**Results Dissemination:** The results of this project have been presented at AAAI (a major AI conference) and at numerous invited talks given by the PI.

**Honors and Awards:** Recipient of Lady Davis Fellowship, Hebrew University, 2018.

Elected Chair-Elect of the Section on Information, Computing, & Communication of AAAS (American Association for the Advancement of Science): term of office 2018-2021

Jin Yuelin Lecturer at Tsinghua University, 2017.

Selected Fellow of the Game Theory Society, 2017.

**Protocol Activity Status:**

**Technology Transfer:** Nothing to Report

### PARTICIPANTS:

**Participant Type:** Faculty

**Participant:** Joseph Halpern

**Person Months Worked:** 3.00

Project Contribution:

**Funding Support:**

**RPPR Final Report**  
as of 20-Sep-2018

International Collaboration:  
International Travel:  
National Academy Member: N  
Other Collaborators:

**Participant Type:** Graduate Student (research assistant)

**Participant:** Matvey Soloviev

**Person Months Worked:** 2.00

**Funding Support:**

Project Contribution:  
International Collaboration:  
International Travel:  
National Academy Member: N  
Other Collaborators:

**Participant Type:** Graduate Student (research assistant)

**Participant:** Meir Friendenberg

**Person Months Worked:** 1.00

**Funding Support:**

Project Contribution:  
International Collaboration:  
International Travel:  
National Academy Member: N  
Other Collaborators:

**Participant Type:** Graduate Student (research assistant)

**Participant:** Smaranda Sandu

**Person Months Worked:** 1.00

**Funding Support:**

Project Contribution:  
International Collaboration:  
International Travel:  
National Academy Member: N  
Other Collaborators:

**ARTICLES:**

## RPPR Final Report as of 20-Sep-2018

**Publication Type:** Journal Article      Peer Reviewed: Y      **Publication Status:** 1-Published

**Journal:** Topics in Cognitive Science

Publication Identifier Type:

Publication Identifier:

Volume: 6

Issue: 2

First Page #: 245

Date Submitted:

Date Published:

Publication Location:

**Article Title:** Decision theory with resource-bounded agents

**Authors:**

**Keywords:** decision making, resource-bounded agents, computation

**Abstract:** There have been two major lines of research aimed at capturing resource-bounded players in game theory. The first, initiated by Rubinstein \citeyear{Rub85}, charges an agent for doing costly computation; the second, initiated by Neyman \citeyear{Ney85}, does not charge for computation, but limits the computation that agents can do, typically by modeling agents as finite automata. We review recent work on applying both approaches in the context of decision theory. For the first approach, we take the objects of choice in a decision %where players are charged for the ``complexity" of the Turing machine and charge players for the ``complexity" of the Turing machine chosen (e.g., its %joe6 %running time). We show how this approach can be used to explain running time). This approach can be used to explain well-known phenomena like \emph{first-impression-matters biases} (i.e., people tend to put more weight on evidence they hear early on) %joe6 and \emph{belief polarization} (two

**Distribution Statement:** 1-Approved for public release; distribution is unlimited.

Acknowledged Federal Support:

**Publication Type:** Journal Article      Peer Reviewed: Y      **Publication Status:** 1-Published

**Journal:** Artificial Intelligence

Publication Identifier Type:

Publication Identifier:

Volume: 209

Issue: 0

First Page #: 1

Date Submitted:

Date Published:

Publication Location:

**Article Title:** A logic for reasoning about ambiguity

**Authors:**

**Keywords:** ambiguity, agreement, semantics

**Abstract:** Standard models of multi-agent modal logic do not capture the fact that information is often \emph{ambiguous}, and may be interpreted in different ways by different agents. We propose a framework that can model this, and consider different semantics that capture different assumptions about the agents' beliefs regarding whether or not there is ambiguity. We examine the expressive power of logics of ambiguity compared to logics that cannot model ambiguity, with respect to the different semantics that we propose.

**Distribution Statement:** 1-Approved for public release; distribution is unlimited.

Acknowledged Federal Support:

**Publication Type:** Journal Article      Peer Reviewed: N      **Publication Status:** 1-Published

**Journal:** British Journal for the Philosophy of Science

Publication Identifier Type:

Publication Identifier:

Volume: 0

Issue: 0

First Page #: 0

Date Submitted:

Date Published:

Publication Location:

**Article Title:** Graded causation and defaults

**Authors:**

**Keywords:** causality, normality

**Abstract:** Recent work in psychology and experimental philosophy has shown that judgments of actual causation are often influenced by consideration of defaults, typicality, and normality. A number of philosophers and computer scientists have also suggested that an appeal to such factors can help deal with problems facing existing accounts of actual causation. This paper develops a flexible formal framework for incorporating defaults, typicality, and normality into an account of actual causation. The resulting account takes actual causation to be both graded and comparative. We then show how our account would handle a number of standard cases.

**Distribution Statement:** 1-Approved for public release; distribution is unlimited.

Acknowledged Federal Support:

# RPPR Final Report

## as of 20-Sep-2018

**Publication Type:** Journal Article      Peer Reviewed: Y      **Publication Status:** 1-Published

**Journal:** Journal of Artificial Intelligence Research

Publication Identifier Type:      Publication Identifier:

Volume: 49      Issue: 0      First Page #: 143

Date Submitted:      Date Published:

Publication Location:

**Article Title:** A procedural characterization of solution concepts in games

**Authors:**

**Keywords:** solution concepts, knowledge-based programs

**Abstract:** We show how game-theoretic solution concepts such as Nash equilibrium, correlated equilibrium, rationalizability, and sequential equilibrium can be given a uniform definition in terms of a knowledge-based program with counterfactual semantics. In a precise sense, this program can be viewed as providing a procedural characterization of rationality

**Distribution Statement:** 1-Approved for public release; distribution is unlimited.

Acknowledged Federal Support:

**Publication Type:** Journal Article      Peer Reviewed: Y      **Publication Status:** 1-Published

**Journal:** Games and Economic Behavior

Publication Identifier Type: DOI      Publication Identifier: 10.1016/j.geb.2014.05.012

Volume: 87      Issue: 0      First Page #: 0

Date Submitted:      Date Published:

Publication Location:

**Article Title:** Conditional belief types

**Authors:**

**Keywords:** type space, hypothetical knowledge, conditional probability, counterfactuals

**Abstract:** We study type spaces where a player's type at a state is a conditional probability on the space. We axiomatize these spaces using conditional belief operators, examining three additional axioms of increasing strength. First, introspection, which requires the agent to be unconditionally certain of her beliefs. Second, echo, according to which the unconditional beliefs implied by the condition must be held given the condition. Third, determination, which says that the conditional beliefs are the unconditional beliefs that are conditionally certain. Echo implies that conditioning on an event is the same as conditioning on the event being certain, which formalizes the standard informal interpretation of conditional probability. The game-theoretic application of our model, discussed within an example, sheds light on a number of issues in the analysis of extensive form games. Type spaces are closely related to the sphere models of counterfactual conditionals and to models of hypothetical knowledge

**Distribution Statement:** 1-Approved for public release; distribution is unlimited.

Acknowledged Federal Support:

### BOOKS:

**Publication Type:** Book      Peer Reviewed: N      **Publication Status:** 1-Published

Publication Identifier Type: ISBN      Publication Identifier: 978026203502

Book Edition:      Volume:      Publication Year: 2016      Date Received: 21-Aug-2016

Publication Location:

Publisher: MIT Press

**Book Title:** Actual Causality

**Authors:** Joseph Y. Halpern

**Editor:**

Acknowledged Federal Support: Y

### CONFERENCE PAPERS:

## RPPR Final Report as of 20-Sep-2018

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** Fifteenth Conference on Theoretical Aspects of Rationality and Knowledge(TARK 2015)  
Date Received: 21-Aug-2016 Conference Date: 04-Jun-2015 Date Published:  
Conference Location: Pittsburgh, PA  
**Paper Title:** Bayesian games with intentions  
**Authors:** Adam Bjorndahl, Joseph Y. Halpern, Rafael Pass  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** Fifteenth International Conference on Principles of Knowledge Representation and Reasoning (KR 2016)  
Date Received: 21-Aug-2016 Conference Date: 25-Apr-2016 Date Published: 25-Apr-2016  
Conference Location: Cape Town, South Africa  
**Paper Title:** Sequential equilibrium and imperfect recall  
**Authors:** Joseph Y. Halpern, Rafael Pass  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)  
Date Received: 21-Aug-2016 Conference Date: 08-May-2016 Date Published: 08-May-2016  
Conference Location: Singapore  
**Paper Title:** Decentralised Norm Monitoring in Open Multi-Agent Systems  
**Authors:** Natasha Alechina, Joseph Y. Halpern, Ian A. Kash, Brian Logan  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** 32nd Conference on Uncertainty in AI (UAI 2016)  
Date Received: 21-Aug-2016 Conference Date: 25-Jun-2016 Date Published: 25-Jun-2016  
Conference Location: New York City, NY  
**Paper Title:** MDPs with unawareness in robotics,  
**Authors:** Nan Rong, Joseph Y. Halpern, Ashutosh Saxena  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** 17th ACM Conference on Electronic Commerce  
Date Received: 21-Aug-2016 Conference Date: 24-Jul-2016 Date Published: 24-Jul-2016  
Conference Location: Maastricht, Netherlands  
**Paper Title:** Computational Extensive-Form Games  
**Authors:** Joseph Y. Halpern, Rafael Pass, Lior Seeman  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** 35th Annual ACM Symposium on Principles of Distributed Computing  
Date Received: 21-Aug-2016 Conference Date: 25-Jul-2016 Date Published: 25-Jul-2016  
Conference Location: Chicago, IL  
**Paper Title:** Rational consensus: Extended Abstract,  
**Authors:** Joseph Y. Halpern, Xavier Vilaca  
Acknowledged Federal Support: **Y**

**RPPR Final Report**  
as of 20-Sep-2018

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)  
Date Received: 29-May-2017      Conference Date: 06-Feb-2017      Date Published: 02-Feb-2017  
Conference Location: San Francisco, CA  
**Paper Title:** Incentivising monitoring in open normative systems  
**Authors:** N. Alechina, J. Y. Halpern, I. Kash, B. Logan,  
Acknowledged Federal Support: **Y**

**Publication Type:** Conference Paper or Presentation **Publication Status:** 1-Published  
**Conference Name:** Sixteenth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017)  
Date Received: 29-May-2017      Conference Date: 08-May-2017      Date Published: 08-May-2017  
Conference Location: Sao Paulo, Brazil  
**Paper Title:** Causality, responsibility, and blame in team plans  
**Authors:** N. Alechina, J. Y. Halpern, B. Logan,  
Acknowledged Federal Support: **Y**



Nothing to report in the uploaded pdf (see accomplishments).