PREDICTIVE SENSOR TASKING AND DECISION SUPPORT FOR SPACE SITUATIONAL AWARENESS USING EVIDENTIAL REASONING

A Dissertation
Presented to
The Academic Faculty

By

Andris D. Jaunzemis

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Aerospace Engineering

Georgia Institute of Technology

May 2018

PREDICTIVE SENSOR TASKING AND DECISION SUPPORT FOR SPACE SITUATIONAL AWARENESS USING EVIDENTIAL REASONING

Approved by:

Dr. Marcus J. Holzinger, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Karen Feigh
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Eric Johnson
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Travis Blake
Principal, Physical Sciences
Kairos Ventures

Dr. K. Kim Luu
Aerospace Engineer
Air Force Research Laboratory

Date Approved: March 28, 2018

The best thesis defense is a good thesis offense.

Randall Munroe, XKCD #1403

To my parents, whose encouragement and support got me here;

To Samantha, whose unconditional love keeps me going;

Et ad maiorem Dei gloriam.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Marcus Holzinger, for many hours of guidance on research, graduate school, careers, and life in general. Without a doubt, your emphasis on first-principles approaches and mathematical rigor has made me a better engineer. Thank you for investing in me as a researcher and as an individual.

I would also like to thank the rest of my committee for their guidance through the thesis process. Your comments and suggestions have helped both shape this work and make it stronger, and your encouragement has helped drive me toward this milestone.

Of course, I would not have made it through the assignments, deadlines, and nights at the observatory without my colleagues in the Holzinger research group. I appreciate the many useful discussions about research problems, but also the much-needed distractions from work. I look forward to reuniting with you all at conferences and other future events!

Finally, this research was partially supported by the National Science Foundation Graduate Research Fellowship Program (NSF GRFP grant DGE-1148903) and Lockheed Martin Corporation. The opinions, findings, and conclusions or recommendations expressed in this material are the author's alone and do not necessarily reflect the views of the National Science Foundation or Lockheed Martin.

TABLE OF CONTENTS

Acknowledgments v

List of Tables xiii

List of Figures xv

Chapter 1: Introduction and Motivation 1

1.1 Binary Hypothesis Testing for Anomaly Detection 5

1.2 Ignorance-Reduction Criterion for Sensor Tasking 6

1.3 Evidence-Gathering for Hypothesis Resolution using Judicial Evidential Reasoning 8

1.4 Cognitive Systems Engineering Applied to Decision Support in SSA 10

1.5 Organization and Relevant Literature 11

1.6 List of Publications 11

Chapter 2: Development of Requirements for SSA Decision Support 15

2.1 Cognitive Systems Engineering 15

2.1.1 Situation Awareness 15

2.1.2 Decision Support Systems 16

2.1.3 Cognitive Work Analysis 17

2.2 Work Domain Analysis Applied to SSA 18

2.2.1 Description of the Work Domain 19
2.2.2 Abstraction Hierarchy 21
2.2.3 SSA Work Domain Decomposition 23
2.2.4 SSA Environment Decomposition 26
2.2.5 Insights for SSA DSS Development 30
2.3 Control Task Analysis Applied to SSA 31
2.3.1 Decision Ladders 31
2.3.2 ConTA Application: Information Fusion and Sensor Allocation . . . 34
2.4 Conclusions 39

Chapter 3: Binary Hypothesis Testing for Anomaly Detection 40

3.1 Data Association in SSA 41
3.2 UCT Association Scenario 44
3.3 Gaussian Mixture Model Approximation for Control Distance Metric 45
3.4 Anomaly Detection using Binary Hypothesis Testing 52
3.5 Binary Hypotheses for Mahalanobis Distance 56
3.6 Simulated Results 58
3.6.1 Trajectory Optimization 59
3.6.2 PDF Generation 59
3.6.3 Anomaly Detection 60
3.7 Gaussian Mixture Model Approximation Validation 61
3.7.1 Inclination Change 65
3.7.2 Phasing Maneuvers 69

3.8 Empirical Data Results 70

3.8.1 Station-Keeping 71

3.9 Operational Application 72

3.10 Conclusions 73

Chapter 4: Spacecraft Custody Maintenance and Anomaly Detection using Evidential Reasoning 75

4.1 Dempster-Shafer Theory 76

4.1.1 Belief Functions 77

4.1.2 Combination Rules 79

4.1.3 The Curse of Dimensionality 81

4.1.4 Placing Bets from Evidence 82

4.1.5 Binary Hypothesis BPAs 83

4.2 Ignorance Criterion for Tasking Decisions 84

4.3 Multi-Sensor Multi-Target Ignorance-Reduction Tasking 85

4.3.1 Greedy Minimum-Ignorance Optimization 87

4.3.2 Receding Horizon Minimum-Ignorance Optimization 87

4.4 Application to SSA Sensor Tasking 89

4.4.1 Custody and Anomaly Discrimination Framework 90

4.4.2 SSA Dempster-Shafer Experts 91

4.4.3 Tasking 93

4.4.4 Anomaly Discrimination 95

4.4.5 Custody Reacquisition 97

4.5 SSA Sensor Evidence 98

4.5.1 Uncertainty and Reachability 99
4.5.2 Cloud Detection 100
4.5.3 Sky Brightness 102
4.6 Simulation Results 103
4.6.1 Case 1: Clear Observation, Nominal Dynamics 108
4.6.2 Case 2: Unclear Observation, Nominal Dynamics 111
4.6.3 Case 3: Clear Observation, Anomalous Dynamics 113
4.6.4 Case 4: Unclear Observation, Anomalous Dynamics 114
4.6.5 Comparison to Covariance-Minimization 114
4.7 Conclusions 117

Chapter 5: Evidence Gathering for Hypothesis Resolution using Judicial Evidential Reasoning 131

5.1 Ambiguity Aversion 132
5.1.1 Plausibility Transformation for Decision-Making 132
5.1.2 Entropy for Decision-Making 133
5.2 Hypothesis Abstraction 134
5.3 General Evidence-Gathering Problem Definition 135
5.4 Evidence-Gathering for Hypothesis Entropy Reduction 137
5.4.1 Normalized Jirousek-Shenoy Entropy 139
5.4.2 Optimization Formulation 140
5.4.3 Computational Complexity 140
5.5 Implementation Considerations 141
5.5.1 Unbiased Hypothesis Resolution 142

5.5.2 Sub-Problem Definition 143
5.5.3 Combating Confirmation Bias 145
5.5.4 Resolving Combined Schedule Incongruity 146
5.5.5 Efficient Minimax Optimization 147
5.5.6 Hypothesis Pruning via Entropy Stopping Condition 148
5.5.7 Judicial Evidential Reasoning Summary 149
5.6 Examples 153
5.6.1 Case 1: Single JER Agent-Pair 153
5.6.2 Case 2: Multiple JER Agent-Pairs 156
5.6.3 Example Summary 159
5.7 Simulation Results 159
5.7.1 Scenario Description 159
5.7.2 Dynamics 160
5.7.3 Belief Function Models 161
5.7.4 Evidence to Belief Function Mappings 163
5.7.5 Case 1: Nominal Maneuver 164
5.7.6 Case 2: Propulsion Non-Start 165
5.7.7 Case 3: Propulsion System Explosion 166
5.7.8 Case 4: Collision with Object in LEO 168
5.7.9 Comparison to Entropy-Greedy Scheduler 171
5.7.10 Discussion 173
5.8 Conclusion 174

Chapter 6: Development of a Prototype DSS for SSA 176

6.1 Design Requirements Addressed 176

6.2 DSS Design 178

6.2.1 Sensor Tasking Schedulers 179

6.2.2 Simulation Environment 180

6.3 Human-in-the-Loop Data Collection 181

6.4 Test Scenarios 184

6.5 Results and Discussion 190

6.5.1 NASA-TLX 190

6.5.2 Cognitive Support 191

6.6 Conclusions 193

Chapter 7: Future Work 200

7.1 Non-SSA Applications of JER 200

7.2 Improved Confirmation Bias Metrics in JER 201

7.3 Principle of Equal Effort and Asynchronous Evidence Gathering 202

7.4 Further Application of Decision Support Requirements 202

7.5 Hypothesis Generation and Evidence Mapping 203

Chapter 8: Conclusions 204

Appendix A: Anomaly Detection Sensitivity Study Results 209

Appendix B: SSA DSS Human-in-the-Loop Experiment Materials 216

B.1 Informed Consent 216

References 230

LIST OF TABLES

1.1 Related literature and relevance 12

3.1 Boundary condition orbital elements for GMM inclination change 62

3.2 Timing comparison for validation scenario 63

3.3 Timing comparison by algorithm segment 65

4.1 Simulated sensor locations and optical parameters 105

4.2 Simulated space object parameters 106

5.1 Ellsberg's paradox belief structures and entropy 134

5.2 Example Case 1: Basic probability assignments for diagnostic tests 154

5.3 Example Case 2: Basic probability assignments for diagnostic tests 156

5.4 Example Case 2: Results based on true hypothesis realization 158

6.1 Summary of hypotheses and propositions for simulations 185

6.2 Sensor parameters, scenario A 185

6.3 Space object parameters, scenario A 186

6.4 Hypothesis parameters, scenario A 186

6.5 Sensor parameters, scenario B 186

6.6 Space object parameters, scenario B 186

6.7 Hypothesis parameters, scenario B 187

6.8 Sensor parameters, scenario C 187

6.9 Space object parameters, scenario C 187

6.10 Hypothesis parameters, scenario C 187

6.11 Sensor parameters, scenario D 188

6.12 Space object parameters, scenario D 188

6.13 Hypothesis parameters, scenario D 188

6.14 Sensor parameters, scenario E 189

6.15 Space object parameters, scenario E 189

6.16 Hypothesis parameters, scenario E 189

LIST OF FIGURES

1.1 Monthly number of objects in Earth orbit by object type (NASA Orbital Debris Quarterly News, February 2018) 2

1.2 Sensors in the Space Surveillance Network (SSN) 3

1.3 General sensor tasking frameworks 9

2.1 Abstraction hierarchy model decomposition (Naikar, 2013) 22

2.2 Abstraction hierarchy of the SSA work domain 27

2.3 Abstraction hierarchy of the SSA environment 29

2.4 Control task analysis decision ladder template, adapted from Vicente [59] . 33

2.5 Requirements derivation for SSA information fusion for hypothesis resolution 35

2.6 Requirements derivation for SSA sensor allocation for hypothesis resolution 37

3.1 Maneuver detection scenario 44

3.2 Minimum-fuel control distance anomaly detection and characterization framework 45

3.3 Partitioned PDF using multiple trajectories 52

3.4 One-sided Binary Hypothesis Testing Illustration 54

3.5 Theoretical scenarios for control cost binary hypotheses 55

3.6 Theoretical scenarios for Mahalanobis distance binary hypotheses 58

3.7 Comparison of control cost distributions using the non-linearized and approximate methods 63

3.8 Mahalanobis distance distributions for non-Gaussian boundary conditions . 64

3.9 Probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated inclination change maneuver, $P_{FA} = 0.05$. 67

3.10 Probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated phasing maneuver, $P_{FA} = 0.05$. 70

3.11 Cross-track acceleration for Galaxy 15 satellite, real data 71

3.12 Probability of anomaly contours vs uncertainty scaling parameter and observation gap, real-data (WAAS) inclination change maneuver, $P_{FA} = 0.05$. 71

4.1 Uncertainty and reachability volumes interrogated by field of view 99

4.2 Reachability trendlines fit to reference data [84, 85] 100

4.3 All-Sky cloud detection 102

4.4 Propagated orbit and sensor geometries (triangles and boxes indicate simulation initial and final positions, respectively) 107

4.5 Case 1 - Sensor tasking schedule and anomaly ignorance 112

4.6 Case 1 - Belief assignment 119

4.7 Case 2 - Sensor tasking schedule and anomaly ignorance 120

4.8 Case 2 - Belief assignment results 121

4.9 Case 3 - Sensor tasking schedule and anomaly ignorance 122

4.10 Case 3 - Belief assignment results 123

4.11 Case 4 - Sensor tasking schedule and anomaly ignorance 124

4.12 Case 4 - Belief assignment results 125

4.13 Case 1: Covariance-minimization 126

4.14 Case 1: Covariance-minimization sensor tasking schedule and anomaly ignorance . . . 126

4.15 Case 1: Covariance-minimization (limited) . . . 127

4.16 Case 1: Covariance-minimization (limited) schedule and anomaly ignorance 127

4.17 Case 4: Covariance-minimization . . . 128

4.18 Case 4: Covariance-minimization sensor tasking schedule and anomaly ignorance . . . 128

4.19 Case 4: Covariance-minimization (limited) . . . 129

4.20 Case 4: Covariance-minimization (limited) sensor tasking schedule and anomaly ignorance . . . 129

4.21 Anomaly hypothesis ignorance comparison for ignorance-reduction and covariance-minimization methods. . . . 130

5.1 Predictive and reactive evidence-gathering. . . . 135

5.2 Sample alpha-beta pruning . . . 148

5.3 Judicial Evidential Reasoning algorithm . . . 152

5.4 Example Case 1: Brute force evaluation tree . . . 155

5.5 Example Case 1: Alpha-beta evaluation tree . . . 155

5.6 Nominal GTO transfer orbit and target GEO orbit. . . . 161

5.7 Possible causes for GTO insertion failure . . . 161

5.8 Case 1: nominal maneuver (baseline) . . . 165

5.9 Case 1: nominal maneuver (baseline), hypothesis resolutions (solid line for belief, dashed line for plausibility) . . . 166

5.10 Case 2: propulsion non-start . . . 167

5.11 Case 2: propulsion non-start, hypothesis resolutions (solid line for belief, dashed line for plausibility) . . . 168

5.12 Case 3: propulsion explosion 169

5.13 Case 3: propulsion explosion, hypothesis resolutions (solid line for belief, dashed line for plausibility) 169

5.14 Case 4: collision in LEO 170

5.15 Case 4: collision in LEO, hypothesis resolution 170

5.16 Entropy-greedy scheduler: propulsion explosion 172

5.17 Entropy-greedy scheduler: propulsion explosion, hypothesis resolutions (solid line for belief, dashed line for plausibility) 172

6.1 Screen-shot of the prototype SSA DSS 182

6.2 Box-and-whisker plot illustration 190

6.3 NASA-TLX questionnaire response statistics, prompts 1-3 195

6.4 NASA-TLX questionnaire response statistics, prompts 4-6 196

6.5 Cognitive support questionnaire response statistics, prompts 1-3 197

6.6 Cognitive support questionnaire response statistics, prompts 4-6 198

6.7 Cognitive support questionnaire response statistics, prompts 7-8 199

A.1 Control distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated inclination change maneuver. . . 210

A.2 Mahalanobis distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated inclination change maneuver. 211

A.3 Control distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated phasing maneuver. 212

A.4 Mahalanobis distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated phasing maneuver. 213

A.5 Control distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, real-data (WAAS) inclination change maneuver. 214

A.6 Mahalanobis distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, real-data (WAAS) inclination change maneuver. 215

B.1 Human subject study consent form approved by Georgia Tech Institute Review Board - page 1 217

B.2 Human subject study consent form approved by Georgia Tech Institute Review Board - page 2 218

B.3 Human subject study consent form approved by Georgia Tech Institute Review Board - page 3 219

SUMMARY

Situation awareness is the perception of elements in the environment, comprehension of their meaning, and projection of their status into the future. Space situational awareness (SSA) is particularly concerned with accurately representing state knowledge of space objects to resolve potential threats, such as collision. Tracking techniques used in the space surveillance system still rely largely on models and applications from the 1950s and 1960s, while the number of tracked objects continues to grow with improved sensor technologies and ease-of-access to space. This work frames the SSA sensor tasking problem to interrogate specific hypotheses using evidential reasoning. First, cognitive systems engineering practices are applied to derive cognitive work and information relationship requirements for SSA decision-support systems and provide insight on the utility of hypothesis-based methods in SSA. To evaluate hypothesis-based methods for SSA, the spacecraft anomaly detection problem is formulated as a binary hypothesis test using distance metrics while accounting for non-Gaussian boundary conditions to improve applicability to non-linear orbital dynamics. Next, a sensor tasking criterion is developed to gather the evidence that minimizes ambiguity, or ignorance, in hypothesis resolution. The application of evidential reasoning provides a rigorous framework for quantifying ambiguity and allows inclusion of diverse SSA sensors. Building upon this method, a generalized evidence-gathering framework, Judicial Evidential Reasoning (JER), is proposed for hypothesis resolution tasks. JER also accounts for confirmation bias by applying a principle of equal effort. Resource allocation is a non-linear, high-dimensional, mixed-integer problem, so JER also applies adversarial optimization techniques to address computational tractability concerns. Finally, a prototype SSA decision support system is developed based on the derived requirements to evaluate workload and situation awareness impacts of hypothesis-based tasking. This work aims to enable predictive sensor tasking to provide decision-quality information and improve decision-maker situation awareness and workload.

CHAPTER 1
INTRODUCTION AND MOTIVATION

In the 60 years since the launch of Sputnik 1, humans have steadily learned to leverage the "final frontier" to suit our needs. Currently, there are over 20,000 trackable objects in the space object catalog [1, 2] ranging from decommissioned rocket bodies to active telecommunications assets to university science and technology experiments (see Fig. 1.1). These numbers are expected to grow significantly due to improved tracking capabilities, new launches, and continued debris generation [3]. While Earth orbit is a vast volume, useful or strategic orbit regimes (e.g. low Earth orbit (LEO), Geostationary Earth Orbit (GEO), sun-sychronous LEO) have quickly become congested and contested [4]. With such diverse involvement in the space arena, there is a large economic and national security incentive to understand the space environment to ensure continued operation of assets.

Space situational awareness (SSA) is concerned with accurately representing the state knowledge of objects in the space environment to provide better prediction capabilities for threats such as potential conjunction events. Maintaining SSA is essential to the command and control missions of the Joint Space Operations Center (JSpOC) [7]. Discourse and activity in SSA increasingly focuses on decision-making in the presence of limited resources, uncertain information, and a contested space environment. The number of trackable space objects is continually growing with expanded use of small spacecraft technologies [8] and increased sensor capabilities. Growing clutter poses safety concerns, accentuated by the high-profile LEO collision event in 2009 between a COSMOS satellite and an active Iridium satellite [9]. Future SSA sensor tasking needs to focus on providing decision-makers with "actionable knowledge required to predict, avoid, deter, operate through, recover from, and/or attribute cause to the loss and/or degradation of space capabilities and services" [4].

Furthermore, establishing protocols and regulations in the use of space depends upon

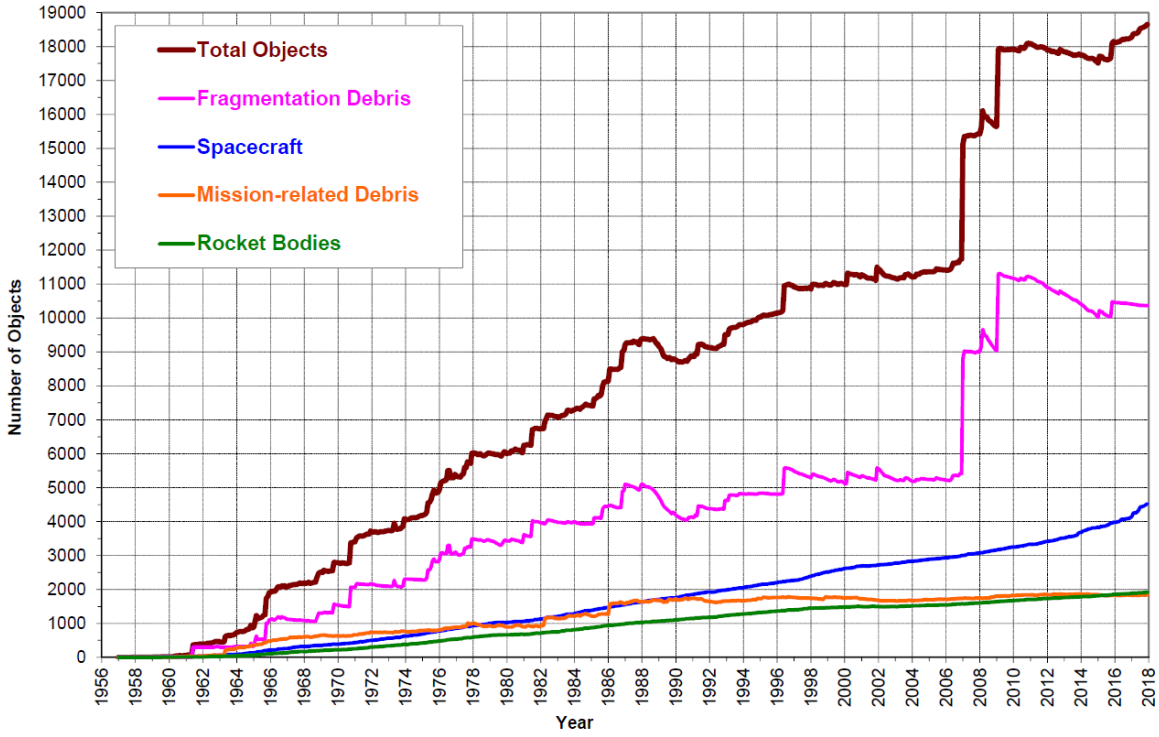


Figure 1.1: Monthly number of objects in Earth orbit by object type (NASA Orbital Debris Quarterly News, February 2018)

the "availability of quantifiable and timely information regarding the behavior of resident space objects" [4]. Constraints imposed by non-linear orbital dynamics and the disparity between the number of space objects and the number of sensors hinder the ability to reliably provide information on maneuvers or other events. Integrating and fusing non-traditional sensor data is crucial for SSA [10, 7], and increasing emphasis is being placed on algorithms and processes that have an ability to ingest disparate data from many sources and fuse an understanding of the greater situation of the space domain.

Figure 1.2 shows the distribution of sensors in the Space Surveillance Network (SSN) that are used regularly to track RSOs in the near-Earth environment (extending beyond geosynchronous orbit). Tracking techniques used in space surveillance still largely rely upon models and applications from the 1950s and 1960s [11]. As access to space becomes more affordable and the space object population increases, the amount of data required to maintain situational awareness greatly increases [12]. Increased data needs make the

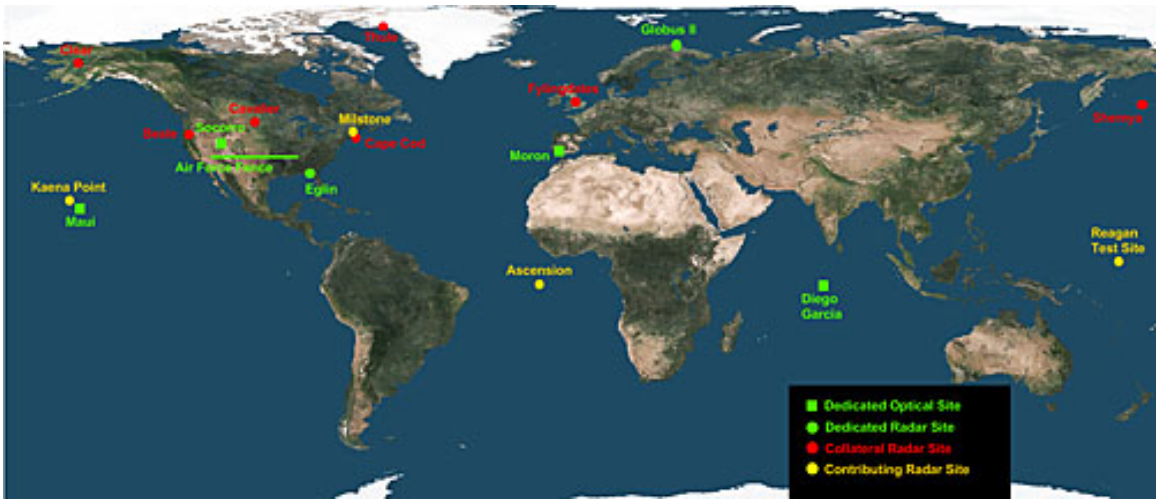


Figure 1.2: Sensors in the Space Surveillance Network (SSN)

human-in-the-loop involvement in space surveillance particularly troublesome, motivating the development of autonomous tasking capabilities. For instance, current space object custody tasking requires human analysis of candidate tasking schedules while incorporating constraints such as observation conditions (e.g. sky brightness, cloud cover). In the event that an object is not detected, a human analyst may be required to inspect the observation conditions visually before declaring lost custody or anomaly. This approach is reactive and rigid; however, adaptive approaches, such as the Dynamic Data Driven Application System (DDDAS) paradigm, can create more flexible algorithms that can incorporate additional data at runtime [13]. This necessitates a more automated approach to data collection and processing that incorporates auxiliary sensor data to operate in a more predictive manner and dynamically adjust the algorithm objectives and actions.

The availability of low-cost, high-accuracy, steerable sensors [15] provides an opportunity for improved sensor management, at both the individual sensor and network levels, to improve SSA knowledge and capabilities [16]. DDDAS-like algorithms operated on these networks enable predictive data collection to improve capabilities for space object custody, anomaly detection, new object detection, and resolution of other SSA hypotheses.

The task of gathering evidence to support claims or resolve hypotheses naturally extends far beyond SSA to awareness and decision-making in any domain. A decision-

maker's cognitive frame is determined by the things that are known, things that are not known, things that cannot be known, and things the decision-maker does not want to know [17]. Former United States Secretary of Defense Donald Rumsfeld famously articulated a similar classification in a February 2002 briefing: known knowns, known unknowns, and unknown unknowns [18]. National security and intelligence professionals have long used an analysis technique known as the Johari window [19] for classification into categories for open (known knowns), hidden (unknown knowns), blind (known unknowns), and unknown (unknown unknowns) information. When it comes to decision-making, knowledge and non-knowledge are equally constructive; while known threats may lead to negative outcomes, ignorance is also identified as a distinct kind of danger [17].

Known unknowns result from phenomena which are recognized but poorly understood; on the other hand, unknown unknowns are phenomena which cannot be expected because there has been no prior experience or theoretical basis for expecting the phenomena [20]. In operation, sensor effort must be allocated to the resolution of both. Resolving known unknowns leverages existing understanding to gain more specific knowledge with gathered evidence. This can be approached with predictive tasking schema, hypothesizing events and leveraging knowledge of the dynamics and uncertainties to reduce the search space and gather the most useful information. Unknown unknowns may still manifest as very real threats, and only very general searching strategies over a wide search space may be used. This type of tasking is considered reactionary, relying on specific evidence of the event before it can be further interrogated and resolved. Both approaches attempt to transition unknown information into the known-known category, as outlined in Fig. 1.3(b).

The application of rigorous evidence fusion techniques that resolve specific threat hypotheses has the potential to address data scarcity problems and improve predictive tasking for SSA. Modern information fusion techniques, such as evidential reasoning and Dempster-Shafer theory [21], provide an opportunity for improved hypothesis resolution SSA sensor tasking is a high-dimensional, multi-objective, non-linear, mixed-integer opti-

mization problem, so the application of modern high-dimensional and combinatorial optimization techniques enables efficient solution of tractable SSA problems.

Thesis Statement: Formulating the SSA sensor tasking problem to interrogate specific hypotheses using evidential reasoning yields decision-quality information, enables predictive tasking, and improves decision-maker situation awareness and workload.

1.1 Binary Hypothesis Testing for Anomaly Detection

Timely detection of anomalous events is crucial to responsiveness in follow-up tracking and post-anomaly orbit characterization [22]. Detecting low-thrust maneuvers is especially difficult since subtle state changes may propagate into vast state discrepancies [23, 24]. Data association hypothesizes an association between two uncorrelated tracks (UCTs), and the correlation likelihood can be assessed using numerous methods, including batch least-squares [23, 25], finite-set statistics [26], and admissible regions [27, 28, 29].

Distance metrics, such as Mahalanobis distance, can also be employed to quantify a state discrepancy between UCTs [30]. The minimum-fuel control distance metric, in particular, allows for data association and maneuver detection while incorporating propulsive effort, allowing reconstruction of the maneuvers [24]. Previous work develops a control cost distributions distribution and compares the mean maneuver size to the uncertainty in the control cost distribution [24]. This contribution addresses shortcomings of the original formulation in the case of non-Gaussian boundary conditions and limits in the error rates (e.g. false detection, missed detection). Non-Gaussian boundary conditions arise, even from initially Gaussian uncertainty distributions, due to propagation through non-linear orbital dynamics. Gaussian mixtures allow modeling of arbitrary non-Gaussian distributions while still preserving the computational advantage of modeling distributions as opposed to particle-filters. Additionally, a binary hypothesis testing formulation allows prescription of allowable false alarm rates.

Contribution 1: A binary hypothesis testing approach to anomaly detection with non-

Gaussian boundary conditions for both control distance and Mahalanobis distance metrics.

1.2 Ignorance-Reduction Criterion for Sensor Tasking

In general, the sensor tasking or sensor scheduling problem addresses how to obtain, process, and utilize information about the state of the environment [31]. Potential SSA needs include maintaining catalogs of space object state observations [26, 16], detecting maneuvers or other anomalies [32], and estimating control modes or behavior [33, 4]. The SSA sensor tasking problem is a high-dimensional, multi-objective, mixed-integer, nonlinear optimization problem, so current approaches focus on tractable sub-problems (e.g. single objectives or hypotheses, limited target objects, limited sensors). For instance, maintaining a catalog of space object estimates requires observations of many different space objects. Information-maximizing methods, as characterized through covariance estimates, minimize state estimate uncertainty for all catalog objects [34, 26]. Other objectives may require more data of specific targets or events. Space object association may be handled by quantifying a state anomaly or maneuver required to associate two uncorrelated tracks (UCTs) [24, 32], classification methods may employ taxonomies trained on representative space object feature sets to categorize space objects [35], and attitude or control mode estimation requires many observations of a single object to develop a light curve, a time-history of photons received from the target space object [97]. These competing objectives are generally not complementary, especially given limited sensor resources, so different objectives may prefer different tasking approaches.

Many existing sensor tasking approaches aim to maintain low overall uncertainty (e.g. information-maximum), but this tasking does not necessarily support the needs of a decision-maker. This motivates an approach that encodes tasking opportunities and decision-making priorities as hypotheses that can be interrogated by evidence. If a potential target's orbit and operational capabilities (or lack-thereof) are well-known, it might not be necessary to minimize its associated uncertainty. Conversely, many consecutive follow-ups might be de-

sired on a newly-acquired object to fully characterize its orbit, or on an object approaching a congested volume of space (such as a GTO object approaching apogee). Hypothesis-driven approaches are not new to SSA; for instance, multiple hypothesis testing (MHT) techniques have been applied to object detection within electro-optical images [36, 37, 38]. Applied to sensor-tasking, hypothesis-driven approaches enable a predictive mode of tasking to answer specific relevant questions.

Additionally, sensor information must be fused into a coherent understanding of the environment via association, correlation, and combination [31]. In classical Bayesian approaches, sensor data is used to form deterministic probabilities placed on event hypotheses under the assumption that the only possible realizations of this hypothesis are true or false. However, in complex decision-making contexts, information is not always represented well in a strictly binary manner, needing to account for uncertain information and ambiguity. An expert might be able to confirm or refute a given set of hypotheses, but it cannot attribute belief to any hypotheses for which it is not an expert. For this reason, evidential reasoning methods, such as Dempster-Shafer theory, quantify ambiguity, leading to more realistic modeling of human analyst processes [39, 40, 41]. Dempster-Shafer theory has gained significant traction in various applications, including classification [42, 43], monitoring and fault detection [44, 45], and decision-making [46].

This contribution formulates a tasking criterion based on the reduction of residual ambiguity, gathering evidence that yields a more precise understanding of the relevant hypotheses. In comparison to a covariance-based sensor tasking scheme, the ignorance-reduction approach is able to resolve the anomaly and custody hypotheses to equal quality using far fewer observations. When the number of actions is constrained to be the same between the algorithms, ignorance-reduction chooses actions that gather stronger evidence, resolving the hypotheses quicker and to better quality. Additionally, methods are presented for reliably encoding and fusing SSA sensor data as evidence experts.

Contribution 2: A sensor tasking ignorance-reduction criterion and a framework using

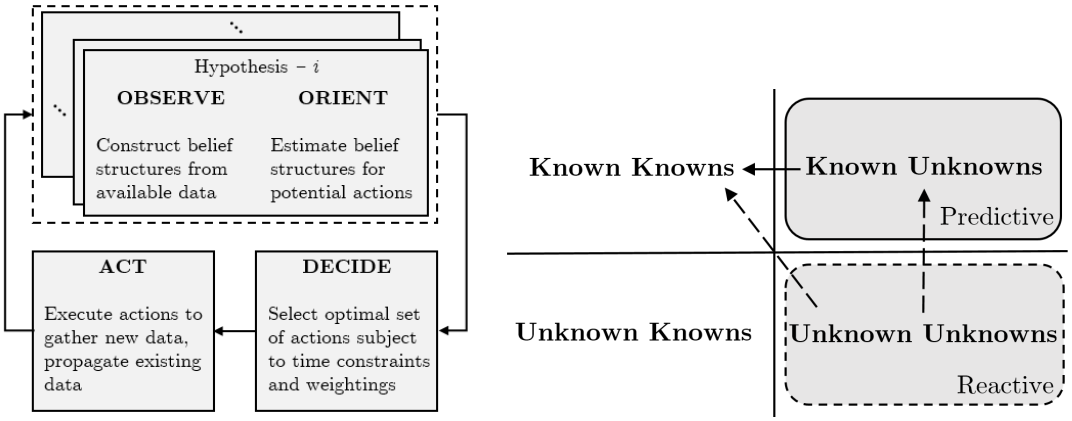
SSA sensors as evidence experts to resolve specific hypotheses.

1.3 Evidence-Gathering for Hypothesis Resolution using Judicial Evidential Reasoning

The curse of dimensionality is one of the primary criticisms of the Dempster-Shafer approach [47]; the computational complexity of evidence combination generally increases exponentially with the number of hypotheses [40]. Approximations [48] and exact implementations [40] have shown linear complexity for hierarchical evidence. Monte-Carlo methods [49] and consonant methods on chain-like hypothesis structures [50] enable feasible implementations of the Dempster-Shafer formalism. Comparisons of these approximation methods for decision-making find that there is no definitive "best" approximation as they often restrict application to cases with low conflict in evidence or collapse the belief intervals into points, eliminating the ability to represent ambiguity [51]. In order to be applicable in a decision-making scenario, with real operational objectives and time constraints, an evidential-reasoning-based tasking framework must be able to support many hypotheses and reliably support human decision-making cognition.

In command and control scenarios, an elegantly simple but extensively applicable model of rational human behavior is the Observe-Orient-Decide-Act (OODA) loop [52, 53]. This ongoing cycle encompasses gathering data from observations, processing the data to understand the situational reality, making decisions from this understanding, and acting on these decisions [53]. The OODA loop is readily applicable to the hypothesis-based evidence-gathering problem, as shown in Fig. 1.3(a): measurements are gathered and processed into belief structures based on the hypotheses of interest.

The ignorance-reduction criterion developed in the previous contribution is well-suited to resolve relatively low numbers of hypotheses, applicable to the task of maintaining custody of a subset of satellites. For general operational applications in SSA, with over $1,000$ active assets and nearly $19,000$ total tracked objects, and there may be tens or hundreds of



(a) OODA loop descriptions. (b) Reactive and predictive tasking.

Figure 1.3: General sensor tasking frameworks

hypotheses of interest for each space object. With so many potential tasking options and so many hypotheses to investigate, as well as time constraints on responding to some opportunities or threats, the computational complexity of a brute-force approach is daunting. Therefore, the larger sensor allocation problem is decomposed into separate sub-problems to address each hypothesis, and the sub-problem solutions are combined to quickly find a near-optimal tasking [54]. This framework also leverages game-theoretic adversarial optimization to improve computational complexity over a brute-force approach.

The developed approach, called Judicial Evidential Reasoning (JER), hinges upon three primary considerations: hypothesis abstraction, ambiguity aversion, and confirmation bias. The use of a hypothesis abstraction supports human decision-making strengths of planning and strategy, off-loading processing work to the algorithm and fusing evidence into intuitive hypothesis resolutions. Recognizing the need to account for ambiguity aversion in decision-making, the use of Dempster-Shafer theory allows for quantification of evidence ambiguity. Finally, applying a principle of equal effort through an alternating-turn adversarial optimization scheme avoids confirmation bias induced by improper prior beliefs or evidence uncertainty and ambiguity, avoiding fixation on incorrect propositions.

Contribution 3: A generalized evidence-gathering approach to multi-hypothesis resolution, applied in both non-SSA and SSA scenarios, that efficiently computes near-optimal

evidence-gathering schedules while accounting for ambiguity and confirmation bias.

1.4 Cognitive Systems Engineering Applied to Decision Support in SSA

In order to be applicable in real operational scenarios, a proposed sensor tasking framework and associated decision-support system must reliably support human cognition and expertise. In application, the upper and lower bounds of the belief function may not provide intuitive usefulness in presentation to an analyst or decision-maker. Typically the belief structure is either converted to a probability distribution (as a Bayesian approximation of the belief structure) [55] or collapsed to a probability formation that allows a decision-maker to place bets on each hypothesis given the available evidence using familiar Bayesian constructs [56, 57]. However, the previous contributions showed the usefulness in incorporating ambiguity in hypothesis resolution for decision-making in sensor scheduling. This contribution aims to show the added cognitive benefits of a hypothesis-based tasking abstraction, as well as the effects of explicitly and rigorously quantifying hypothesis resolution quality on decision-maker behavior.

This work applies cognitive systems engineering practices [58], primarily cognitive work analysis [59], to identify purposes, capabilities, and constraints in the SSA work domain. A work domain analysis identifies capabilities and constraints present within both the SSA operator's work domain and the SSA environment as a whole, and a control task analysis aids in the derivation of a set of decision-support design requirements for information fusion and sensor allocation tasks. A prototype decision-support system that leverages these requirements is also presented and evaluated.

Contribution 4: Application of cognitive systems engineering techniques to derive requirements for a prototype SSA decision-support system, as well as analysis of operator situation awareness and workload using this prototype.

1.5 Organization and Relevant Literature

The contributions of this work aim to improve SSA decision-making through hypothesis-based, predictive sensor tasking. Table 1.1 summarizes the relevant literature for each contribution.

The thesis begins with the cognitive systems engineering analysis from the fourth contribution to provide context on goals and challenges in the SSA work domain before delving into specific approaches and applications. The third chapter addresses the first contribution, developing a binary hypothesis testing approach for anomaly detection with non-Gaussian boundary conditions using two different distance metrics: Mahalanobis distance and control distance. The fourth chapter addresses the second contribution, incorporating evidential reasoning to fuse SSA sensor data and presents an ignorance-reduction criterion for sensor tasking. The fifth chapter addresses the third contribution, which extends the evidential reasoning approach to develop a generalized evidence-gathering framework to resolve multiple hypotheses while mitigating computational complexity, evidence ambiguity, and confirmation bias. The sixth chapter returns to the cognitive systems engineering work in the first contribution, leveraging the design requirements to develop a prototype decision support system for SSA relevant to the hypothesis-based methods in the previous contributions, investigating decision-maker situational awareness, cognitive support, and workload.

1.6 List of Publications

The publications resulting from work related to this thesis are listed below in chronological order, separated by publication type.

Peer-Reviewed Journal Articles

J.1 A. D. Jaunzemis, M. V. Mathew, and M. J. Holzinger, "Control Cost and Mahalanobis Distance Binary Hypothesis Testing for Spacecraft Maneuver Detection,"

Table 1.1: Related literature and relevance

		UCT Correlation and Anomaly Detection	Evidence-based Hypothesis Resolution	Efficient Many-Hypothesis Resolution	Cognitive Systems Engineering for SSA
Existing Literature (by reference index)	Holzinger et al. [24]	*Control distance metric, single-hypothesis anomaly detection*			
	Hill et al. [30]	*Mahalanobis distance spacecraft anomaly detection*			
	Hobson [16]		*Sensor management for RSO custody*	*Unequal cost multiple hypothesis testing for space surveillance telescope*	
	Dempster [41], Shafer [39], Yager [75], Smets et al [57]		*Develops probability intervals, DST, TBM, pignistic belief*	*Approximation methods for evidence combination with hierarchical evidence, Monte-Carlo and consonant methods for computational efficiency*	
	Hussein et al. [34]		*Covariance-reduction sensor tasking*	*Observe-orient-decide-act loop*	
	Fruh and Jah [35]		*Taxonomy for space object classification*	*Integer and combinatorial optimization*	
	Zingarelli et al. [36], Hardy et al. [37]		*Multiple-hypothesis testing for space surveillance telescope*	*Ambiguity aversion*	
	Gordon and Shortliffe [48], Shafer and Logan [40], Kreinovich et al. [49], Dubois and Prade [50]			*Dempster-Shafer entropy definitions*	*Decision support system guidelines*
	Boyd [52], Angerman [53]				*Cognitive work analysis, knowledge representation for decision-making*
	Nemhauser and Wolsey [54]				*Applications of cognitive work analysis, decision support requirement derivation*
	Ellsberg [94]				*Human-centric SSA*
	Jirousek and Shenoy [74]				*JSpOC cognitive task analysis*
	Keen [105], Shim [106]				
	Rasmussen [115, 122], Vicente [59]				
	Burns et al [119], Miller et al. [104]				
	Ianni et al. [120]				
	Aleva and McCracken [121]				
My Publications	SFM '15 JGCD '16	*Non-Gaussian BHT for anomaly detection, control and Mahalanobis*	*Anomaly probability for custody discrimination and re-acquisition*	*Hypothesis-based sensor allocation via ignorance-reduction*	*Decision-quality hypothesis knowledge representation using entropy-reduction, ambiguity aversion*
	SFM '16 AMOS Tech '16 JAIS '18		*Ignorance-reduction criterion, SSA sensors as DST experts*	*Tractable evidence-gathering for multi-hypothesis resolution via JER*	
	ICSSA '17 FUSION '18 JIF '18				*SSA decision support requirements derivation, analysis of prototype*
	AMOS Tech '18 JCEDM '18				

Journal of Guidance, Control, and Dynamics, 2016, Vol. 39, No. 9, pp. 2058-2072, doi:10.2514/1.G001616

J.2 A. D. Jaunzemis, M. J. Holzinger, and K. Kim Luu, "Sensor Tasking for Spacecraft Custody Maintenance and Anomaly Detection using Evidential Reasoning," Journal of Aerospace Information Systems, 2018, Vol. 15, No. 3, pp. 131-156, doi:10.2514/1.I010584

J.3 A. D. Jaunzemis, M. J. Holzinger, M. W. Chan, and P. P. Shenoy, "Evidence Gathering for Hypothesis Resolution using Judicial Evidential Reasoning," Information Fusion, submitted March 2018.

J.4 A. D. Jaunzemis, K. M. Feigh, M. J. Holzinger, and M. W. Chan, "Cognitive Systems Engineering Applied to Decision Support in Space Situational Awareness," Journal of Cognitive Engineering and Decision Making, to be submitted May 2018

Conference Presentations

C.1 A. D. Jaunzemis, M. V. Mathew, and M. J. Holzinger, "Control Metric Maneuver Detection with Gaussian Mixtures and Real Data," 25th AAS/AIAA Space Flight Mechanics Meeting, Williamsburg, VA, January 2015.

C.2 A. D. Jaunzemis, M. J. Holzinger, "Evidential Reasoning Applied to Single-Object Loss-of-Custody Scenarios for Telescope Tasking," 26th AAS/AIAA Space Flight Mechanics Meeting, Napa, CA, February 2016.

C.3 A. D. Jaunzemis, M. J. Holzinger, "Evidence-Based Sensor Tasking for Space Domain Awareness," Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, HI, September 2016.

C.4 A. D. Jaunzemis, D. Minotra, M. J. Holzinger, K. M. Feigh, M. W. Chan, P. P. Shenoy, "Judicial Evidential Reasoning for Decision Support Applied to Orbit In-

sertion Failure," 1st IAA Conference on Space Situational Awareness, Orlando, FL, November 2017.

C.5 A. D. Jaunzemis, M. J. Holzinger, M. W. Chan, P. P. Shenoy, "Evidence Gathering for Hypothesis Resolution using Judicial Evidential Reasoning," submitted to 21st International Conference on Information Fusion, Cambridge, UK, July 2018.

C.6 A. D. Jaunzemis, D. Minotra, K. M. Feigh, M. J. Holzinger, "Cognitive Systems Engineering Applied to Decision Support in SSA," submitted to Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, HI, September 2018.

CHAPTER 2

DEVELOPMENT OF REQUIREMENTS FOR SSA DECISION SUPPORT

Research in space situational awareness (SSA) increasingly focuses on gathering and processing actionable knowledge, primarily "to predict, avoid, deter, operate through, recover from, and/or attribute cause to the loss and/or degradation of space capabilities and services" [4]. This chapter applies cognitive systems engineering methods to derive requirements for decision support in SSA.

2.1 Cognitive Systems Engineering

This section introduces relevant terminology and frameworks in cognitive engineering to provide a background for the development of requirements for supporting human-in-the-loop decision-making in SSA.

2.1.1 Situation Awareness

Endsley defines situation awareness as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" [5], or, more simply, as "knowing what is going on around you" [6]. Inherent in Endsley's definition of situation awareness is an understanding of what is important [6], so the design of support systems for situation awareness must begin with establishing goals and purposes of the work domain. Endsley makes an important distinction between knowledge states and information acquisition processes relevant to the discussion of situation awareness [6]:

> "Situation as defined above is a state of knowledge about a dynamic environment. This is different than the processes used to achieve that knowledge. Dif-

ferent individuals may use different processes (information acquisition methods) to arrive at the same state of knowledge, or may arrive at different states of knowledge based on the same processes due to differences in the comprehension and projection of acquired information or the use of different mental models or schemata."

Often, decision support systems are measured for effectiveness based on performance measurements relating a user's response to some input conditions. However, as an intermediate variable between the input and output of the system, "the measurement of situation awareness provides far greater diagnosticity than is typically available from performance measures" [6]. Measurements of situation awareness, along with traditional performance measures, provide more insight into the operator's cognitive processes and in-situ understanding of the work domain. In general, situational awareness can be quantified as the degree of correspondence between a set of human judgments and distribution of the true system or environmental states [100, 101], and similarly SSA is concerned with accurately estimating the distribution of orbital states (e.g. position and velocity), primarily for near-Earth objects. In order to be quantifiable, the cognitive process measurements must have a basis in observable variables: one method uses subjective reports of retrospective memory, another uses subjectively reported measures of situation awareness [102], and another uses probes embedded within simulation studies in real-time, interactive contexts, freeze simulation and query actors on state of controlled system [6, 103]. Therefore, situation awareness measurement and analysis is typically done through in-situ surveys conducted during human-in-the-loop experiments [125].

2.1.2 Decision Support Systems

Systems engineers face a difficult design problem in effectively managing complex sociotechnical systems, comprised of both the technological systems and the people or organizations that operate within them [104]. In the late 20th century, Keen presented a road-map

for decision support system (DSS) research, focusing on utilizing emerging software tools to build semi-expert artificial intelligence systems and emphasizing the value and role of experts in DSS [105]. Shim et al. updated Keen's agenda for the 21st century, highlighting increased proficiency with technology and an associated expectation of more functionality in DSS technology [106]. Shim encouraged researchers to identify areas where tools are needed to transform uncertain and incomplete data, along with qualitative insights, into useful knowledge, along with further exploiting software and technology tools.

When considering building a DSS, the design approach for the user interface must be carefully considered. The effectiveness of a decision aid depends on relationships between the representation, the domain and associated tasks, and the characteristics of the agent [107]. Designers cannot anticipate all the possible scenarios that could arise and must therefore design displays that support effective problem solving even when novel or unanticipated scenarios are encountered [107]. To this end, it is important to frame the goal of such a design task as first helping the user to focus attention on a potentially important event and then providing integrated displays that help the user to construct a deeper understanding of the context [108].

2.1.3 Cognitive Work Analysis

Militello et al. [58] define cognitive systems engineering (CSE) as "an approach to the design of technology, training, and processes intended to manage cognitive complexity in sociotechnical systems." CSE aims to provide the designer with "a realistic model of how the human functions cognitively" [109]. A multitude of CSE methodologies have emerged in recent decades to inform system design by modeling human cognitive functions [104], including cognitive work analysis [110, 59], contextual design [111], hierarchical task analysis [112], and naturalistic decision-making [113]. Unfortunately, the bulk of applications of CSE methods in industry has been limited to incorporating the insights as graphical user interface elements [104].

One of the more widely-adapted CSE frameworks is cognitive work analysis (CWA), which is a framework for establishing characteristics and constraints of the work domain [110, 59]. CWA differs from other types of work analysis by focusing more on how work may be driven by constraints imposed by the domain and less on how the work is actually accomplished [59]; in other words, the emphasis is on how the work could be done, not on how work is done or should be done [114]. It is therefore a useful design tool for a new system intended to support expert work, defined as the ability to compose a process needed for a specific task as a sequence of familiar subroutines that are useful in different contexts [115]. CWA provides an approach to "characterize the constraints that define the cognitive requirements and challenges, and the knowledge, skills, and strategies that underlie both expert performance and the error-vulnerable performance of domain practitioners" [116].

The traditional cognitive work analysis, as defined by Vicenti [59], consists of five phases or dimensions with different analysis boundaries, summarized below. The first phase, work domain analysis (WDA), analyzes relationships between the purposes, priorities, functions, and resources in the domain. The second phase, control task analysis (ConTA), analyzes activities in specific situations or tasks. The third phase, strategies analysis (SA), analyzes strategies for executing an activity. The fourth phase, social organization and cooperation analysis (SOCA), analyzes the distribution of work amongst individuals and teams, as well as communication required between these entities. The final phase, worker competencies analysis (WCA), analyzes the perceptual and cognitive capabilities and limits of humans in the domain. CWA research efforts to date predominantly implement the first two phases, leading them to be the most matured analysis techniques [114].

2.2 Work Domain Analysis Applied to SSA

This work begins by applying the first phase of CWA, the work domain analysis, to analyze the broader purposes of decision support in SSA and the means available to accomplish

those goals. WDA can be applied in to inform DSS design in a number of ways. For instance, the ecological design approach uses WDA results to inform specific user interface design solutions [117]. An alternate approach is to use the WDA results to derive requirements for supporting complex work in the domain [114]. Since well-understood goals are central to the development of situation awareness [6], this work follows the latter approach, applying the principles of WDA to uncover purposes, capabilities, and constraints within the SSA work domain.

2.2.1 Description of the Work Domain

The work domain relevant to SSA consists of numerous social and technical components. SSA is particularly concerned with accurately representing the state knowledge of objects in the space environment to provide better prediction capabilities for threats such as potential conjunction events. Operators in SSA operations centers, such as the Joint Space Operations Center (JSpOC), often start with limited training in orbital mechanics, sensor phenomenologies, data fusion, or other relevant SSA fields. The operators are then responsible for aggregating data on a diverse space object population, ranging from active satellites to orbital debris, and conduct analyses to predict events (e.g. conjunctions) or schedule follow-on observations to maintain a catalog of space object state estimates. This data is gathered from a diverse network of sensors, some or many of which may be controlled by entirely separate entities, which poses difficulties in gathering timely data on specific events. For instance, sensors in the space surveillance network (SSN) are geometrically diverse (recall Fig. 1.2) and relatively sparse, especially as compared to the number of RSOs tracked. For instance, collections from an object in low Earth orbit (LEO) from one particular radar sensor might occur over two 5-minute-long observable passes in one day. This leaves significant time wherein anomalies (e.g. maneuvers, on-orbit break-ups) can occur without being directly observed.

In general, the sensor tasking problem addresses how to obtain, process, and utilize

information about the state of the environment [31]. The SSA sensor tasking problem is a high-dimensional, multi-objective, mixed-integer, non-linear optimization problem, and current approaches focus on tractable sub-problems (e.g. single objectives, limited target objects, limited sensors). Potential SSA sensor tasking needs include maintaining catalogs of space object state observations [26, 16], detecting maneuvers or other anomalies [32], and estimating control modes or behavior [71, 72, 4]. These objectives are generally not complementary, especially given limited sensor resources, and the different objectives require different tasking approaches.

Discourse and activity in SSA increasingly focuses on decision-making in the presence of limited resources, uncertain information, and a contested space environment [4]. Therefore, the goal of aggregating all this data is to support SSA decision-making. Support for decision-making must provide quantifiable and timely evidence of behaviors related to specific hypotheses (e.g. threats). To support this hypothesis resolution activity, existing approaches largely focus on collecting observables to identify physical states or parameters. However, many complex hypotheses require RSO behavior prediction that takes into account other RSOs, physics knowledge, and indirect information from non-standard sources. As such, an active avenue of research in SSA focuses on the use of information fusion and emerging technologies to ingest varied, sparse datasets to form a coherent story of the space environment, as well as techniques for better conveying this story to decision-makers.

Additionally, presenting all relevant information (available sensors, location of target space objects and the most up-to-date information on them, and active working hypotheses on the space environment) to a decision-maker in SSA creates a big-data and data-visualization problem. Problematically, the collected data products are also affected by adverse observation conditions, uncertainties, biases, and unobservable states that may contribute to ambiguity in evidence. Advancements in the fusion of all relevant data, tracking data and operator information, into a coherent user interface that communicates available

courses of action provide an avenue for increased intuition in SSA decision-making.

The primary challenges of the SSA work domain can be summarized using dimensions of complexity adapted from Vicente [59]:

- **Large problem space:** High number of interacting variables.
- **Dynamic:** Constantly varying states, potentially long response times between measurement opportunities.
- **High-risk:** Errors may lead to catastrophic results.
- **Social:** Multiple organizations, with competing interests in the use of space, vying for SSA data.
- **Distributed:** Geographically disparate sensor networks and organizations.
- **Uncertainty:** Sensor bias, measurement noise, and unobservability result in probabilistic and less-than full-state knowledge.
- **Disturbances:** Operators are expected to understand anomalous behavior and bring system back within nominal conditions.
- **Automation:** Operators expected to monitor and intervene quickly and decisively in off-nominal conditions.

To derive guidelines that support decision-making and address these complexities, the following sections apply CWA techniques to the SSA domain.

2.2.2 Abstraction Hierarchy

Work domain analyses are employed as the first step of CWA to establish broad understandings and identify constraints that exist within the domain [104]. The result of a WDA is traditionally a model known as the abstraction hierarchy, which provides a graphical representation of linkages between purposes, priorities, functions, and resources of the domain

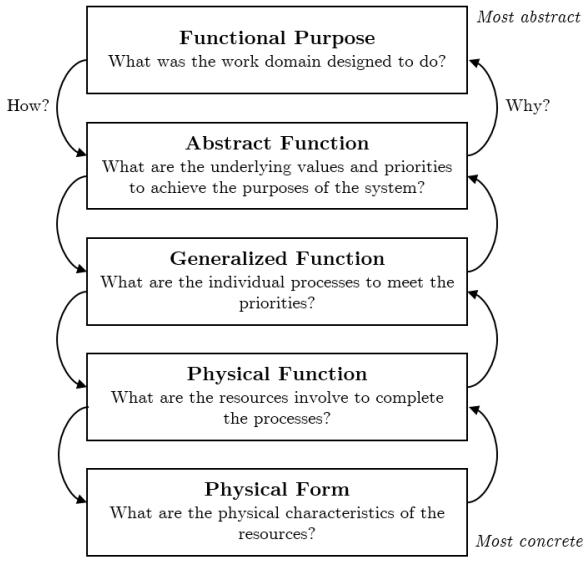


Figure 2.1: Abstraction hierarchy model decomposition (Naikar, 2013)

[115, 59]. The abstraction hierarchy assesses the means-ends relationships inherent to the work domain based on the purposes of the actor(s), but explicitly does not consider the actions of the actor. In particular, the abstraction hierarchy focuses on structural means-ends relationships within the objects in the work domain. By representing the work domain from multiple levels of abstraction, an analyst can view the domain at varying levels of detail. The functional purpose and other high level abstractions provide broad overviews of the system and intended goals, whereas the lower level abstractions define attributes of the physical objects with which the actor interacts.

The structural relationships between elements in adjacent levels of the hierarchy can be summarized as follows: each level simultaneously provides a means (the "how") to elements in the level above, and an end (the "why") to the level below. The traditional five levels of decomposition included in an abstraction hierarchy [115] are summarized in Fig. 2.1, adapted from [118]. Beginning with the overall purposes of the work domain and progressing down the decomposition levels, the elements in the hierarchy become more concrete, arriving at the physical characteristics of the domain resources.

The following summaries provide added detail on each level of the decomposition [115,

114]:

- **Functional Purposes:** reason(s) or purpose(s) of the system.
- **Abstract Functions:** principles or priorities of the work domain that are preserved, conserved, maximized, or minimized (e.g. conservation of energy).
- **Generalized Functions:** functions that must be present for the functional purpose of the work domain to be fulfilled.
- **Physical Functions:** capabilities of the physical elements within the work domain.
- **Physical Forms:** properties of physical elements within the work domain (e.g. physical form, configuration).

It is often useful to further decompose the work domain into multiple abstraction hierarchies to separately examine different aspects of the environment and domain activities at these varying levels of detail. For instance, Burns et al. [119] used a system of three abstraction hierarchies to model naval command and control, and Miller et al. [104] used two abstraction hierarchies to separately model the environment and work domain for extra-vehicular activities (EVA) with time-delay. Similarly, the SSA problem can be appropriately decomposed using two abstraction hierarchies, and the following sections will discuss the elements of these decompositions in greater detail.

2.2.3 SSA Work Domain Decomposition

The first system examined is the work domain for SSA operators. This decomposition is adapted from preliminary work [82], constructed through interviews with subject-matter experts with experience in SSA and decision support. The resulting modified decomposition is shown in Fig. 2.2.

Functional purposes: The overarching goals of SSA historically related primarily to maintaining safety of the more-than $1,300$ active satellites in Earth orbit. However, the US

Air Force has also recognized the space domain to be a valuable asset in maintaining national security concerns through communications and surveillance. Therefore, the purposes goals of any decision-support system in SSA can be summed up by ensuring continued safe operation in the space environment and addressing national security concerns.

Abstract functions: All objects in Earth orbit are subject to orbital dynamics, an active area of continued research. To a first order approximation, these dynamics are summarized using laws of Keplerian motion; however, increasing demands on spacecraft capabilities, and thereby sensing capabilities, necessitates higher-fidelity models that incorporate other environmental effects such as atmospheric drag and solar radiation pressure. The sensors used for gathering evidence are also subject to dynamics and phenomenology dependent upon the type of sensor used. For instance, radar sensors exploit different phenomenologies than electro-optical sensors, making radar more effective for range and range-rate measurements while electro-optical sensors are more effective for angle and angle-rate measurements. One of the primary tasks of decision-makers in SSA is to assign priorities to drive operations. In some cases, this means prioritizing data collection of certain assets, such as a military or communications satellite, over less-critical objects, such as university satellites or previously well-tracked objects. As the focus of SSA increasingly turns toward decision-making and gathering actionable information to resolve hypotheses, these hypotheses and their relative priorities also become an increasing focus of decision-support efforts. SSA operators must also be concerned with the efficiency of the sociotechnical system workflow: sensor and operator resources should neither be under-utilized or over-utilized.

Generalized functions: The primary functions required to successfully perform SSA activities are listed at the generalized functions level. SSA operators are often concerned with detecting events or anomalies that occur, a difficult task considering the disparity between the number of trackable objects and the number of available sensors. Pursuant to the space asset safety goal, operators are particularly concerned with predicting and avoiding potential conjunction or collision events. The primary means of gathering information to

address hypotheses related to the safety and security goals is through sensor allocation, which leverages both orbital dynamics and sensor phenomenologies to gather and fuse relevant evidence. Operators have many existing methods to choose from when it comes to processing detections to maintain a catalog of space object data, and techniques for correlating these detections is still an active area of research. The operators must also be aware of the fact that, due to anomalous events and non-linearities in the dynamical environment, the accuracy of this information is subject to degradation, requiring further data collection. Importantly, SSA operations centers are often not comprised of actual spacecraft or sensor operators. Instead, the SSA operators aggregate and fuse observation information to update the catalogs and predict events, while also generating lists of desired observations to address decision-making needs. Therefore, the dissemination of this information is crucial to successful SSA operation.

Physical functions: At the physical functions level, the functional capabilities and constraints imposed by the SSA work domain are listed. The computational resources required to perform associate tracks, update catalogs, and detect or predict events are a primary consideration toward the functional capabilities of any SSA operations center. Similarly, the sensor network comprises the capability to gather data to inform these computations. Signal processing also plays an important role in both the reception of data (e.g. electro-optical signal processing for detection) and the transmission of information (e.g. alerts or sensor tasking requests). The personnel available to perform the functions at the generalized function level also add a constraint on SSA operations. Finally, the space object catalog resources and individual space object ephemerides provide a means of disseminating information about the space object population. This is primarily accomplished through public-facing catalogs that do not contain any classified information, such as the US Air Force's Space-Track website.

Physical Form: The final level of decomposition lists work domain constraints and capabilities imposed by physical characteristics of the domain. As mentioned previously,

the disparity between the number of trackable space objects and the number of sensors places a significant constraint on SSA operations. Additionally, observational constraints are imposed by the relative geometries between sensors and space objects, subject to orbital dynamics. The identification of certain space objects as higher-priority targets (to resolve hypotheses or improve catalog estimates) helps constrain this problem. Finally, the condition of the measurement signals imposes added constraints, as uncertainty and bias in state estimation and measurement processing leads to ambiguity in evidence.

2.2.4 SSA Environment Decomposition

The second system examined is the SSA environment, modeled similar to the corresponding environment hierarchies in the naval [119] and EVA [104] studies. Unlike operators in the work domain, the environment does not have any functional purpose or goals, so the top level of the traditional decomposition is omitted. However, there are still many dynamic elements to the SSA environment that impact decision support design considerations. Due to similarities in the relevant phenomenologies and constraints, the results of this decomposition, shown in Fig. 2.3, closely resemble the EVA decomposition [104].

Abstract functions: The laws of conservation of mass, momentum, and energy govern the physical domain of the orbit environment. These predictable laws impose constraints on everything from spacecraft orbits to sensor capabilities.

Generalized functions: Several physical processes, governed by the laws of conservation in the abstract functions level, affect SSA operations. Orbital dynamics prescribe the motion of objects in orbit, including spacecraft, debris, and celestial bodies such as planets and stars. The dynamics of these objects primarily drive data collection by limiting observability to specified time-spans. Atmospheric processes also affect observability through the quality of the data gathered. Similarly, space weather (such as solar storms or high-energy particles) can impact both spacecraft and sensor operation. Finally, the capability to transmit and receive electromagnetic signals provides a basis for observation and

Functional Purpose	Space Asset Safety	National Security						
Abstract Function	Orbital Dynamics	Sensor Phenomenology	Space Object Asset Priority	Hypotheses and Priorities	Workflow Efficiency			
Generalized Function	Event Detection	Conjunction Risk Assessment	Information Fusion	Sensor Allocation	Catalog Maintenance	Accuracy Degradation	Uncorrelated Track Processing	Information and Alert Dissemination
Physical Function	Computational Resources	Sensor Network Resources	Signal Transmission and Processing Capability	Personnel Capability	Space Object Catalog	Individual Ephemerides	Public Catalog	
Physical Form	Number, Type, and Location of Tracked Objects	Number, Type, and Location of High-Priority Assets	Number, Type, and Location of Sensors	Signal Characteristics (uncertainty, ambiguity)				

Figure 2.2: Abstraction hierarchy of the SSA work domain

communication.

Physical functions: In the environment decomposition, elements of the physical function level pertain to operational environments of the operators' systems [104]. Both engineered and natural objects constrain the transmission of signals by either producing a signal for observation, providing a path for data transmission, or obscuring that path to prevent transmission. The atmosphere provides similarly important constraints on observations through "atmospheric seeing" as electromagnetic plane waves emitted or transmitted by objects in space are distorted when passing through the atmosphere. Conversely, the vacuum of space provides a more pristine environment for the transmission of electromagnetic signals, so sensors or assets beyond the atmosphere can avoid some of these constraints.

Physical Form: The environment contains several physical elements that significantly affect SSA operations. For instance, Earth's gravitational and electromagnetic fields drive propagation and safe operation of orbiting assets, providing predictable but still stochastic dynamics. The solar radiation environment not only affects objects through high-energy particle radiation, but also through the luminosity of the Sun that illuminates targets and drives electro-optical observation opportunities. Naturally the locations and types of man-made objects in the environment can drive observation opportunities for those or other objects and transmission of data between different parts of the SSA network. Similarly, the location and types of celestial bodies may transmit or constrain data collection, but also provide added data points in some measurements: for instance, an electro-optical system typically leverages the (nearly) inertially fixed locations of stars to further constrain angle-measurements. Finally, turbulence and adverse weather play a large role in both ground-based sensor operation, including signal degradation that affects both scientific and communication data.

Abstract Function	Conservation of Mass	Conservation of Energy	Conservation of Momentum			
Generalized Function	Orbital Dynamics	Atmospheric Processes (air movement, temperature, weather)	Space Weather (radiation, solar pressure, electromagnetism)	Electromagnetic Signal Generation and Transmission		
Physical Function	Engineered Objects (signal transmission capabilities/limitations)	Natural Objects (signal transmission limitations)	Atmosphere (signal transmission capabilities/limitations)	Vacuum (signal transmission capabilities)		
Physical Form	Gravitational Field	Electromagnetic Field	Solar Radiation (luminosity, radiation)	Location and Type of Celestial Bodies (stars, planets, natural debris)	Location and Type of Man-Made Objects (spacecraft, man-made debris)	Atmospheric Conditions (turbulence, pressure, temperature, clouds)

Figure 2.3: Abstraction hierarchy of the SSA environment

2.2.5 Insights for SSA DSS Development

The work domain analysis conducted through the abstraction hierarchies above identifies capabilities and constraints present within both the SSA operator's work domain and the SSA environment as a whole. These results can be used to develop relationships and perspectives related to important aspects of the combined SSA work domain and environment.

One specific take-away that can be derived from these decompositions is the importance of data fusion from disparate sensor resources and various signal characteristics, including considerations for uncertainty, ambiguity, and unobservability. State estimates are only as good as the quality of the physical data products, so the signal processing and information fusion must be able to account for these difficult aspects of the SSA observation problem. More importantly, however, this data must be fused to address actionable decision-maker needs derived from the SSA domain purposes of asset safety and national security.

A useful insight from this analysis comes from comparing existing sensor allocation methodologies to the goals of the work domain. Many existing or proposed approaches focus on minimizing catalog-wide covariance estimates. This involves fusing the data (signals), generated using the available sensors, to maintain a prescribed covariance accuracy dependent upon the priority of the asset. This sensor allocation addresses the function of catalog maintenance, as well as the goal of ensuring space asset safety by improving the accuracy of conjunction assessments; however, is not necessarily well-suited to resolving other hypotheses that may pertain to national security or space object safety. In particular, unless the hypothesis can be directly restated in terms of covariance estimates, an algorithm predicated on state uncertainty minimization will not be able to prioritize that hypothesis. Therefore, there is a need for sensor allocation approaches that not only address the catalog maintenance function, but also directly address other decision-maker hypotheses related to safety and national security.

Additionally, the proliferation of different sensor technologies means that the physical form of sensor data is widely varied. In some cases, evidence toward certain hypotheses

may even be extracted from non-traditional sources such as news articles. An effective DSS for SSA must be able to ingest data from disparate sources and fuse all that data to update hypothesis knowledge states that address the decision-maker concerns.

The design insights gleaned from this analysis are summarized as follows:

- Existing sensor allocation methods (e.g. covariance-based tasking) lack robust translation to domain goals.
- Information fusion must incorporate data from disparate sensor phenomenologies.

For further insight into the development of design requirements for an SSA DSS, this work continues with the second stage of the cognitive work analysis.

2.3 Control Task Analysis Applied to SSA

Using the results of the work domain analysis, the second phase of the cognitive work analysis, the control task analysis (ConTA), is conducted to further develop insights for decision-support requirements. Previous work by the NASA directorate of Human Effectiveness has investigated how new fusion technologies could be incorporated into the SSA workflow [120, 7], including a ConTA study used to inform designs for several prototype screens for evaluation by Joint Space Operations Center (JSpOC) operators [121]. The full results of this analysis are under distribution restriction [121]. In contrast, this work seeks to derive generalized requirements for SSA DSS development.

2.3.1 Decision Ladders

ConTA emphasizes the actions that the worker should undertake to accomplish a particular task, while also encouraging flexibility and expertise by not adhering to a strict linear, procedure-like approach. The decision ladder, a popular choice of control task modeling, maps information processing actions and states of knowledge throughout a control task to

model the cognitive processes required to complete the task. Figure 2.4 shows a sample decision ladder in the style of Rasmussen and Vicente [115, 59].

Beginning in the bottom-left of the ladder, the *analysis* phase involves ingestion of alerts and observations to identify the system state. The *judgment* phase, at the top of the ladder, models the selection of a particular target goal through the consideration of options and their consequences. Then, descending toward the bottom-right of the ladder, the *planning* phase of the task selects actions to execute based on the stated goals, and the task terminates in execution that plan. It is important to note that not all control tasks need to utilize the entirety of the ladder; for instance, some tasks may not involve any planning and execution so the task can terminate after analysis and judgment. Similarly, not all tasks begin with an activation or alert and may be best modeled starting from a particular system state, for instance. In this way, the decision ladder approach avoids enforcing a linear approach to modeling control tasks.

To further encourage flexibility in task modeling, the decision ladder may include shortcuts in the form of associative leaps (from one state of knowledge state directly to another) and shunts (from a state of knowledge directly to a non-adjacent processing activity). In practice, these shortcuts arise from experience as an expert begins to recognize familiar situations: the expert is able to employ skill- and rule-based behavior to develop shortcuts and avoid cognitively expensive tasks in upper portions of the decision ladder, which are associated with knowledge-based behavior [122, 59]. In the template in Fig. 2.4, an associative leap is shown for a case where, once an alert is received, the operator immediately knows, through experience, the system state and can bypass observation and identification. Similarly, a shunt is shown for a case where, given a particular system state, the operator immediately knows, through experience, the appropriate choice of task and can bypass the judgment phase of control task. These considerations for flexibility are what allow designers to use decision ladders to support expert work.

The control task decision ladders can be leveraged to generate two different types of

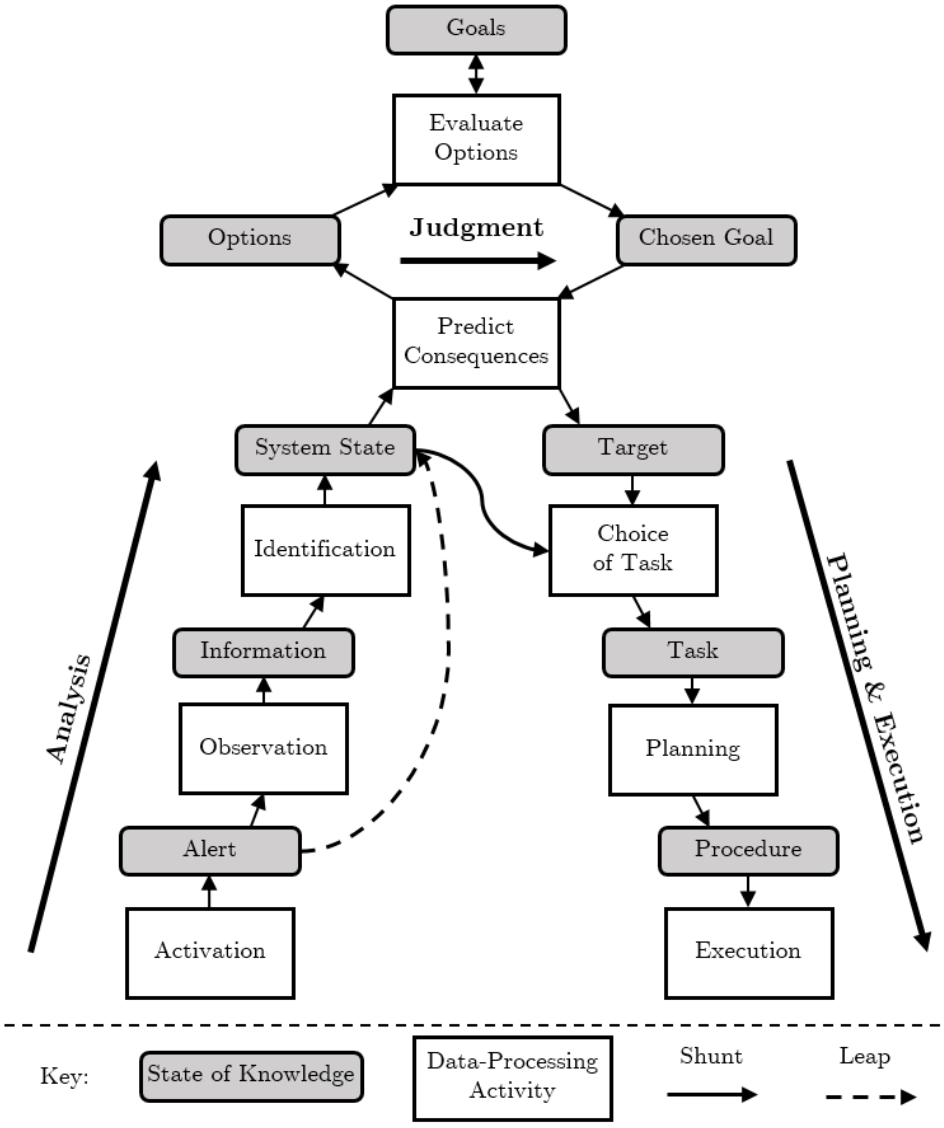


Figure 2.4: Control task analysis decision ladder template, adapted from Vicente [59]

design requirements [104, 123, 124]. A cognitive work requirement (CWR) specifies cognitive demands, tasks, and decisions that must be supported by the DSS. An information relationship requirement (IRR) specifies the context for required data, which translates that data into the actionable information that the decision-maker requires.

Miller et al. [104] demonstrate how to translate from the states of knowledge in the decision ladder to CWRs and IRRs. In their study, each state of knowledge generates at least one CWR, and each CWR has a corresponding IRR. A similar approach is followed in this work. For each applicable decision ladder state, states of knowledge articulate ques-

tions relevant to the SSA operator. Each state of knowledge generates a CWR outlining some functionality the DSS must provide to address this state of knowledge. Similarly, the CWR generates an IRR more explicitly stating the data products required to generate the necessary information.

2.3.2 ConTA Application: Information Fusion and Sensor Allocation

Recalling the SSA work domain abstraction hierarchy in Fig. 2.2, any number of these elements may be identified for further inspection through ConTA. This thesis focuses on information fusion and sensor allocation to support SSA decision making, so the ConTA will similarly focus on these functions. Information fusion and sensor allocation may be related in one decision ladder, as the information fusion function generally relates to the analysis phase while sensor allocation relates to judgment, planning, and execution.

Additionally, the WDA analysis above identified that the formulation and resolution of hypotheses can form a means to the ends of maintaining space asset safety and addressing national security concerns. Therefore, this application of ConTA will focus on the hypothesis resolution abstract function as the means to address these SSA purposes.

The CWRs and IRRs for the information fusion portion of this task (analysis) are shown in Fig. 2.5. The following sections describe each decision ladder state in more detail to provide context for the CWRs and IRRs.

Alert: The SSA information fusion process begins by receiving alerts to any new data. This may come in the form of direct sensor data from previous task execution, resulting in successful detections and correlations, uncorrelated tracks, or missed detections. Additionally, data may be derived from alternate sources such as a satellite operator reporting anomalous behavior or a program that parses articles or databases for keywords. The availability of any new information that may aid in resolving hypotheses should raise alerts to operators.

Information: Next, the acquired data must be processed into information that can be

Decision Ladder States	States of Knowledge	Cognitive Work Requirements	Information Relationship Requirements
Alert	• What new data has been acquired?	• **CWR-1:** The DSS shall track incoming data over time from diverse data sources.	• **IRR-1:** Track and store observational and auxiliary data as well as operator reports.
Information	• How do acquired observational data compare with expected data?	• **CWR-2:** The DSS shall assess acquired observational data in comparison to expected values.	• **IRR-2:** Calculate deviation of observation data (e.g. angles, range, light-curve) from expected values.
	• What is the evidence corresponding to the acquired data?	• **CWR-3:** The DSS shall translate observational data into evidence.	• **IRR-3:** Map data sources to hypothesis evidence to compute acquired evidence.
	• What is the strength of the acquired evidence?	• **CWR-4:** The DSS shall estimate the strength of the acquired evidence.	• **IRR-4:** Associate actual observation conditions with acquired data to determine evidence ambiguity.
System State	• What are the updated system states?	• **CWR-5:** The DSS shall assess states for objects, sensors, and hypotheses.	• **IRR-5:** Update estimates for space object states, sensor performance, observation conditions, and hypothesis resolution.
	• What hypotheses still need to be resolved?	• **CWR-6:** The DSS shall track hypothesis resolution in comparison to prescribed resolution thresholds.	• **IRR-6:** Update prior hypothesis resolutions using acquired evidence.
	• What are the constraints on the remaining hypotheses?	• **CWR-7:** The DSS shall associate hypothesis resolution and observational constraints with remaining resolution time horizon.	• **IRR-7:** Estimate evidence-gathering actions available with good conditions within hypothesis time horizons.

Figure 2.5: Requirements derivation for SSA information fusion for hypothesis resolution

used to update the system state. This involves careful consideration of the incoming data, as well as the expected data, to develop mappings to actionable evidence. Available data may include individual detections from radar or electro-optical sensors, detection and correlation probabilities that can be used to update state estimates, light-curves that can be used to estimate attitude or control modes, and auxiliary data such as background sky brightness or atmospheric transmittance. In processing the acquired data, the strength of the evidence should also be taken into consideration. For instance, a "missed detection" alert may indicate a loss of custody; however, it may also simply indicate poor observation conditions, in which case the evidence toward any custody hypotheses should be considered vacuous. The data-to-evidence mappings must be complete enough to account for complex conditions such as the one described above, but also flexible enough to allow for the inclusion of data and evidence that was not anticipated in the design phase.

System State: Once evidence has been extracted from the acquired data, this evidence must be fused to update the system states. In a hypothesis-based scheme, this includes updating prior hypotheses using the new evidence to arrive at posterior hypothesis knowledge states. The DSS should clearly identify which hypotheses are still unresolved, and as well as any constraints on these remaining hypotheses (e.g. resolution time horizons, line-of-sight opportunities for sensors).

After the system state has been updated, the information fusion task (the analysis phase of the decision ladder) is complete. Sensor allocation picks up from this stage to follow through with judgment, planning, and execution. The CWRs and IRRs for this task are shown in Fig. 2.6, and details on the decision ladder states are shown below.

Options, Goals, and *Chosen Goal:* The judgment phase consists of an iterative process wherein the overall SSA goals and purposes are considered along with options related to addressing these goals. The primary goals for SSA operators are to ensure space asset safety and address national security concerns, and these goals may be modeled as a constantly evolving set of hypotheses and associated priorities as determined by SSA operators

Decision Ladder States	States of Knowledge	Cognitive Work Requirements	Information Relationship Requirements
Options	• What updates are required to the considered set of hypotheses?	• **CWR-8:** The DSS shall enable operators to modify list of hypotheses as required.	• **IRR-8:** Support creation of new hypotheses, removal of resolved hypotheses.
Goals	• What are the relative priorities between the hypotheses?	• **CWR-9:** The DSS shall provide capability for operators to adjust hypothesis priorities.	• **IRR-9:** Adjust hypothesis priorities based on user input, update estimated resolutions.
Chosen Goal	• What are the estimated hypothesis resolutions?	• **CWR-10:** The DSS shall assess expected hypothesis resolution based on current prioritization.	• **IRR-10:** Estimate expected observation data, map to evidence, update expected hypothesis resolutions.
Target	• What resources are available to address the prioritized hypotheses?	• **CWR-11:** The DSS shall assess the evidence, and related resources, available to resolve hypotheses.	• **IRR-11:** Generate lists of possible allocations to address chosen goals, consider operational and organizational constraints.
Task	• What evidence should be gathered to address prioritized hypotheses?	• **CWR-12:** The DSS shall propose the optimal set of evidence to resolve the hypotheses.	• **IRR-12:** Select optimal set of evidence to gather from available evidence.
Procedure	• What specific sensor actions or data requests are required to gather the necessary evidence?	• **CWR-13:** The DSS shall generate specific actions and requests required to reach target hypothesis resolution.	• **IRR-13:** Generate list of actions and requests required to gather evidence.
	• What deviations from the current allocation are required to execute these actions?	• **CWR-14:** The DSS shall assess required changes in sensor allocation and alert operators of expected impacts of these changes.	• **IRR-14:** Compare proposed sensor allocation to existing allocation, evaluate resultant changes in expected data/evidence.

Figure 2.6: Requirements derivation for SSA sensor allocation for hypothesis resolution

and decision-makers. Therefore, the DSS must allow operators to use information from the system state to spawn new hypotheses as necessary. Similarly, operators must be able to identify and confirm resolved hypotheses to remove them from future tasking consideration, if appropriate. Operators must also consider which hypothesis priorities should be modified in order to address the overall goal of resolving all the hypotheses within prescribed tolerances. When adjusting the priorities, the consequences of any changes should be clearly articulated to the decision-maker. This involves forecasting evidence returns from candidate action sequences and estimating the updated hypothesis state. This iterative process terminates when decision-makers are satisfied with the estimated resolution.

Target: Once the set of goals and priorities have been adjusted satisfactorily, the sensor allocation continues by determining the evidence required to address the selected goals. The DSS must consider operational constraints such as sensor availability, line-of-sight, and observation conditions, as well as organizational constraints (e.g. requesting data from sensors controlled by other organizations). Therefore, the DSS should identify the evidence available to address the chosen goals and priorities, as well as the resources required to gather that evidence. Recall the following design insight from the WDA: the DSS must be able to ingest and incorporate data from disparate sources. As such, the DSS should be able to leverage these disparate data sources in task planning.

Task and *Procedure:* From the available evidence identified in the previous stages, an optimal set of evidence must be selected that best addresses the prioritized hypotheses. This generates an action procedure to gather or request the necessary evidence. In addition to gathering primary data products required for hypothesis resolution, any auxiliary data that may be used to build context for the data should be considered. Recall that auxiliary data can play a significant role in the data-to-evidence mapping of the *Information* phase. Any required deviations from nominal or previously-issued tasking assignments must also be clearly articulated to inform operators of required modifications and any expected impacts on previously planned activities.

In total, this ConTA of the information fusion and sensor allocation functions generated 14 cognitive work requirements and 14 corresponding information relationship requirements.

2.4 Conclusions

By carefully considering the goals of operators in SSA, as well as constraints imposed by the domain, this application of CSE methods develops several insights and design requirements for decision-support in SSA. The WDA provides insights that existing sensor allocation and information fusion approaches, primarily predicated on state uncertainty minimization, do not provide a robust or clear mapping to decision-maker goals of maintaining space asset safety and addressing national security concerns. This indicates an opportunity for improvement by allocating sensors to address specific hypotheses related to these SSA domain goals.

The following chapters will investigate several applications of this hypothesis abstraction to support SSA goals, developing information fusion and sensor allocation methods using hypotheses. Chapter 6 will revisit the design requirements generated through the ConTA to design a prototype DSS for further investigation of the hypothesis-based tasking approach.

CHAPTER 3
BINARY HYPOTHESIS TESTING FOR ANOMALY DETECTION

Correlating on-orbit observations and detecting space object maneuvers is a challenging endeavor in Space Situational Awareness (SSA). Predicting conjunction events is a difficult task [60], but recent events, such as the Iridium-Cosmos on-orbit collision, highlight the mutual interest that national and private operators share for accurate object correlation and maneuver detection capability [9]. Detecting maneuvers is particularly important when tracking active objects for which no operational information is available, as detecting maneuvers in real-time is required to adequately react to anomalies or possible conjunctions [22, 61]. Timely detection of maneuvers allows for responsiveness in follow-up tracking, which is crucial for post-maneuver orbit characterization [22]. This task is especially difficult in low-thrust maneuvers, where state change is more subtle [23].

This chapter develops an algorithm for spacecraft dynamic anomaly detection using distance metrics. The control distance metric provides a method for computing the distance between two state distributions resulting from different uncorrelated tracks (UCTs) [24]. This is accomplished by linearizing about a nominal optimal trajectory connecting the mean states of the state probability density function (PDFs). Due to the non-linear nature of the general orbit determination problem, these PDFs can be non-Gaussian, particularly after propagation for some time [62]. Therefore, this work extends the previous control distance metric using a Gaussian mixture model (GMM) approximation for application to non-Gaussian boundary conditions.

This work further modifies the approach to control distance anomaly detection by introducing a binary hypothesis structure, using control cost distributions from both the maneuvered and quiescent (non-maneuvered) trajectories. Using the extended control distance metric approach, control cost distributions are computed for the binary hypothesis

pair of a non-maneuvered and a maneuvered trajectory. The control cost PDFs are used in a binary hypothesis test, subject to a desired false alarm rate, to determine the probability that an anomaly has occurred. The primary contributions in this chapter are: 1) a computationally tractable GMM approximation to computing control cost distributions for non-Gaussian boundary condition probability densities, 2) a binary hypothesis testing framework for anomaly detection using the control distance metric that permits specification of false alarm rates, 3) an analogous binary hypothesis testing framework for anomaly detection using Mahalanobis distance, and 4) the implementation of these two methods with quantitative performance comparisons drawn between the control distance metric and Mahalanobis distance using both synthetic and empirical data. These metrics are compared using a GEO spacecraft in both North-South and East-West station-keeping test cases.

After development of the novel theoretical contributions, implementation details for both control distance and Mahalanobis distance detection methods are discussed, followed by simulation results using synthetic data for a GEO spacecraft performing both North-South and East-West station-keeping maneuvers. Corresponding results using empirical data are also presented. Finally, the results from the synthetic and real data simulations are distilled into potential operational applications.

3.1 Data Association in SSA

Data association algorithms for object correlation and maneuver detection have been well explored in literature. The data association task hypothesizes an association and attempts to compute a measure of the probability that the hypothesis is true. Methods often focus on admissible regions or probabilistic approaches to compare uncorrelated tracks and detect maneuvers. For instance, Tommei et al. address object correlation and orbit determination with admissible region-based methods and a virtual debris algorithm that were applied to optical observations and radar observations [27]. Maruskin et al. also use admissible regions for object correlation by mapping admissible regions to Delaunay or Hamiltonian

orbit elements and by comparing an observation with an earlier estimate propagated forward in time [28]. Fujimoto et al. employ highly constrained probability distributions in Poincare orbit element space, where distributions are defined by admissible-region maps such that the intersection between admissible regions, or lack thereof, reflects the correlation between respective observations [29].

DeMars et al. approach the SSA data association problem using finite-set statistics for multi-object estimation. The adaptive entropy-based Gaussian-mixture information synthesis (AEGIS) approach is used to track objects while utilizing finite-set statistics (FISST) to account for uncertainty, false alarms and missed detections = [26]. Kelecy and Jah apply batch least-squares and extended Kalman filter based strategies to detect and reconstruct low thrust finite maneuvers [23]. Huang et al. relate UCTs using a nonlinear least squares iterative process to optimally estimate maneuvers and correlate objects following a maximum a posteriori criterion [25].

Likewise, there are many distance or pseudo-distance metrics that may be used to measure the discrepancy between two state distributions (e.g. Mahalanobis distance) [63]. Problematically, existing metrics do not directly quantify the level of propulsive effort required to cause the observed state change. The problem of associating UCTs over large time periods is particularly difficult when resident space objects (RSOs) maneuver during observation gaps. Even relatively small station-keeping maneuvers at geostationary Earth orbit (GEO) can result in position discrepancies of many kilometers after an observation gap. UCT correlation is further confounded by state estimate uncertainties [24]. Since both the initial and final UCTs are best estimates, with associated PDFs, correlation is difficult particularly in densely-populated regions of the space environment.

Holzinger et al. propose a minimum-fuel control distance metric to approach data association and maneuver detection while considering propulsive effort and reconstructing maneuvers [64]. Since on-board fuel remains a scarce commodity for operational spacecraft, operators are likely to execute fuel-optimal, or near-fuel-optimal, maneuvers [24].

Under the assumption of optimal control, multiple deterministic UCTs can be related by computing the control effort required for a trajectory to meet those boundary conditions. This approach necessitates the reconstruction of a minimum-fuel trajectory consistent with the a priori information and new observations. Holzinger et al. have shown, through the properties of strict positivity, symmetry, and triangle inequality, that control performance is a metric [24], allowing objective comparisons to other commonly used distance metrics. The control cost distributions required can be computed along each relevant trajectory by considering boundary condition uncertainty [64]. In previous work, a single cost distribution was developed using the maneuvered trajectory, testing the anomaly hypothesis by comparing the size of the observed maneuver to the amount of uncertainty in control cost due to uncertainty in the boundary conditions [64]. This straight-forward approach encounters problems when attempting to draw conclusions regarding error rates, making comparisons to other metrics incomplete.

A commonly-used statistical approach in anomaly detection that incorporates error rates is binary hypothesis testing, wherein integration over a pair of PDFs allows determination of false alarm and missed detection probabilities. Binary hypothesis testing has been implemented for anomaly detection in various fields, such as signal processing [65]. One variant of binary hypothesis testing, the Neyman-Pearson approach, devises the most powerful likelihood-ratio test for a given significance level and threshold [66]. In Neyman-Pearson detector implementation, these thresholds are selected through analysis of a number of observations with associated PDFs; however, in applications involving only one observation, this method reduces to a more basic form of binary hypothesis testing, which is the case for UCT association where only one PDF is available at each time epoch.

Evaluating metrics for anomaly detection requires selection of a representative subsection of the infinite continuum of possible maneuvers. Since optical observations are primarily useful for space objects at high altitudes, such as GEO, maneuvers relevant to GEO spacecraft are particularly interesting. Spacecraft in GEO are assigned to specific

longitude slots for their operational lifetime, but are also subject to a number of perturbations that must be rejected using station-keeping maneuvers. For instance, North-South station-keeping maneuvers adjust inclination, which is primarily perturbed by third-body gravitational effects, to maintain an equatorial orbit. East-West station-keeping maneuvers adjust true longitude (or phase), which is primarily perturbed by Earth oblateness sectoral harmonics (i.e. J_{22}), to maintain the spacecraft's GEO slot. These two primary station-keeping maneuvers form a representative subset of maneuvers that could potentially be encountered in operation.

3.2 UCT Association Scenario

The following notional scenario is relevant to the task of associating a pair of UCTs to detect maneuvers. As pictured in Fig. 3.1, an uncorrelated track $\mathrm{UCT_A}$ at time t_0 is

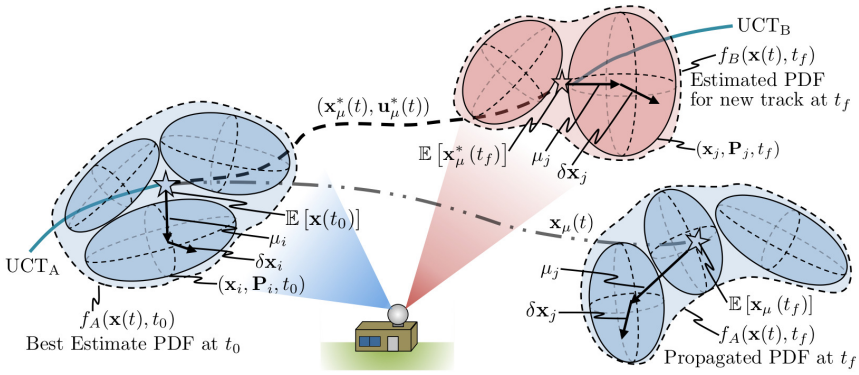


Figure 3.1: Maneuver detection scenario

used to generate a PDF, $f_A(\mathbf{x}(t_0))$. Using the mean state, $\mathbf{x}_\mu(t_0) = \mathbb{E}[\mathbf{x}(t_0)]$, the track is propagated to time t_f under assumed quiescent dynamics, $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t) = \mathbf{0}, t)$, yielding the propagated mean state $\mathbf{x}_\mu(t_f)$ and its associated PDF $f_A(\mathbf{x}(t_f))$. Another uncorrelated track at time t_f, $\mathrm{UCT_B}$, yields its own PDF $f_B(\mathbf{x}(t_f))$ and mean state $\mathbf{x}^*_\mu(t_f) = \mathbb{E}[\mathbf{x}^*(t_f)]$. An optimal maneuvered trajectory is generated connecting the mean states of the UCTs, $\mathbf{x}_\mu(t_0)$ and $\mathbf{x}^*_\mu(t_f)$, under the same dynamics with non-zero control, $\dot{\mathbf{x}}^*(t) =$

$\mathbf{f}(\mathbf{x}^*(t), \mathbf{u}^*(t), t)$. This yields the maneuvered trajectory $\mathbf{x}^*_\mu(t)$ and associated control $\mathbf{u}^*_\mu(t)$. Additional notation is included in Fig. 3.1 to describe the Gaussian components of the state PDFs.

The anomaly detection algorithm is assembled as shown in Fig. 3.2. The inputs to the

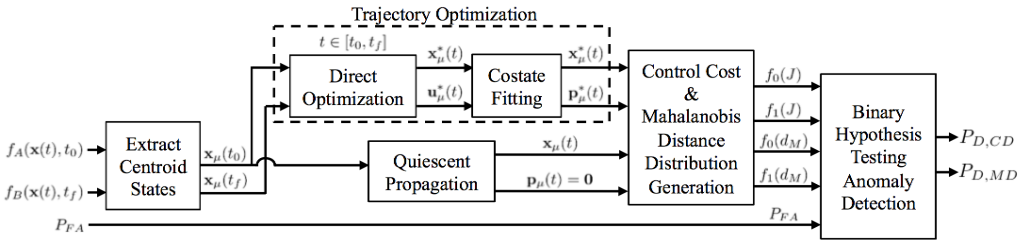


Figure 3.2: Minimum-fuel control distance anomaly detection and characterization framework

algorithm are a pair of UCTs, represented by PDFs at times t_0 and t_f. The output of the algorithm is the probability that the anomaly hypothesis is true, P_D, indicating whether an anomaly has been detected. Stated differently, this is the probability that something outside the modeled quiescent dynamics (e.g. a maneuver) has occurred, under the preliminary assumption that the two UCTs are associated with the same object. Anomaly probability is calculated using two different methods, the control distance metric and Mahalanobis distance, yielding two anomaly probabilities for this study. This allows the analogous methods to be compared in both anomaly detection sensitivity and error rates.

3.3 Gaussian Mixture Model Approximation for Control Distance Metric

The anomaly detection algorithm begins with a trajectory optimization routine, which uses mean states of the boundary conditions to generate boundary time-fixed optimal connecting trajectories. The UCT pair is considered a two-point boundary value problem (TPBVP), and a trajectory is computed to optimally connect the two UCTs, minimizing a chosen cost function J. In this study, the trajectory is optimized with respect to the quadratic control cost, shown in Eq. (3.1) [67]. This particular cost function satisfies the metric properties of

non-negativity, coincidence, symmetry, and triangle inequality.

$$J(\mathbf{u(t)}; t_0, t_f) = \frac{1}{2} \int_{t_0}^{t_f} \mathbf{u}(\tau)^{\mathrm{T}} \mathbf{u}(\tau) \, \mathrm{d}\tau. \tag{3.1}$$

The quadratic control cost function is ideal for variable specific impulse (VSI) engines, often used in low-thrust applications. Alternate cost functions could also be implemented, such as an impulsive cost function, as long as they also satisfy metric properties. In operation, there is an infinite continuum of possible (optimal or sub-optimal) control trajectories that operators might use. This trajectory optimization strategy provides a lower-bound on the required maneuver to transition between the given states, subject to the chosen cost function. This is similar to a reachability analysis, which imposes an upper-bound on the reachable state-space given an assumed propulsive capability.

Derivation of the GMM approximation for the control distance metric follows an approach similar to the Gaussian-restricted derivation [24]. Using the optimal state and costate trajectory ($\mathbf{x}^*_\mu(t)$, $\mathbf{p}^*_\mu(t)$), state uncertainties are incorporated to generate an approximate probability distribution of the control distances associated with propagation between the boundary conditions. This is accomplished by linearizing about the nominal optimal trajectory and applying perturbations within the boundary condition distributions to develop control cost distributions.

The probability density of a state $\mathbf{x}$ in a GMM is defined using the weighted sum of a set of k multivariate Gaussian, or normal, density functions, as in Eq. (3.2)

$$f(\mathbf{x}; \mathbf{x}_{\mu,1}, \mathbf{P}_1, \ldots, \mathbf{x}_{\mu,k}, \mathbf{P}_k) = \sum_{i=1}^{k} w_i f(\mathbf{x}; \mathbf{x}_{\mu,i}, \mathbf{P}_i) \tag{3.2}$$

where $f(\mathbf{x}; \mathbf{x}_{\mu,i}, \mathbf{P}_i)$ is the density of state $\mathbf{x}$ in the i^{th} Gaussian component of the GMM and $w_1, \ldots, w_k$ are the GMM weightings such that $w_i \geq 0$ and $\sum_{i=1}^{k} w_i = 1$. A multivariate random variable, $\boldsymbol{\zeta}$, sampled from this mixture is chosen by first sampling a random variable $\mathcal{U} \sim \mathrm{Uniform}(0, 1)$. Given $\mathcal{U}$, the j^{th} Gaussian component is selected such

that $\mathcal{U} \in \left[\sum_{i=1}^{j-1} w_i, \sum_{i=1}^{j} w_i\right]$. Then $\boldsymbol{\zeta}$ is sampled from that j^{th} Gaussian component: $\boldsymbol{\zeta} \sim \mathcal{N}(\mathbf{x}_{\mu,j}, \mathbf{P}_j)$. The centroid state, or expectation, of the initial boundary condition is computed as the weighted average of the mean states of each Gaussian component, as shown in Eq. (3.3):

$$\mathbf{x}_\mu(t_0) = \mathbb{E}\left[\mathbf{x}(t_0)\right] = \mathbb{E}\left[\sum_{i=1}^{n} w_i \mathbf{x}_i(t_0)\right] = \sum_{i=1}^{n} w_i \mathbf{x}_{\mu,i}(t_0) \tag{3.3}$$

The centroid state of the final boundary condition can be computed similarly. The centroid states enable the trajectory to be reduced to a two-point boundary value problem, similar to the previous technique [64].

The control cost distribution is constructed by linearizing about the optimal trajectory, $(\mathbf{x}^*_\mu(t), \mathbf{u}^*_\mu(t))$, connecting the centroid states and sampling initial and final states from the non-Gaussian boundary conditions, which introduces new perturbing terms. Included in Fig. 3.1 are notional depictions of the key variables required for the GMM approximation. The deviation of the mean state of an individual Gaussian component from the centroid state of its mixture is defined as $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ for the initial and final boundary conditions, respectively. This deviation term adds a new perturbation to the control cost distribution.

The quadratic control cost function, in Eq. (3.4), can be expanded by decomposing the control effort $\mathbf{u}(t)$ into three components as shown in Eq. (3.5):

$$J = \frac{1}{2}\int \mathbf{u}(\tau)^{\mathrm{T}}\mathbf{u}(\tau)\mathrm{d}\tau \tag{3.4}$$

$$\mathbf{u}(t) = \mathbf{u}^*_\mu(t) + \delta\mathbf{u}_{ij}(t) + \delta\mathbf{u}(t) \tag{3.5}$$

where $\mathbf{u}^*_\mu(t)$ represents the optimal control associated with the mean trajectory $\mathbf{x}^*_\mu(t)$, $\delta\mathbf{u}_{ij}(t)$ represents the control perturbation due to the $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ variations in mean state of Gaussian initial component i and Gaussian final component j from the centroid states, and $\delta\mathbf{u}(t)$ represents the control perturbation due to $\delta\mathbf{x}_0$ and $\delta\mathbf{x}_f$, variations in the state

sampled from boundary condition uncertainty. From Holzinger et al. [24], the optimal control effort can be written as a function of the costates using Eq. (3.6):

$$\mathbf{u}(t) = -\frac{\partial \mathbf{f}}{\partial \mathbf{u}}^{\mathrm{T}} \left(\mathbf{p}^*_\mu(t) + \delta\mathbf{p}_{ij}(t) + \delta\mathbf{p}(t)\right) \tag{3.6}$$

where $\mathbf{p}^*_\mu(t)$ represents the optimal costate associated with the mean trajectory $\mathbf{x}^*_\mu(t)$, $\delta\mathbf{p}_{ij}(t)$ represents the costate perturbation due to $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$, and $\delta\mathbf{p}(t)$ represents the costate perturbation due to $\delta\mathbf{x}_0$ and $\delta\mathbf{x}_f$. There exists a function $\boldsymbol{\Lambda}(t, t_0) \in \mathbb{R}^{n\times 2n}$, shown in Eq. (3.7), that maps variations in initial and final states to variations in the costate at time t, $\delta\mathbf{p}(t)$.

$$\boldsymbol{\Lambda}(t, t_0) = \begin{bmatrix} \boldsymbol{\Phi}_{px}(t, t_0) & -\boldsymbol{\Phi}_{pp}(t, t_0)\boldsymbol{\Phi}_{xp}(t_f, t_0)^\dagger \boldsymbol{\Phi}_{xx}(t_f, t_0)\boldsymbol{\Phi}_{pp}(t, t_0)\boldsymbol{\Phi}_{xp}(t_f, t_0)^\dagger \end{bmatrix} \tag{3.7}$$

Note that this function is composed of portions of the state transition matrix partitioned as

$$\begin{bmatrix} \delta\mathbf{x}(t) \\ \delta\mathbf{p}(t) \end{bmatrix} = \boldsymbol{\Phi}(t, t_0) \begin{bmatrix} \delta\mathbf{x}(t_0) \\ \delta\mathbf{p}(t_0) \end{bmatrix} \tag{3.8}$$

$$= \begin{bmatrix} \boldsymbol{\Phi}_{xx}(t, t_0) & \boldsymbol{\Phi}_{xp}(t, t_0) \\ \boldsymbol{\Phi}_{px}(t, t_0) & \boldsymbol{\Phi}_{pp}(t, t_0) \end{bmatrix} \begin{bmatrix} \delta\mathbf{x}(t_0) \\ \delta\mathbf{p}(t_0) \end{bmatrix} \tag{3.9}$$

where $\boldsymbol{\Phi}(t, t_0)$ is the state transition matrix mapping variations $\delta\mathbf{x}$ and $\delta\mathbf{p}$ to time t about the optimal trajectory. Also note that, while the pseudoinverse term $\boldsymbol{\Phi}_{xp}(t_f, t_0)^\dagger$ is not guaranteed to exist for arbitrary systems, its existence implies controllability through the optimal control problem. This portion of the state transition matrix determines how variations in the costates affect the state, or in other words whether the state is controllable. In this paper, controllability is assumed, so for present purposes the pseudoinverse is also assumed to exist [24].

Using $\boldsymbol{\Lambda}(t, t_0)$, the components $\delta\mathbf{u}_{ij}(t)$ and $\delta\mathbf{u}(t)$ can be computed as shown in Eqns.

(3.10) and (3.11).

$$\delta \mathbf{u}_{ij}(t) = -\frac{\partial \mathbf{f}}{\partial \mathbf{u}}^{\mathrm{T}} \mathbf{\Lambda}(t, t_0) \begin{bmatrix} \boldsymbol{\mu}_i \\ \boldsymbol{\mu}_j \end{bmatrix} \tag{3.10}$$

$$\delta \mathbf{u}(t) = -\frac{\partial \mathbf{f}}{\partial \mathbf{u}}^{\mathrm{T}} \mathbf{\Lambda}(t, t_0) \begin{bmatrix} \delta \mathbf{x}_i \\ \delta \mathbf{x}_j \end{bmatrix} \tag{3.11}$$

The following terms are defined for ease of notation:

$$\boldsymbol{\mu}_{ij} = \begin{bmatrix} \boldsymbol{\mu}_i \\ \boldsymbol{\mu}_j \end{bmatrix} \tag{3.12}$$

$$\mathbf{P}_{ij} = \begin{bmatrix} \mathbf{P}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_j \end{bmatrix} \tag{3.13}$$

$$\delta \mathbf{z}_{ij} = \begin{bmatrix} \delta \mathbf{x}_i \\ \delta \mathbf{x}_j \end{bmatrix} \tag{3.14}$$

Note that $\boldsymbol{\mu}_{ij} \in \mathbb{R}^{12\times 1}$ is a constant vector for each (i,j) boundary condition pair. Similarly, $\mathbf{P}_{ij} \in \mathbb{R}^{12\times 12}$ is a constant matrix for each (i,j) boundary condition pair. The zero-mean random variable $\delta \mathbf{z}_{ij}$ is sampled from the i and j boundary condition uncertainties such that $\delta \mathbf{x}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_i)$ and $\delta \mathbf{x}_j \sim \mathcal{N}(\mathbf{0}, \mathbf{P}_j)$.

Since $\delta \mathbf{z}_{ij}$ is independent of time τ, the approximate quadratic control cost for a single term of the GMM connecting initial distribution i to final distribution j can be expressed

as seen in Eq. (3.15) by substituting the definitions in Eqs. (3.16-3.18) into Eq. (3.4).

$$J_{ij} \approx J^* + \boldsymbol{\omega}(t_f,t_0)^{\mathrm{T}}\delta\mathbf{z}_{ij} + 2\boldsymbol{\mu}_{ij}^{\mathrm{T}}\boldsymbol{\Omega}(t_f,t_0)\delta\mathbf{z}_{ij} + \boldsymbol{\omega}(t_f,t_0)^{\mathrm{T}}\boldsymbol{\mu}_{ij} \tag{3.15}$$
$$+ \boldsymbol{\mu}_{ij}^{\mathrm{T}}\boldsymbol{\Omega}(t_f,t_0)\boldsymbol{\mu}_{ij} + \delta z^{\mathrm{T}}\boldsymbol{\Omega}(t_f,t_0)\delta\mathbf{z}_{ij}$$
$$J^* = \frac{1}{2}\int_{t_0}^{t_f} \mathbf{u}^*(\tau)^{\mathrm{T}}\mathbf{u}^*(\tau)\mathrm{d}\tau \tag{3.16}$$
$$\boldsymbol{\omega}(t_f,t_0) = \int_{t_0}^{t_f} \boldsymbol{\Lambda}(\tau,0)^{\mathrm{T}}\frac{\partial\mathbf{f}}{\partial\mathbf{u}}\mathbf{u}^*(\tau)\mathrm{d}\tau \tag{3.17}$$
$$\boldsymbol{\Omega}(t_f,t_0) = \frac{1}{2}\int_{t_0}^{t_f} \boldsymbol{\Lambda}(\tau,0)^{\mathrm{T}}\frac{\partial\mathbf{f}}{\partial\mathbf{u}}\frac{\partial\mathbf{f}}{\partial\mathbf{u}}^{\mathrm{T}}\boldsymbol{\Lambda}(\tau,0)\mathrm{d}\tau \tag{3.18}$$

J^* is the quadratic control cost of the optimal trajectory $(\mathbf{x}_\mu^*(t), \mathbf{p}_\mu^*(t))$ without boundary-condition variations. The terms $\boldsymbol{\omega}(t,t_0)$ and $\boldsymbol{\Omega}(t,t_0)$ are defined relative to the optimal trajectory $(\mathbf{x}_\mu^*(t), \mathbf{u}_\mu^*(t))$ connecting the centroid states. Combining terms, the cost J_{ij} in Eq. (3.15) can be re-written in a format similar to the strictly Gaussian result from Holzinger et al. [24], as shown in Eq. (3.19).

$$J_{ij} = J^* + \delta J_{ij} \tag{3.19}$$
$$\approx J^* + \boldsymbol{\omega}(t_f,t_0)^{\mathrm{T}}(\boldsymbol{\mu}_{ij} + \delta\mathbf{z}_{ij}) + (\boldsymbol{\mu}_{ij} + \delta\mathbf{z}_{ij})^{\mathrm{T}}\boldsymbol{\Omega}(t_f,t_0)(\boldsymbol{\mu}_{ij} + \delta\mathbf{z}_{ij}) \tag{3.20}$$

where δJ_{ij} is the variational control cost due to uncertainties in the i^{th} initial and j^{th} final terms of the boundary conditions. From Holzinger et al. [24] Appendix B, the analytic first and second moments of the variational control cost, δJ_{ij}, are:

$$\mathbb{E}\left[\delta J_{ij}\right] = \mu_{J,ij} \approx \boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\mu}_{ij} + \boldsymbol{\mu}_{ij}^{\mathrm{T}}\boldsymbol{\Omega}_{ij}\boldsymbol{\mu}_{ij} + \mathrm{Tr}\left[\boldsymbol{\Omega}_{ij}\mathbf{P}_{ij}\right] \tag{3.21}$$
$$\mathbb{E}\left[\delta J_{ij}^2\right] = \sigma_{J,ij}^2 \approx \left(\boldsymbol{\omega}^{\mathrm{T}} + 2\boldsymbol{\mu}_{ij}^{\mathrm{T}}\boldsymbol{\Omega}\right)^{\mathrm{T}}\mathbf{P}_{ij}\left(\boldsymbol{\omega}^{\mathrm{T}} + 2\boldsymbol{\mu}_{ij}^{\mathrm{T}}\boldsymbol{\Omega}\right) + 2\mathrm{Tr}\left[\boldsymbol{\Omega}\mathbf{P}_{ij}\boldsymbol{\Omega}\mathbf{P}_{ij}\right] \tag{3.22}$$

The control cost considering all i initial boundary conditions and j final boundary condi-

tions is then the weighted sum of the individual costs between each i and j:

$$J \approx J^* + \sum_{i=1}^{N_0} \sum_{j=1}^{N_f} w_i w_j \delta J_{ij} \tag{3.23}$$

where w_i is the weight of the i^{th} initial boundary condition and w_j is the weight of the j^{th} final boundary condition. Thus, the analytic expected value of the initial and final Gaussian sum boundary condition may be written as:

$$\begin{aligned}
\mathbb{E}\left[J\right] &\approx \mathbb{E}\left[J^* + \sum_{i}^{N_0} \sum_{j}^{N_f} w_i w_j \delta J_{ij}\right] \\
&\approx J^* + \sum_{i}^{N_0} \sum_{j}^{N_f} w_i w_j \mathbb{E}\left[\delta J_{ij}\right] \\
&\approx J^* + \sum_{i}^{N_0} \sum_{j}^{N_f} w_i w_j \mu_{J,ij}
\end{aligned} \tag{3.24}$$

which leads to the expected value of the total control cost distribution in Eq. (3.25):

$$\mu_J \approx J^* + \sum_{i}^{N_0} \sum_{j}^{N_f} w_i w_j \left(\boldsymbol{\omega}^{\mathrm{T}} \boldsymbol{\mu}_{ij} + \boldsymbol{\mu}_{ij}^{\mathrm{T}} \boldsymbol{\Omega} \boldsymbol{\mu}_{ij} + \mathrm{Tr}\left[\boldsymbol{\Omega} \mathbf{P}_{ij}\right]\right) \tag{3.25}$$

This yields the approximate expected value of the control cost distribution connecting two GMM boundary conditions, and therefore can be used to approximate control cost distributions for non-Gaussian boundary conditions. Note that, in the case where the boundary conditions each consist of a single Gaussian component, the summation and weighting terms drop out and $\boldsymbol{\mu}_{ij} = 0$, recovering the Gaussian expressions for control cost distributions from Holzinger et al. [24].

Note that the approximations shown here assume that the Gaussian components are within the region of convergence of the map $\boldsymbol{\Lambda}(t_0, t_f)$. This assumption is more likely to be valid when the Gaussian components of the GMM are more closely packed, as is the case after propagation of an initially Gaussian state estimate. However, if the distance be-

tween the GMM components increases, this assumption may no longer be valid. In this case, the formulation may be modified by partitioning the GMM as seen in Fig. 3.3, approximating the original PDF as a mixture of GMMs. Preliminary work [68] demonstrated

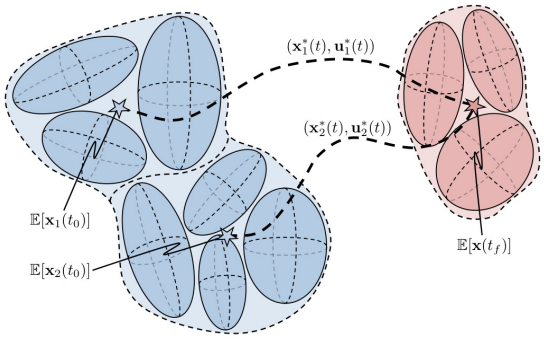


Figure 3.3: Partitioned PDF using multiple trajectories

this intermediate method, linearizing about multiple optimal trajectories computed between N_0 initial and N_f final Gaussian boundary conditions. Similarly, the theory above can be applied using optimal trajectories connecting N_0 initial condition GMMs and N_f final condition GMMs, such that the components of each GMM partition are within the region of convergence of the relevant $\mathbf{\Lambda}(t_0, t_f)$. This approach increases computational complexity due to the combinatorial nature of this approach, requiring $N_0 \times N_f$ optimal trajectories, but is still more efficient than fully non-linearized optimal trajectory sampling. The combinatorial method may be preferred based on the boundary conditions provided, and cost distributions developed using a combinatorial method are still valid for use in the remaining theory.

3.4 Anomaly Detection using Binary Hypothesis Testing

The existing implementation of control cost maneuver detection forms an anomaly hypothesis using only the control cost distribution from the maneuvered trajectory, computing the probability that the deterministic optimal control cost is detectable over the uncertainty [69]. While the cost distribution for the maneuvering hypothesis is well defined from the

previous method, an opposing distribution for the null hypothesis was not developed. As a result, the previous method does not allow specification of acceptable error rates, making comparisons to existing distance metrics, such as Mahalanobis distance, incomplete.

This study modifies the anomaly detection method by applying a binary hypothesis structure using separate PDFs for the null and alternative hypotheses. In order to apply binary hypothesis testing, a pair of mutually exclusive hypotheses, a null hypothesis $\mathcal{H}_0$ and an alternative hypothesis $\mathcal{H}_1$, must be developed along with corresponding probability density functions, $f_0(x)$ and $f_1(x)$. Given PDFs for the binary hypotheses, the following probabilities can be computed:

$$P_{FA} = \int_{\nu}^{\infty} f_0(x)\mathrm{d}x \tag{3.26}$$

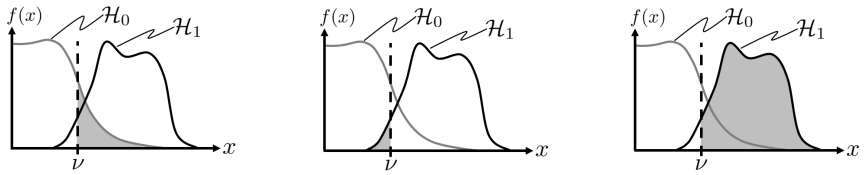
$$P_{FN} = \int_{-\infty}^{\nu} f_1(x)\mathrm{d}x \tag{3.27}$$

$$P_D = \int_{\nu}^{\infty} f_1(x)\mathrm{d}x \tag{3.28}$$

where ν is a selected threshold, P_{FA} is the probability of false alarm (Type I error), P_{FN} is the probability of false negative or missed detection (Type II error), and P_D is the probability of detection. Noting that each integral shares the same integration threshold ν, the probabilities in Eqs. (3.27-3.28) can be related by $P_{FN} + P_D = 1$. Also note that, assuming both PDFs have been normalized, Eq. (3.27) is equivalently defined as the cumulative distribution function (CDF) for the $\mathcal{H}_1$ hypothesis, while Eqs. (3.26) and (3.28) are complementary CDFs, or tail distributions. Figure 3.4 notionally depicts the computation of these probabilities from the $\mathcal{H}_0$ and $\mathcal{H}_1$ PDFs. The threshold ν is typically selected in order to match a maximum allowable rate of false alarms, P_{FA}, using the null hypothesis PDF and Eq. (3.26).

For the anomaly detection problem, the binary hypotheses are formulated as follows:

$\mathcal{H}_0$ (Null Hypothesis): Observed trajectory adequately explained by quiescent state propagation with boundary condition uncertainty.



(a) P_{FA} (b) P_{FN} (c) P_D

Figure 3.4: One-sided Binary Hypothesis Testing Illustration

$\mathcal{H}_1$ (Alternative Hypothesis): Observed trajectory not adequately explained by quiescent state propagation with boundary condition uncertainty.

Recalling the scenario from Fig. 3.1, the binary distributions for the null and alternate hypotheses are drawn from the maneuvering and non-maneuvering trajectories, respectively.

In the event that the null hypothesis is accepted, the change in state between observations is explained solely by uncertainty in the boundary conditions. Therefore, the associated control cost PDF $f_0(J)$ is derived from a quiescent propagated trajectory:

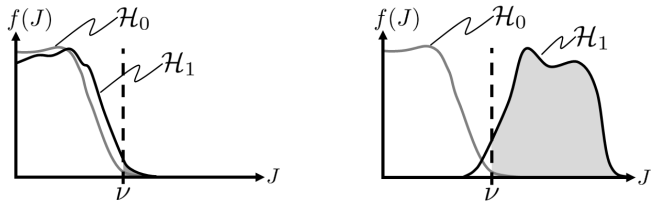
$$f_0(J) \approx \sum_{i=1}^{N_0} \sum_{j=1}^{N_0} w_i w_j \mathcal{N}\left(\mu_{J,ij}, \sigma_{J,ij}^2\right) \tag{3.29}$$

The GMM approximation for control distance is applied along the quiescent trajectory such that each initial Gaussian component $i = 1, ..., N_0$ has a corresponding final component $j = 1, ..., N_0$.

Alternately, if the null hypothesis is rejected, then the change in state between observations is too large to be adequately explained solely by uncertainty in the boundary conditions. The associated control cost PDF $f_1(J)$ is derived from the maneuvering trajectory:

$$f_1(J) \approx \sum_{i=1}^{N_0} \sum_{j=1}^{N_f} w_i w_j \mathcal{N}\left(\mu_{J,ij}, \sigma_{J,ij}^2\right) \tag{3.30}$$

Note that here the GMM approximation for control cost distribution is applied along the optimal trajectory connecting the two input UCT boundary conditions, so the final condition



(a) Significant overlap, low P_D (b) Small overlap, high P_D

Figure 3.5: Theoretical scenarios for control cost binary hypotheses

Gaussian components j have changed between Eqns. (3.29) and (3.30).

This formulation was selected by analyzing a number of different theoretical scenarios to ensure the entire space of possible binary hypotheses revealed the desired behavior, as illustrated in Fig. 3.5. For instance, in the case of a small maneuver, the null and alternative hypothesis PDFs will mostly overlap. Since the allowable rate of false alarm, P_{FA}, is likely to be small, the corresponding threshold ν will cause the anomaly probability P_D from Eq. (3.28) to be small as well. However, in the event that a large maneuver has occurred, the alternative hypothesis PDF will be shifted far to the right of the null hypothesis PDF. Using the same P_{FA} and corresponding ν, the anomaly probability P_D is large, indicating that a maneuver has likely occurred.

The following approach applies binary hypothesis testing to spacecraft anomaly detection using control cost PDFs and prescribed acceptable false alarm and missed detection rates:

1) Construct control cost PDFs for binary hypotheses, $f_0(J)$ and $f_1(J)$, using Eqs. (3.29) and (3.30).

2) Using allowable false alarm rate P_{FA}, compute integration threshold cost ν using Eq. (3.26).

3) Compute anomaly probability P_D using Eq. (3.28) with $f_1(J)$ and ν from previous step.

Once P_D is computed, a final thresholding process can be used to determine whether to flag as an anomaly.

4a) If threshold detection probability for anomaly is directly prescribed, $P_{D,thresh}$ is given.

4b) If allowable false negative (missed detection) rate prescribed, $P_{D,thresh} = 1 - P_{FN}$.

5) If $P_D >= P_{D,thresh}$, flag as anomaly.

3.5 Binary Hypotheses for Mahalanobis Distance

The Mahalanobis distance is a measure of the distance between two points within a Gaussian distribution, scaled by the covariance. A metric related to Mahalanobis distance has been proposed by Hill et al. to identify outliers when comparing UCTs [30]. The particular metric shown in Eq. (3.31) compares the new PDF at time t_f, $f_B(\mathbf{x}) = \mathcal{N}(\mathbf{x}_{\mu,B}, \mathbf{P}_B)$ with a previous estimate propagated to time t_f, $f_A(\mathbf{x}) = \mathcal{N}(\mathbf{x}_{\mu,A}(t_f), \mathbf{P}_A(t_f))$. In the case of maneuver detection, $\mathbf{P}_A(t_f) + \mathbf{P}_B$ is the distance matrix accounting for the combined uncertainty of both distributions [30].

$$d_M\left(\mathbf{x}_{\mu,A}, \mathbf{x}_{\mu,B}, \mathbf{P}_A(t_f) + \mathbf{P}_B\right) = \sqrt{\left(\mathbf{x}_{\mu,A} - \mathbf{x}_{\mu,B}\right)^{\mathrm{T}} \left(\mathbf{P}_A(t_f) + \mathbf{P}_B\right)^{-1} \left(\mathbf{x}_{\mu,A} - \mathbf{x}_{\mu,B}\right)} \tag{3.31}$$

Since Mahalanobis distance has been proposed as a potential data association and anomaly detection metric, an analogous formulation for binary hypothesis testing is also developed using Mahalanobis distance. The null hypothesis considers the state distribution $\mathbf{f}_A(\mathbf{x}(t_f))$, the initial distribution propagated quiescently to time t_f. Given a point sampled from distribution $\mathbf{f}_A$ at t_f, namely $\mathbf{x}_A(t_f) \sim \mathbf{f}_A(\mathbf{x}(t_f))$, Eq. (3.32) computes the Mahalanobis distance from this point to the quiescent distribution $\mathbf{f}_A$.

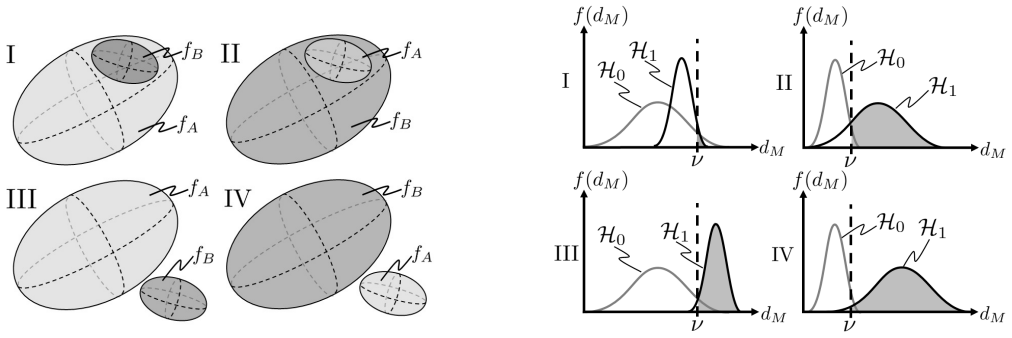
$$d_{M|\mathcal{H}_0} = \sqrt{\left(\mathbf{x}_A(t_f) - \mathbb{E}\left[\mathbf{x}_A(t_f)\right]\right)^{\mathrm{T}} \mathbf{P}_A^{-1} \left(\mathbf{x}_A(t_f) - \mathbb{E}\left[\mathbf{x}_A(t_f)\right]\right)} \tag{3.32}$$

The alternative hypothesis considers the state distribution $\mathbf{f}_B(\mathbf{x}(t_f))$ from the new observation at time t_f. Given a point sampled from distribution $\mathbf{f}_B$, namely $\mathbf{x}_B(t_f) \sim \mathbf{f}_B(\mathbf{x}(t_f))$, Eq. (3.33) computes the Mahalanobis distance from this point to the quiescent distribution $\mathbf{f}_A$.

$$d_{M|H_1} = \sqrt{\left(\mathbf{x}_B(t_f) - \mathbb{E}\left[\mathbf{x}_A(t_f)\right]\right)^{\mathrm{T}} \left(\mathbf{P}_A(t_f) + \mathbf{P}_B\right)^{-1} \left(\mathbf{x}_B(t_f) - \mathbb{E}\left[\mathbf{x}_A(t_f)\right]\right)} \tag{3.33}$$

The use of $(\mathbf{P}_A(t_f) + \mathbf{P}_B)$ as the distance matrix follows the convention set by Hill et al. per Eq. (3.31) to describe the distance between maneuvered and quiescent distributions [30].

Similar to the control cost distance, this interpretation using Mahalanobis distance is justified by considering the theoretical cases of quiescent propagated state distribution $\mathbf{f}_A(\mathbf{x}(t_f))$ and new state distribution $\mathbf{f}_B(\mathbf{x}(t_f))$ as shown in Fig. 3.6. Cases I and III refer to a situation where uncertainty in the initial spacecraft state estimate is high, but uncertainty in the new estimate is reduced. Cases II and IV refer to a situation where uncertainty in the initial estimate is low, but uncertainty in the new estimate is significantly larger. In case I, the propagated distribution envelops the new state distribution, so the Mahalanobis distance PDFs show significant overlap, yielding a low probability of anomaly. In case II, the new state distribution envelops the propagated distribution, so there is once again significant overlap. The mean of the distributions are offset and uncertainty in the new observation is larger, leading to a non-negligible anomaly probability and indicating that an anomaly likely occurred between the observations to cause the change. In both cases III and IV, the state distributions no longer overlap, causing the Mahalanobis distance distributions to be further separate as well, leading to high anomaly probabilities. In each case, this binary hypothesis Mahalanobis distance formulation effectively determines the anomaly probability as designed, providing an analogous formulation for comparison to the control distance metric.



(a) State distributions

(b) Mahalanobis distance PDFs

Figure 3.6: Theoretical scenarios for Mahalanobis distance binary hypotheses

For GMM boundary conditions, Li et al. show that minimizing the Kullback-Leibler divergence yields the distance matrix in Eq. (3.34) [70].

$$\mathbf{P}_\mu(\mathbf{x}) = \left[\frac{\sum_{i=1}^{n} w_i f(\mathbf{x}; \mathbf{x}_{\mu,i}, \mathbf{P}_i) \mathbf{P}_i^{-1}}{\sum_{j=1}^{n} w_j f(\mathbf{x}; \mathbf{x}_{\mu,j}, \mathbf{P}_j)} \right]^{-1} \tag{3.34}$$

Therefore the Mahalanobis distance analog for GMMs is defined in Eq. (3.35) [70].

$$d_{GM}(\mathbf{x}, \mathbf{x}_\mu; \mathbf{P}_\mu^{-1}) = \sqrt{(\mathbf{x} - \mathbf{x}_\mu)^{\mathrm{T}} \mathbf{P}_\mu^{-1} (\mathbf{x} - \mathbf{x}_\mu)} \tag{3.35}$$

In the case where the GMM only has a single component, Eq. (3.34) reduces to the covariance of the individual distribution, and the standard Mahalanobis distance is recovered.

3.6 Simulated Results

The algorithm as described in Fig. 3.2 is implemented in MATLAB to evaluate its performance and effectiveness. Specific implementation details for each portion of this block diagram are explained below.

3.6.1 Trajectory Optimization

Using the input UCT state PDFs, the deterministic two-point boundary value problem between the expected-value states is formulated into an optimization problem, discretizing the simulation into a user-defined number of time-steps, which is solved using the constrained minimization function in MATLAB, *fmincon()*. The decision variable for this minimization is a stacked vector of the thrust accelerations at each discrete time instant, and the thrust accelerations are held constant for each discrete time step. Keplerian dynamics, along with a number of user-selectable perturbation accelerations (J_2, J_{22}, J_3, lunar gravitational, and solar gravitational perturbations), are enforced between steps of the trajectory as equality constraints to ensure the generated trajectory dynamics are accurate. Since the partial derivatives of the dynamics with respect to the decision variables (thrust accelerations) are well known, the gradient of the constraint is supplied to the optimization function to improve convergence. The output of the direct optimization step is a nominal optimal direct trajectory of states and controls connecting the UCTs, which is refit to an optimal indirect trajectory of states and costates. The generated optimal trajectory is validated using the nonlinear dynamics to numerically integrate the proposed control vector and quantify the error between the integrated final condition and the specified final UCT boundary condition. In addition, the quiescent trajectory from the initial condition is computed under the same dynamics with the assumption that no control input is used.

3.6.2 PDF Generation

To accurately construct the control cost PDFs for GMM boundary conditions, a localized Monte-Carlo-like method is employed to sample from the boundary conditions and apply the GMM control cost approximation. This is done by selecting one of the boundary condition Gaussian components from the GMM randomly with a probability based on the weighting for that particular Gaussian component. Once the Gaussian components are selected for the initial and final states, the deviation $\boldsymbol{\mu}_{ij}$ and covariance $\mathbf{P}_{ij}$ terms are known.

The zero-mean random variable $\delta\mathbf{z}_{ij}$ is selected using the covariance information for the chosen Gaussian components. With all the required terms gathered, the quadratic cost is computed using Eq. (3.4). When generating the numerical PDFs, if a negative-valued cost is generated, that sample is discarded and another sample is drawn. The number of samples required to construct this distribution depends on the scenario. In this study, the control cost distribution is sampled $n_{samp} = 10000$ times, and the samples are used to construct the approximate control cost PDF. While exhaustively sampling a 12-dimensional space such as $\delta\mathbf{z}_{ij}$ is computationally restrictive, this lower sampling was shown in previous studies to be sufficient for reconstructing the 1-dimensional control cost and Mahalanobis distance distributions in a similar scenario [68]. This process is performed using both the quiescent and maneuvering trajectories to form both null hypothesis $f_0(J)$ and alternative hypothesis $f_1(J)$ PDFs. This process could equivalently be performed analytically by forming a GMM using $\mu_{J,ij}$, and $\sigma_{J,ij}$ from Eqs. (3.21-3.22) along with w_i and w_j for each boundary condition. The localized Monte-Carlo-like approach is chosen for simplicity of implementation. Mahalanobis distance PDFs are constructed in a similar manner. Samples from each boundary condition are used in conjunction with Eqs. (3.32-3.33) to develop PDFs for each hypothesis.

3.6.3 Anomaly Detection

Once PDFs are obtained for each hypothesis, anomaly detection is performed the same for both control distance and Mahalanobis distance. CDFs for both null and alternative hypotheses are computed, noting that the lower limit of integration becomes 0 since the distance cannot be negative.

$$F_0(x) = \int_0^x f_0(y)\mathrm{d}y \tag{3.36}$$

$$F_1(x) = \int_0^x f_1(y)\mathrm{d}y \tag{3.37}$$

Applying Eq. (3.26), the input allowable false alarm rate P_{FA} is used to compute the threshold ν by interpolating on the null hypothesis CDF, F_0. Using the threshold ν and applying Eq. (3.28), the probability of anomaly is computed by interpolating on the alternative hypothesis CDF, F_1.

3.7 Gaussian Mixture Model Approximation Validation

The GMM approximation promises a more computationally tractable method for addressing non-Gaussian boundary conditions; however, it still must provide an accurate reconstruction of the uncertainty cost distribution. To validate the GMM approximation of control cost distributions, a synthetic scenario is constructed for a GEO spacecraft with a non-Gaussian boundary condition state distribution, represented using GMMs. The orbital elements for the boundary conditions are listed in Table 3.1. This particular scenario is selected to illustrate the ability to generate and handle non-Gaussian control cost distributions using the binary hypothesis testing approach outlined above. The initial condition occurs 30 minutes before the ascending node passage, and the 3 Gaussian terms vary only in inclination. The final condition occurs 30 minutes after the ascending node passage, with 2 Gaussian terms varying only in inclination. Note that the asymmetry in the GMMs ($N_0 = 3, N_f = 2$) is entirely allowed by the GMM approximation formulation. These particular boundary conditions represent a 1 hour observation gap wherein the observed spacecraft has performed a small inclination correction. The $1-\sigma$ boundary condition uncertainties are initialized at 10 meters in position and 10 centimeters-per-second in velocity. These values are selected to generate multi-modal cost distributions when scaled by α to emphasize the generality of the analytical contributions.

For validation purposes, an alternate method is used to generate control cost distributions through direct sampling of the boundary conditions and non-linearized trajectory optimization. For each run, a state is selected at random from the initial and final GMM boundary conditions. The optimal control between the chosen boundary conditions is com-

Table 3.1: Boundary condition orbital elements for GMM inclination change

	Initial Condition			**Final Condition**	
Parameter	$i = 1$	$i = 2$	$i = 3$	$j = 1$	$j = 2$
Weighting, w	0.34	0.33	0.33	0.5	0.5
Semi-major Axis, a (km)	42164	42164	42164	42164	42164
Eccentricity, e	0	0	0	0	0
Inclination, i (deg)	0.015	0.025	0.035	0.00	0.01
Long. of Asc. Node, Ω (deg)	0	0	0	0	0
Arg. of Periapsis, ω (deg)	0	0	0	0	0
True Anomaly, θ (deg)	352.5	352.5	352.5	7.5	7.5

puted, and the resulting control cost is collected before selecting another pair of boundary conditions. This process is repeated $n_{samp} = 5000$ times to sample the control cost space, binning the results to construct a control cost PDF. Note that $\delta\mathbf{z}_{ij}$ is a 12-dimensional space, so the curse of dimensionality computationally restricts complete sampling; however, previous work showed 5000 samples to sufficiently reconstruct the 1-dimensional control cost distributions [68]. Since this method does not make the simplifying assumption of linearizing about a best-estimate trajectory, it more accurately generates the actual control cost distribution between the GMM boundary conditions, at the expense of much longer computation times.

Figure 3.7 shows the normalized PDF and associated CDF resulting from both these methods. The solid lines indicate the non-linearized propagation, labeled "True," and the dashed lines indicate the control cost distribution generated using the GMM approximation method, labeled "Aprx." Inspection of the normalized PDFs shows significantly non-Gaussian control cost distributions, as expected given the separation in the boundary condition components: transitioning from initial component $i = 1$ to final component $j = 1$ uses significantly less fuel than from initial component $i = 3$ to final component $j = 1$. Despite the differences in approach, the control cost PDFs and CDFs agree well between the non-linearized and approximated methods, demonstrating the efficacy of contribution 1: the GMM control cost approximation.

Since the CDFs are nearly identical, the binary hypothesis testing algorithm yields a

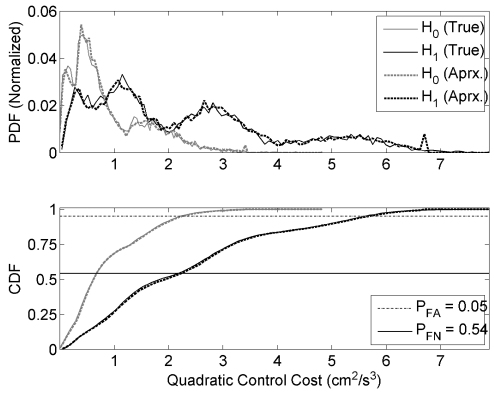


Figure 3.7: Comparison of control cost distributions using the non-linearized and approximate methods

Table 3.2: Timing comparison for validation scenario

Method	**Complexity**	**Actual Time (s)**
Non-Linearized	$\mathcal{O}(n_{samp})$	53450
GM Approximation	$\mathcal{O}(c)$	2.835

nearly identical probability of anomaly using either method. To further demonstrate the applicability of contribution 2, the binary hypothesis testing algorithm, a sample calculation of anomaly probability is illustrated using Fig. 3.7. Assuming an allowable false alarm rate of 5%, or $P_{FA} = 0.05$, the threshold ν is calculated from Eq. (3.26) by interpolating on the $\mathcal{H}_0$ CDF to find the cost where $F_0(J) = 1 - P_{FA}$. Given the threshold cost, the false negative probability is calculated using Eq. (3.27) by interpolating on the $\mathcal{H}_1$ CDF, such that here $P_{FN} = F_1(\nu) = 0.54$. Recalling that $P_D = 1 - P_{FN}$, the probability of anomaly for this scenario is therefore $P_D = 0.46$, or 46%.

Given the similarity of the PDFs, the computational complexity savings of the GMM approximation method offers a significant benefit inherent in the first contribution. Table 3.2 shows the time required to perform the full anomaly detection algorithm for both of these scenarios. These results are in-line with expectation, as the trajectory optimizer is expected to be the computational bottleneck. The non-linearized validation method runs in roughly $\mathcal{O}(n_{samp})$ time since a trajectory optimization is required for each sample (with

small variations in the time required for each sample based on the boundary conditions). In comparison, the GMM approximation method only requires a single trajectory optimization, and therefore runs in nearly constant time, or $\mathcal{O}(c)$. Therefore, when considering the computational complexity improvements, the GMM approximation method significantly outperforms the non-linearized approach while maintaining accuracy in the control cost PDF and therefore performing similarly in anomaly detection.

In order to test the efficacy of the Mahalanobis distance GMM method proposed, the same scenario is used to develop PDFs and CDFs for the binary hypotheses using Mahalanobis distance. In comparison to the control distance method, the Mahalanobis distance PDFs from the individual Gaussian components are more distinct, so the combined PDF forms narrow peaks in an extremely non-Gaussian manner as seen in Fig. 3.8. In particular, the null hypothesis distribution covers a wide range of Mahalanobis distance values, since the quiescent propagated distributions are very dissimilar. The wide null hypothesis distribution causes the distance threshold for 95% anomaly confidence to be large, yielding an effectively 0% anomaly detection, similar to Case I in Fig. 3.6.

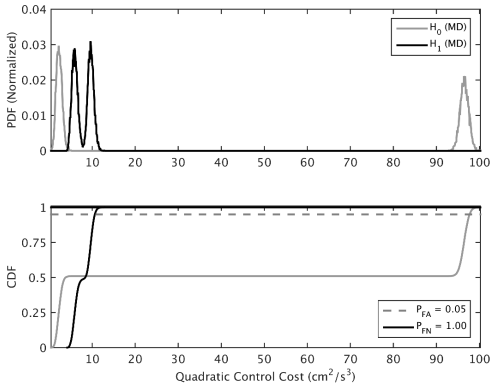


Figure 3.8: Mahalanobis distance distributions for non-Gaussian boundary conditions

Additionally, the Mahalanobis distance GMM formulation is significantly more computationally expensive due to the complicated combined covariance calculation in Mahalanobis distance PDF generation. Without this calculation, in the purely Gaussian case,

Mahalanobis distance is computed quicker than control distance, since Mahalanobis distance does not require trajectory optimization. Adding in the combined covariance calculation for non-Gaussian boundary conditions significantly increases computation time over the control distance method, as shown in Table 3.3.

Table 3.3: Timing comparison by algorithm segment

Algorithm Segment	**Control Distance (s)**	**Mahalanobis Distance (s)**
Trajectory Optimization	1.845	0
PDF Generation	0.762	11.347
Anomaly Testing	0.229	0.225
Total	2.835	11.572

Using the formulations developed in the Theory section, these results demonstrate the ability to take non-Gaussian boundary conditions and form PDFs and CDFs for use in binary hypothesis testing. The control distance method proved to be better in the detection of this particular maneuver, but the following simulation results will delve deeper into performance of each algorithm in different maneuver detection scenarios.

3.7.1 Inclination Change

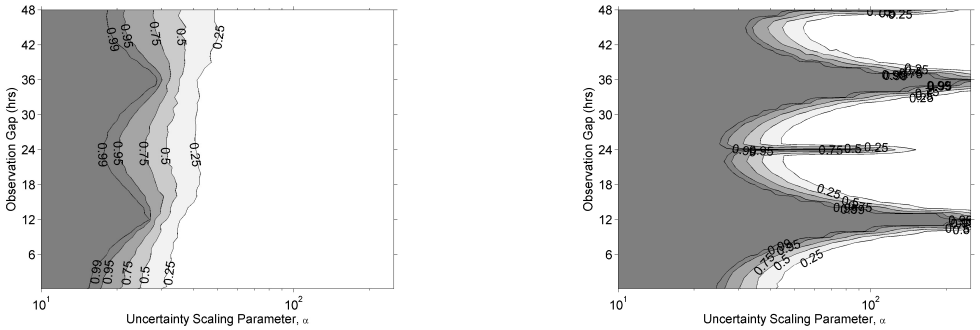
Having shown the ability to accurately construct control cost PDFs for non-Gaussian boundary conditions, the anomaly detection algorithm is next evaluated by parameterizing the problem to assess sensitivity. For the remaining results in this paper, the boundary conditions are simplified to single Gaussian distributions for ease of parameterization and discussion, but similar simulations could be performed using GMMs as shown above.

A simulated scenario is constructed to emulate an inclination change performed at GEO, termed a North-South station-keeping maneuver in operations. Typically, satellites at GEO will be placed into orbit slots and given allowable deviations in the North-South and East-West directions. This particular scenario employs a 0.02 degree inclination change, similar in magnitude to that observed in the available real-world data. The goal of this scenario is to analyze the sensitivity of both the control distance and Mahalanobis dis-

tance binary hypothesis testing methods to changes in observation gap, uncertainty, and false alarm rate. Each of these three parameters are varied systematically in simulation for evaluation.

The observation gap is varied between 10 minutes and 48 hours. Each element of the 6×6 covariance matrix is varied by an uncertainty scaling parameter α. The $1-\sigma$ boundary condition position uncertainty is initialized at 1 meter and varied up to 250 meters using the scaling parameter α. Likewise, the $1-\sigma$ boundary condition velocity uncertainty is also varied between 1 centimeter-per-second and 250 centimeters-per-second using the scaling parameter α. The prescribed false alarm rate is varied between 0.5% and 10%, or $P_{FA} = 0.005$ to 0.10. For each combination of observation gap, boundary condition uncertainty, and prescribed false alarm rate, the anomaly probability is computed using both the control distance and Mahalanobis distance metrics. This produces a 4-dimensional dataset, $(P_D : \delta t, \alpha, P_{FA})$, which is best viewed as a set of contour plots using slices of constant P_{FA}.

Figure 3.9(a) presents a contour plot for a subset of the control distance sensitivity study data using a false alarm probability $P_{FA} = 0.05$. Note that, though the uncertainty scaling parameter α is varied from $1 \leq \alpha \leq 250$, a subset of this range ($\alpha >= 10$) has been plotted to highlight trends. Toward the left of the uncertainty axis, boundary condition uncertainty is low, so it is easier to distinguish between the non-maneuvered and maneuvered trajectories; therefore, anomaly probability is high ($P_D \approx 1$). Increasing boundary condition uncertainty causes more overlap in the control distance distributions, introducing values of $P_D < 1$ as the propagated uncertainty is large enough to account for the new observation. Additionally, a slight increase in anomaly probability can be seen around half-orbit period observation gaps, which coincides with the furthest out-of-plane difference between the quiescent and maneuvered trajectories. At these points, boundary condition uncertainty must be significantly greater to cause overlap between the trajectories, showing increased sensitivity to the maneuver at that observation condition.



(a) Control distance

(b) Mahalanobis distance

Figure 3.9: Probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated inclination change maneuver, $P_{FA} = 0.05$.

Additional contour plots for different false alarm probabilities can be seen in Fig. A.1 of Appendix A. Each subfigure shows a contour plot of anomaly probability for a prescribed false alarm rate, where the darker colors indicate a higher probability of anomaly. Here, false alarm rates of $0.5\%, 1\%, 5\%$, and 10% were used. Direct comparisons can be drawn between the various contour plots in this study by selecting plots with the same P_{FA} and selecting a point $(\alpha, \Delta t)$, comparing the P_D value at that point on each of the plots. As expected, the algorithm declares higher probabilities of anomaly for a fixed observation gap and uncertainty as the allowable rate of false alarm increases because the threshold for maneuver detection is lessened. The remainder of the trends are consistent regardless of false alarm probability, so they are relegated to Appendix A for reference.

Figure 3.9(b) presents the same anomaly probability data using Mahalanobis distance distributions, with a more complete set of plots featured in Fig. A.2 of Appendix A. Similar trends observed with control distance can also be noted for Mahalanobis distance: increasing the rate of false positives increases the probability of anomaly, and increasing uncertainty yields lower confidence in anomaly detection. However, Mahalanobis distance also shows significant variability, specifically with resonances near orbital half-periods (shown 12 and 36 hours for GEO). At these points, the non-maneuvered and maneuvered orbits,

for $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively, are at their furthest separation ($\mathcal{H}_0$ at its maximum out-of-plane distance), making the state difference large. Additionally, a smaller spike can be seen at the orbit periods (24 and 48 hours) where both orbits lie in the equatorial plane, due to significant velocity vector differences between the $\mathcal{H}_0$ and $\mathcal{H}_1$ orbits despite similar positions.

It can be seen that, at some observation gap durations, Mahalanobis distance is significantly more sensitive to inclination change maneuvers, as evidenced by the higher probabilities of anomaly at higher uncertainty. However, this is not always the case, as at times the control distance metric and Mahalanobis distance metric are comparable in detection probabilities. This variation with observation gap time is a dangerous aspect of using Mahalanobis distance to develop maneuver detection thresholds. Since Mahalanobis distance is very dependent upon the observation time, this requires a better understanding and more careful consideration of the dynamics of the spacecraft and what kind of maneuver it would perform at what point in the orbit. Control distance, however, is more consistent with time, allowing the development of a more general-use threshold for the North-South station-keeping maneuver at GEO. Additionally, the reliability of uncertainty quantification is a concern when considering the uncertainty sensitivity advantage of Mahalanobis distance. The Operational Application section below discusses a synergistic implementation using both control distance and Mahalanobis distance to leverage the advantages of both methods.

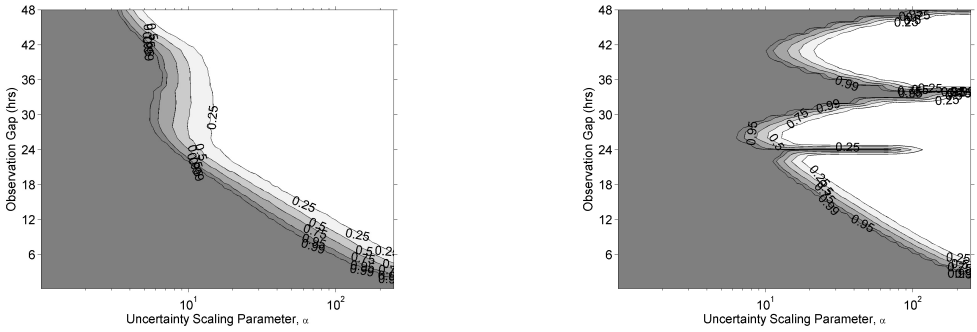
Alternate test cases (omitted from this paper for space and uniformity between examples) showed that both algorithms were more sensitive to velocity uncertainty than position uncertainty. For instance, in the control distance test cases shown here, the algorithm struggles to detect maneuvers above $\alpha = 100$, which corresponds to 100 meters in position uncertainty and 100 centimeters-per-second in velocity uncertainty. Manually changing the boundary condition uncertainties to include cases lower velocity uncertainty showed improved anomaly detection performance even at higher position uncertainty.

3.7.2 Phasing Maneuvers

A similar sensitivity study is conducted for a phasing maneuver at GEO. This maneuver type is termed East-West station-keeping, as it refers to the satellite maintaining a specific longitude over the Earth. A 0.1 degree change in longitude is prescribed, selected to represent drifting completely across a ± 0.05 degree GEO slot. The same parameters (observation gap, boundary condition uncertainty, and false alarm rate) are varied over the same ranges.

Figures 3.10(a) and 3.10(b) present the probability of anomaly contours for the orbit phasing maneuver for both control distance and Mahalanobis distance, respectively. More complete sets of data are featured in Appendix A, Fig. A.3 for control distance and Fig. A.4 for Mahalanobis distance. Once again, some trends hold true for both control distance and Mahalanobis distance: increasing uncertainty decreases anomaly probability while increasing allowable false alarm rate increases anomaly probability. In this case, however, both methods show significant variation with observation gap. Performing a 0.1 degree longitude change over 10 minutes requires significantly more fuel than over 24 hours, where the spacecraft can more effectively utilize natural dynamics and slightly adjust its semi-major axis to transfer to a different point in the orbit. Therefore, maneuvers of this kind over small observation gaps are much easier to detect.

An interesting note in this scenario is that control distance either matches or outperforms Mahalanobis distance for much of the first orbit period, predicting higher probabilities of anomaly for the same uncertainty, observation gap, false alarm rate triplet. However, as the observation gap increases, Mahalanobis distance again shows considerable variation, yielding higher anomaly probabilities than control distance at 34 hours and lower probabilities at 26 hours. Once again, the observation time can be seen to be a significant factor in anomaly detection for Mahalanobis distance, requiring knowledge of the spacecraft's position in its orbit. For control distance, while it is more sensitive to time for East-West maneuvers than North-South, it still shows improved consistency in medium-duration cases,



(a) Control distance

(b) Mahalanobis distance

Figure 3.10: Probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated phasing maneuver, $P_{FA} = 0.05$.

albeit with lower anomaly sensitivity.

3.8 Empirical Data Results

To complement the simulated scenarios, the algorithm is also tested using real operational data, the availability of which drove the construction of the simulated inclination change scenario. The real data, taken from observations of the Galaxy 15 geostationary satellite by the Wide Area Augmentation System (WAAS), spans a month of operation and includes Earth-centered Earth-fixed (ECEF) position and velocity, as well as radial, in-track, and cross-track (RIC) acceleration, as seen by a rotating Hill frame attached to the spacecraft. WAAS is an extremely accurate navigation system that uses a network of ground-based reference stations to measure small variations in GPS satellite signals to develop deviation corrections (DCs). The DCs are then broadcast by GPS satellites to improve position accuracy calculations for WAAS-enabled GPS receivers. The WAAS data is regarded as the "truth" state since more accurate truth (e.g. maneuver information obtained directly from the operators) is not available.

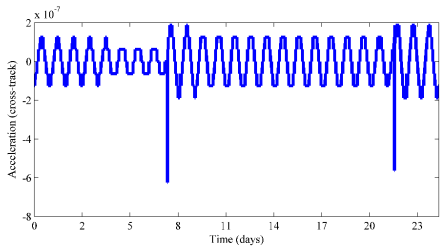
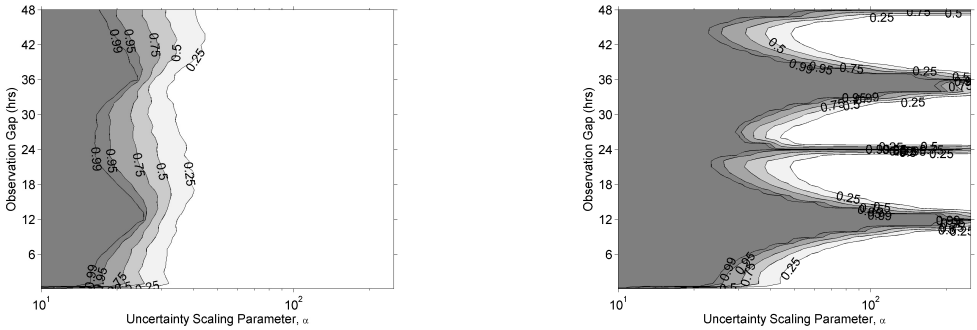


Figure 3.11: Cross-track acceleration for Galaxy 15 satellite, real data

3.8.1 Station-Keeping

Figure 3.11 shows the cross-track acceleration data for the empirical dataset. Inspection of the acceleration data reveals two large anomalous cross-track acceleration events, candidates for North-South inclination station-keeping maneuvers, during days 7 and 22. The selected maneuver, the peak during day 7, resulted in a 0.03 degree inclination change. Simulation initial and final conditions are selected corresponding to the desired observation gap such that the maneuver is always in the middle of the selected time span. For instance, for a 6 hour observation gap, the initial condition is the spacecraft state 3 hours before the maneuver, and the final condition is the state 3 hours after the maneuver. The real-world data is analyzed in a similar manner to the synthetic data by varying observation gap, boundary condition uncertainty, and prescribed false alarm rate.



(a) Control distance

(b) Mahalanobis distance

Figure 3.12: Probability of anomaly contours vs uncertainty scaling parameter and observation gap, real-data (WAAS) inclination change maneuver, $P_{FA} = 0.05$.

The results of the real data sensitivity study are seen in Fig. 3.12(a) and 3.12(b) for control distance and Mahalanobis distance, respectively. This data is nearly identical to that of the simulated inclination change scenario, showing the same trends with respect to all varied parameters. In this scenario, Mahalanobis distance is still more sensitive in its detection of the maneuver at higher uncertainties. However, control distance remains more consistent with respect to observation gap, allowing for improved application to arbitrary space objects without requiring specific knowledge of the object's spot in its orbit. The agreement between these results and Fig. 3.9(a) and 3.9(b) lends confidence to the applicability of this approach in an operational setting. Additional results using different false alarm rates are presented in Appendix A, Fig. A.5 and A.6 for control distance and Mahalanobis distance, respectively. As with the synthetic scenarios, increasing P_{FA} lowers the threshold for anomaly probability calculation and thus increases anomaly probability. The remaining observation gap and uncertainty trends remain unaffected by P_{FA}.

3.9 Operational Application

The simulation results presented in this paper explore a range of different maneuver and boundary condition cases; this section condenses these results and provides recommended operational use cases for each anomaly detection method. In situations where the separation between the Gaussian components of the Mahalanobis distance GMMs are greater than their covariances, combined Mahalanobis distance PDFs tend to be segmented (see Fig. 3.8), which significantly reduces the ability to detect anomalies using the Mahalanobis distance method. Moreover, regardless of the GMM component separations, the Mahalanobis distance method incurs a significant computation time penalty for non-Gaussian boundary conditions due to the costly requirement to recompute the relevant covariance at each intermediate propagation time-step. In contrast, the computational burden of control distance is approximately constant regardless of whether or not boundary conditions are Gaussian. In the Gaussian case, control distance takes approximately 2- to 3-times as long to compute,

and Mahalanobis distance is more sensitive in maneuver detection in most but not all cases. The two methods are complementary; when one performs less effectively, the other remains sensitive to anomalies. The WAAS data supports results generated by the simulated cases, providing empirical support for these results.

Based on these observations, the control distance metric is recommended in applications with non-Gaussian boundary condition uncertainties to avoid significant computation penalties, as well as poorly conditioned anomaly detection in more-segmented GMM cases. For Gaussian boundary conditions, Mahalanobis distance is in general more sensitive to anomalies. However, since neither method always dominates in anomaly detection for Gaussian boundary conditions, the control distance method could be implemented alongside Mahalanobis distance without significant computational complexity increase. Implementing control cost in both cases gracefully handles transitions from Gaussian to non-Gaussian boundary conditions, maintaining performance without introducing significant computational complexity.

3.10 Conclusions

The control distance metric provides a natural means of associating spacecraft observations by subsuming the orbital dynamics into the association metric. This effort modifies the control distance anomaly detection approach to address error-rate shortcomings in the single hypothesis method and relax Gaussian boundary condition assumptions. The inclusion of GMM approximations and a binary hypothesis test approach allow control of allowable error rates and enable comparisons using anomalous and quiescent hypotheses. An analogous set of hypotheses is constructed for Mahalanobis distance and extended to allow for non-Gaussian uncertainties. For both Gaussian and non-Gaussian boundary conditions, simulation results show control distance is able to compute anomaly probabilities at an operationally acceptable computational cost. The Gaussian test cases show Mahalanobis distance to be generally but not uniformly more sensitive to anomalies. However,

Mahalanobis distance anomaly detection is less consistent with observation gap, and the Mahalanobis distance approach is less effective in both anomaly detection and computational complexity with non-Gaussian boundary conditions. This study concludes that, the control distance method is preferred for use with non-Gaussian boundary conditions, while both Mahalanobis distance and control distance should be implemented for Gaussian boundary conditions for added robustness.

CHAPTER 4
SPACECRAFT CUSTODY MAINTENANCE AND ANOMALY DETECTION USING EVIDENTIAL REASONING

The previous chapter applies a hypothesis testing approach to improve an existing method for anomaly detection provided the state estimates have already been updated by successful detection. This chapter continues an analysis of anomaly detection by applying hypothesis-based methods to develop a sensor tasking approach to maintain custody, an assigned responsibility to track a specific subset of space objects.

Applied to sensor-tasking, hypothesis-driven approaches enable a predictive mode of tasking to answer specific relevant questions. This is to say, formulating specific hypotheses of interest a priori allows for tasking that anticipates events related to these hypotheses. The decision-making hypotheses inform what evidence is gathered, with preference toward evidence that resolves the hypotheses quickly. Conversely, reactive tasking uses evidence gathered to formulate and interrogate hypotheses on what events may have occurred. Both modes of tasking are necessary for SSA, as not all events may be hypothesized a priori, requiring reactive tasking. However, a hypothesis-driven sensor tasking approach can also plan ahead for potential events and gather evidence to predict or detect them earlier, requiring less time and sensor resources to reconstruct and understand events.

Once the data is gathered from the sensors, it must be fused into a coherent understanding of the environment via association, correlation, and combination [31]. In classical Bayesian approaches, sensor data is used to form deterministic probabilities placed on event hypotheses under the assumption that the only possible realizations of this hypothesis are true or false. However, in complex decision-making contexts, information is not always best-represented in this strictly binary manner, since some evidence for a particular hypothesis might also involve ambiguity. An expert might be able to confirm or refute a

given set of hypotheses, but it cannot attribute belief mass to any propositions for which it is not an expert. For this reason, evidential reasoning methods such as Dempster-Shafer theory quantify this ambiguity in hypothesis knowledge, leading to more realistic modeling of human analyst processes [39, 40, 41]. Allowing attribution of belief mass to subsets of hypotheses, not only singletons, admits a quantifiable ambiguity in the expert's knowledge of that particular hypothesis resolution.

This chapter develops an ignorance-reduction framework for SSA sensor tasking to gather evidence that quickly resolves hypotheses. The specific contributions of this chapter are as follows: 1) a decision-making criterion using ignorance to rank decision sets, 2) a multi-sensor, multi-target tasking approach utilizing the ignorance decision criterion, 3) a rigorous technique for the formulation of belief structures using SSA sensor data as evidence, and 4) an ignorance-based tasking algorithm to address custody maintenance and anomaly discrimination using SSA sensor data.

4.1 Dempster-Shafer Theory

Gathered evidence must be fused into a coherent understanding of the environment via association, correlation, and combination [31]. Multiple methodologies exist for modeling and reasoning in uncertain domains to provide graphical and numerical representations of uncertainty [55]. One prevailing methodology is Bayesian probability theory, which models knowledge about propositions using true-or-false probabilities. In classical Bayesian approaches, evidence is used to form deterministic probabilities placed on event hypotheses. However, probability theory struggles to express ambiguity in proposition knowledge, often due to some ignorance on the part of the expert or evidence source. For this reason, evidential reasoning methods, such as Dempster-Shafer theory, quantify ambiguity, leading to more realistic modeling of human analyst processes [39, 40, 41].

For illustration, consider an expert with vacuous knowledge, or total ignorance, on a proposition. In probability theory, this is often represented using the principle of non-

information: each state in the proposition state space is assigned equal probability. This equally-likely probability mass function can also arise naturally when an expert has certain knowledge that places equal probability on each state. Therefore, the same probability mass function can represent two very different knowledge states, one with wholly ambiguous information and the other with certain but conflicting evidence, due to the inability to encode ambiguity in Bayesian probability [55]

Dempster-Shafer theory [39] is a formalism for assigning belief mass to hypotheses based on available evidence that has gained significant traction in various applications, including classification [42, 43], monitoring and fault detection [44, 45], and decision-making [46]. Dempster-Shafer theory is considered more expressive than Bayesian probability with respect to ambiguity [74], accomplished by allowing assignment of belief mass to non-singleton propositions. The following sections introduce concepts of Dempster-Shafer theory relevant to this work. For a more complete discussion on important developments in Dempster-Shafer theory, Yager and Liu compiled a book of classic works, reviewed by Dempster and Shafer, on the theory of belief functions [75].

In Dempster-Shafer theory, the possible propositions of a given hypothesis form a set called the frame of discernment, Ω. Elements of 2^{Ω}, the set of all subsets of Ω, are referred to as propositions. The frame of discernment must be a set of mutually exclusive and collectively exhaustive propositions; in other words, exactly one proposition must be true [39].

4.1.1 Belief Functions

An expert's subjective belief in each proposition based on the available evidence is represented through a basic probability assignment (bpa). A bpa m, as defined in Eqn. (4.1),

maps a belief mass to each possible proposition:

$$m \,:\, A \mapsto [0,1], \quad A \in 2^{\Omega} \tag{4.1}$$

$$\sum_{A \in 2^{\Omega}} m(A) = 1 \tag{4.2}$$

$$m(\emptyset) = 0 \tag{4.3}$$

Notationally, $\{\omega_1, \omega_2\}$ is equivalent to $\{\omega_1\} \cup \{\omega_2\}$, the disjunctive combination of propositions ω_1 and ω_2, or "ω_1 or ω_2." The constraint in Eqn. (4.2) enforces the mutually exclusive and collectively exhaustive properties as the belief masses must sum to one, while the constraint in Eqn. (4.3) is similar to Kolmogorov's axiom of zero probability for the empty set.

The set of propositions that have non-zero belief mass are the focal set of the associated bpa. For ease of discussion, a number of common bpas are typically defined based on their focal sets. A vacuous bpa is one in which all the belief mass is assigned to Ω, such that $m(\Omega) = 1$, $m(A) = 0 \,\forall\, A \subset \Omega$. A simple bpa is one in which the focal set consists of only two elements: the entire frame of discernment Ω and one other proposition, as in $m(A) = p$, $m(\Omega) = 1 - p$, $m(B) = 0 \,\forall\, B \in 2^{\Omega} \setminus \{A, \Omega\}$.

Using bpas, Shafer defines notions of belief (or support) and plausibility, which form lower and upper bounds respectively on the probability that a proposition is true given the available evidence [21]. The belief in, or support for, proposition $A \in 2^{\Omega}$ is defined in Eqn. (4.4) as the sum of belief masses attributed to A and its subsets.

$$\mathrm{Bel}_m(A) = \sum_{B|B \subseteq A} m(B) \tag{4.4}$$

The plausibility of proposition $A \in 2^{\Omega}$ is defined in Eqn. (4.5) as the sum of the belief

masses for propositions that do not explicitly refute A:

$$\mathrm{Pl}_m(A) = \sum_{B|B\cap A\neq\emptyset} m(B) = 1 - \mathrm{Bel}_m(\neg A) \tag{4.5}$$

where $\neg A = \Omega \setminus A$ is the negation of A, or "not A." The representations of belief mass m, belief Bel, and plausibility Pl are all interchangeable via the above linear relationships [76]. Also note that, since the proposition Ω (the frame of discernment) represents the disjunctive (or-wise) combination of an exhaustive set of propositions, its belief and plausibility must both be equal to one.

4.1.2 Combination Rules

Numerous methods exist for combining bpas from multiple sources to form a fused bpa [77]. The new bpa behaves just like any other bpa, so a fused estimate of belief and plausibility for each proposition can be obtained. Each combination method exhibits different properties, so implementation should carefully consider use-cases of the fused belief function and characteristics of the evidence sources.

A common bpa combination technique is Dempster's conjunctive rule, which is commutative, associative, and admits the vacuous bpa [39]. Dempster's conjunctive rule of combination, shown in Eq. (4.6), is often represented using the $\oplus$ operator. The belief mass attributed to proposition $A \subseteq \Omega$ after combination of bpas from experts i and j is given as:

$$m_{i\oplus j}(A) = (m_i \oplus m_j)(A) = \frac{\sum_{B\cap C=A} m_i(B)m_j(C)}{1-K} \tag{4.6}$$

$$K = \sum_{B\cap C=\emptyset} m_i(B)m_j(C) \tag{4.7}$$

where K is a term that accounts for conflict between the bodies of evidence.

Evidence conflict occurs when two bpas report belief mass in different propositions. For

illustration, consider the frame $\Omega = \{x_1, x_2\}$ and two bpas, m_1 and m_2, with high belief mass in x_1 and x_2, respectively: $m_1(\{x_1\}) = 0.9$, $m_1(\{x_2\}) = 0$, $m_1(\{x_1, x_2\}) = 0.1$ and $m_2(\{x_1\}) = 0$, $m_2(\{x_2\}) = 0.8$, $m_2(\{x_1, x_2\}) = 0.2$. These bpas are in high conflict ($K = 0.72$), and the fused bpas shows less belief mass in in x_1 or x_2 than either bpa alone: $m_{1\oplus 2}(\{x_1\}) = 0.64$, $m_2(\{x_2\}) = .29$, $m_2(\{x_1, x_2\}) = 0.07$. The use of the conflict term K in Eqn. (4.6) has the effect of distributing conflicting evidence to the null-set, but since support cannot be attributed to the null-set (in classical Dempster-Shafer theory), this belief mass is normalized across the relevant propositions [77].

Some uses of Dempster's rule lead to counter-intuitive results in the presence of extreme conflict, an observation typically referred to as Zadehs paradox [78]. However, the scenario in Zadeh's paradox can be resolved by more carefully adhering to Cromwell's Rule, i.e. not assigning a probability of exactly zero or one to any particular prior. This caveat, with the inclusion of the open-world assumption, i.e. admitting that the actual true event might lie outside the theorized set of possible events, led to the development of the Transferable Belief Model (TBM) as a derivative of Dempster-Shafer theory [57]. The constraints of this particular application allow the classical Dempster-Shafer implementation to be appropriate without applying TBM.

It is important to note that Dempster's rule is not idempotent. Subsequent evidence is assumed to be statistically independent of previous evidence. Therefore, when using Dempster's rule, the evidence must be assumed to be distinct; otherwise, repeated evidence will be heavily weighted in the fused belief mass.

Dempster's rule is not the only combination rule for bpas. For instance, Yager developed a related class of combination rules that, like Dempster's rule, are commutative and not idempotent, but Yager's rule is quasi-associative [79, 80]. The primary difference in Yager's method is the use of a separate probability structure, the ground probability assignment, to pool evidence before conversion to a bpa [77]. Instead of normalizing out conflict, belief mass from conflicting evidence is attributed to the universal-set, the frame

of discernment Ω. As such, Yager's rule is also called the unnormalized Dempster's rule, and, indeed, in the case of no conflict both methods yield the same result [77].

Additional combination rules have been developed that do enforce idempotence, which can be employed for non-distinct bodies of evidence. While the above methods are conjunctive (AND-based) in the attribution of evidence, alternate methods employ disjunctive (OR-based) to handle evidence from varying-reliability sources [77].

4.1.3 The Curse of Dimensionality

One primary criticisms of Dempster-Shafer theory involves the curse of dimensionality [47]; the computational complexity of evidence combination generally increases exponentially with the number of hypotheses [40]. To address these computational complexity concerns, various approximation methods have been developed. Gordon and Shortliffe approximate the results of Dempster's rule in the case where evidence can be arranged in a hierarchical tree-structure [48], while Shafer and Logan propose an exact implementation for similar scenarios [40]; both approaches achieve a linear computational complexity. Wilson proposes a Monte-Carlo algorithm for computing belief on a subset of hypotheses with computation time that also grows linearly with the number of hypotheses and belief structures [81], and Kreinovich et al. similarly argue that Monte-Carlo methods allows for feasible implementations of the Dempster-Shafer formalism [49]. Alternately, consonant methods employed by Dubois and Prade impose a chain-like formation on the hypotheses, such that $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \ldots$ to simplify the computation of belief intervals [50]. Bauer compared many of these approximation methods in the context of decision-making, finding that there is no definitive "best" approximation; rather, the approximations often restrict application to cases with low conflict in evidence or collapse the belief intervals into points, which eliminates a major benefit of Demster-Shafer theory in its ability to represent partial ignorance [51].

4.1.4 Placing Bets from Evidence

In application, the upper and lower bounds of the belief function may not provide intuitive usefulness in presentation to an analyst decision-maker. Typically the belief structure is either converted to a probability distribution (as in the Bayesian approximations) or collapsed to "pignistic" probabilities that allow a decision-maker to place a bet on each proposition given the available evidence [56, 57].

In evidence theory terminology, the creation and combination of belief functions from evidence occurs at the "credal" level [57]. In the credal level, belief may be applied to subsets of the decision space, which is one of the strengths of the evidence theory approach. However, since the elements of the hypothesis state-space are, by definition, mutually exclusive and collectively exhaustive, it is known that exactly one singleton proposition must be true at any given time. Therefore, in order to make decisions based on available evidence, Smets et al. [57] defined a transformation to pignistic probability, which is where the hypotheses are resolved for decision-making. Pignistic probability is effectively a betting probability: the probability that a rational human would assign to an option when required to make a decision given the available evidence. The pignistic probability transform calculates probabilities from a belief function as shown in Eqn. (4.8):

$$\mathrm{BetP}\,(A) = \sum_{X \subseteq \Omega} \frac{|A \cap X|}{|X|} \frac{m(X)}{1 - m(\emptyset)} \tag{4.8}$$

where $|X|$ denotes the number of elements in the set X [57]. This formulation allows $m(\emptyset) \neq 0$ since an open-world assumption is made in the Transferable Belief Model, allowing belief mass to be assigned to the empty set. This transformation can still be used in closed-world applications of Dempster-Shafer theory by ensuring the frame of discernment is collectively exhaustive such that $m(\emptyset) = 0$.

4.1.5 Binary Hypothesis BPAs

From Eqn. (4.6), it can be seen that the computational complexity of the combination of two bpas scales quadratically with the number of propositions in the frame of discernment. The $\mathcal{O}(n^2)$ nature of Dempster's rule means it is computationally preferable to restrict the number of propositions n in Ω. The simplest and most computationally attractive frame of discernment is therefore a binary frame where the two propositions are simply a null and alternate proposition: $\Omega = \{\omega, \neg\omega\}$ where the $\neg$ symbol indicates the negation of ω. Using predicate logic, a complicated frame of discernment can be decomposed into a number of subsets of frames, each addressing smaller portions of information. The important aspect to consider is that the hypotheses must be able to be interrogated through data that is currently available or actionable. The relevant action can then gather evidence to directly interrogate this hypothesis, generating a bpa that represents that particular expert.

Utilizing a binary hypothesis structure with Dempster's conjunctive rule of combination, Eqn. (4.6), allows the combined bpa to be written simply:

$$K_{i,j} = m_i(\omega)m_j(\neg\omega) + m_i(\neg\omega)m_j(\omega) \tag{4.9}$$

$$m_{i\oplus j}(\omega) = \frac{m_i(\omega)m_j(\omega) + m_i(\omega)m_j(\Omega) + m_i(\Omega)m_j(\omega)}{1 - K_{i,j}} \tag{4.10}$$

$$m_{i\oplus j}(\neg\omega) = \frac{m_i(\neg\omega)m_j(\neg\omega) + m_i(\neg\omega)m_j(\Omega) + m_i(\Omega)m_j(\neg\omega)}{1 - K_{i,j}} \tag{4.11}$$

$$m_{i\oplus j}(\Omega) = \frac{m_i(\Omega)m_j(\Omega)}{1 - K_{i,j}} \tag{4.12}$$

This form is a simple operation that can quickly quantify ambiguity in evidence supporting a particular proposition. The usefulness of quantified ambiguity, or ignorance, will be discussed in the following section.

4.2 Ignorance Criterion for Tasking Decisions

Dempster-Shafer and other evidential reasoning theories do not require an expert to report belief in only singleton propositions. Instead, the focal set can contain any subset of the frame of discernment, including the entire frame of discernment itself. However, attributing belief mass to Ω does not increase an analyst's understanding of the situation. Instead, it represents a residual ambiguity, indicating that the expert was unable to attribute that belief to any particular proposition. This admits an ignorance on the part of the expert that is crucial in modeling realistic, uncertain decision-making environments. Similarly, contributing belief to any non-singleton subset of propositions admits some ignorance (e.g. note the indeterminism in the statement: "attribute X belief to either A or B"), since the expert is saying it is unable to further delineate between those propositions based on its available evidence.

When considering potential courses of action, the ideal course leads to a state of perfect knowledge and no residual ambiguity; in other words, all belief is attributed solely to singleton propositions. To support the objectives of decision-makers, it must be possible to interrogate hypotheses with evidence with the goal of confirming or rejecting the hypotheses. This can be alternately formulated as a minimization of ignorance in the hypothesis state-space.

In a similar way that covariance-reduction techniques aim to approach the truth of the estimated state (e.g. orbit) given the available data, a tasking scheme focused on ignorance-reduction will yield the truth-or-falseness of that hypothesis given the available evidence. Importantly, ignorance minimization provides an unbiased method of tasking as it does not favor resolution to any particular proposition in the frame of discernment; it does not prefer tasking that confirms or rejects any specific hypothesis. Instead, it seeks to gather evidence that leads to the least ambiguous knowledge. This formulation is particularly relevant in sensing applications such as the SSA sensor tasking problem; the objective is not to seek

evidence that confirms the nominal state of the space object but rather seek evidence that most efficiently resolves to the true anomaly-state, whether it is nominal or anomalous.

Given a frame of discernment Ω with n mutually exclusive and collectively exhaustive propositions, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$, the ignorance in proposition ω_i based on the evidence in bpa m_j is computed as the difference between the belief mass attributed to ω_i and belief mass attributed to any subset of Ω that contains ω_i:

$$\mathrm{Ig}_{m_j}(\omega_i) = \sum_{\omega_i \subseteq A \subseteq \Omega} m_j(A) - m_j(\omega_i) \tag{4.13}$$

$$= \sum_{\omega_i \subsetneq A \subseteq \Omega} m_j(A) \tag{4.14}$$

$$= \mathrm{Pl}_{m_j}(\omega_i) - \mathrm{Bel}_{m_j}(\omega_i) \tag{4.15}$$

In the case of binary hypotheses, this math simplifies further since the universal-set contains only two elements, i.e. $\Omega = \{\omega_1, \neg\omega_1\}$.

$$\mathrm{Ig}_{m_j}(\omega_1) = \mathrm{Ig}_{m_j}(\neg\omega_1) = m_j(\Omega) \tag{4.16}$$

In this case, the ignorance in frame Ω associated with bpa m_j is simply the belief mass attributed by m_j to the whole frame, since it is the only non-singleton proposition.

4.3 Multi-Sensor Multi-Target Ignorance-Reduction Tasking

The goal in developing this framework is to enable autonomous tasking of a non-homogeneous set of SSA sensors to investigate a hypothesis for a set of space objects. The selected actions will drive the ignorance in this hypothesis toward zero, gathering enough evidence to resolve the hypothesis at the end of the tasking horizon. Note that, while the derivation is given and later applied using an SSA case study, this tasking framework could be applied to any arbitrary hypothesis resolution problem to gather ignorance-minimizing evidence.

The SSA system consists of M sensors and N space objects. At a given time t_k, each sensor $i \in 1, \ldots, M$ can execute an action from a finite set of A allowable actions:

$$\mathcal{S}_k^{(i)} = \left\{ a_1^{(i)}, \ldots, a_A^{(i)} \right\} \tag{4.17}$$

The total set of possible action combinations for all M sensors at time t_k is then given by

$$\mathbb{D}_k = \mathcal{S}_k^{(1)} \times \ldots \times \mathcal{S}_k^{(M)} \tag{4.18}$$

The action taken by sensor i at time t_k is represented by $d_k^{(i)}$. Therefore, the set of actions taken by all M sensors at time t_k is given by the M-tuple:

$$\mathcal{D}_k = \left\{ d_k^{(i)} \right\}_{i=1}^{M} = \left\{ d_k^{(1)}, \ldots, d_k^{(M)} \right\} \in \mathbb{D}_k \tag{4.19}$$

The effect of action $d_k^{(i)}$ on the hypothesis knowledge of space object $j \in 1, \ldots, N$ is determined using Dempster's conjunctive rule of combination. Recall that, due to the associative and commutative properties of Dempster's conjunctive rule, multiple independent pieces of evidence can be combined simply by chaining the combination of their associated belief mass functions.

$$\hat{m}_{k:\mathcal{D}_k}^{(j)} = \bigoplus_{i=1}^{M} \hat{m}_{k:d_k^{(i)}}^{(j)} \quad , \quad d_k^{(i)} \in \mathcal{S}_k^{(i)} \tag{4.20}$$

where $\hat{m}_{k:d_k^{(i)}}^{(j)}$ is the estimated belief-mass function for space object j due to action $d_k^{(i)} \in \mathcal{S}_k^{(i)}$ taken by sensor i at time k. The updated estimated bpa at time t_k for space object j based on decision set $\mathcal{D}_k$ is defined as

$$\hat{m}_{k|k}^{(j)} = m_{k|k-1}^{(j)} \oplus \hat{m}_{k:\mathcal{D}_k}^{(j)} \quad , \quad \mathcal{D}_k \in \mathbb{D}_k \tag{4.21}$$

where $m_{k|k-1}^{(j)}$ is the *a priori* bpa estimate using evidence from time t_{k-1} propagated to

time t_k, and $\hat{m}_{k|k}^{(j)}$ is the *a posteriori* bpa estimate updated with evidence at time t_k. The estimated hypothesis ignorance at time k based on all available evidence from the set of actions $\mathcal{D}_k$ is denoted $\hat{\mathrm{Ig}}_{\hat{m}_{k|k}^{(j)}}(\Omega)$.

4.3.1 Greedy Minimum-Ignorance Optimization

In this general formulation, the goal is to select decisions that minimize ignorance in the hypothesis frame $\Omega = \{\omega_1, \ldots\}$ for each of the N space objects. Therefore, the chosen action should maximally reduce the estimated ignorance over the considered time horizon. Moreover, the chosen set of actions for all sensors at time t_k should maximally reduce the combined estimated ignorance. Since there are multiple objects, this is a multi-objective optimization problem. The proposed solution is to apply a set of weightings, $\mathcal{W} = \{w_1, \ldots, w_N\}, \sum_{j=1}^{N} w_j = 1$, based on relative priorities for each of the N space objects. The optimal decision set at time t_k is then given by Eqn (4.22):

$$\begin{aligned}
\mathcal{D}_k^* &= \arg\min_{\mathcal{D}_k \in \mathbb{D}_k} \sum_{j=1}^{N} w_j \hat{\mathrm{Ig}}_{\hat{m}_{k|k}^{(j)}}(\Omega) \\
&= \arg\min_{\mathcal{D}_k \in \mathbb{D}_k} \sum_{j=1}^{N} w_j \hat{\mathrm{Ig}}\left(\Omega; \hat{m}_{k|k-1}^{(j)} \oplus \hat{m}_{k:\mathcal{D}_k}^{(j)}\right) \\
&= \arg\min_{\mathcal{D}_k \in \mathbb{D}_k} \sum_{j=1}^{N} w_j \hat{\mathrm{Ig}}\left(\Omega; \hat{m}_{k|k-1}^{(j)} \oplus \left(\bigoplus_{i=1}^{M} \hat{m}_{k:d_k^{(i)}}^{(j)}\right)\right)
\end{aligned} \tag{4.22}$$

This formulation selects a decision set to minimize weighted ignorance at the current time step, otherwise known as greedy optimization.

4.3.2 Receding Horizon Minimum-Ignorance Optimization

A related approach is to optimize the set of decisions over a fixed time span of H steps, using a receding horizon approach to minimize ignorance at time t_{k+H}. It is well-known that Dempster's conjunctive rule of combination is not well-suited to propagating belief structures through time since the evidence combined must be independent [80]. This approach

propagates the state estimates through time and computes the belief structure at t_{k+H} based on the propagated state knowledge, avoiding the need to combine belief structures induced by dependent pieces of evidence.

Define the set of possible actions for sensor i from time t_k to t_{k+H} as:

$$\mathcal{S}^{(i)}_{[k,k+H]} = \mathcal{S}^{(i)}_{k} \times \ldots \times \mathcal{S}^{(i)}_{k+H} \tag{4.23}$$

and define the set of possible action combinations for all M sensors from t_k to t_{k+H} as:

$$\mathbb{D}_{[k,k+H]} = \mathcal{S}^{(1)}_{[k,k+H]} \times \ldots \times \mathcal{S}^{(M)}_{[k,k+H]} \tag{4.24}$$

Now the set of decision sets taken by all M sensors from t_k to t_{k+H} is given by:

$$\mathcal{D}_{[k,k+H]} = \{\mathcal{D}_g\}_{g=k}^{k+H} \in \mathbb{D}_{[k,k+H]} = \{\mathcal{D}_k, \ldots, \mathcal{D}_{k+H}\} \quad , \quad H \in \mathbb{Z}^+ \tag{4.25}$$

The optimal set of decision sets between times t_k and t_{k+H} is then given by Eqn. (4.26).

$$\mathcal{D}^*_{[k,k+H]} = \arg \min_{\mathcal{D}_{[k,k+H]}} \sum_{j=1}^{N} w_j \hat{\mathrm{Ig}}_{\hat{m}^{(j)}_{k+H|k+H}}(\Omega) \tag{4.26}$$

The key difference between Eqns. (4.22) and (4.26) is the estimated bpa $\hat{m}^{(j)}_{k+H|k+H}$, which is evaluated based upon the estimated state at t_{k+H}, which depends upon all actions taken from t_k through and including t_{k+H}. This is reasonable since the goal is to minimize ignorance at the horizon time t_{k+H}, not the intermediate steps, so the only relevant bpa for each space object is the one at t_{k+H}. Naturally, if $t_{k+H} = t_k$ so that the horizon length is zero, the receding horizon approach in Eqn. (4.26) collapses to the greedy approach in Eqn. (4.22).

4.4 Application to SSA Sensor Tasking

Having derived the ignorance-reduction tasking scheme, this section demonstrates its practical application using the SSA custody problem as a case-study. In SSA, an assigned responsibility to track a particular space object is referred to as custody. The sensors given custody of a space object must periodically observe the object to verify that its state has not changed. Missed-detection of a space object could be due to any number of causes, including cloud cover, background sky brightness, proximity to a bright sky object, poor space object state predictions, or spacecraft maneuver, to name a few. The latter cause, a controlled maneuver by the space object, is one of many potential "dynamics anomalies" where the space object state is not adequately predicted by the assumed dynamics. A maneuver, in particular, requires timely response by decision-makers to reacquire the space object to maintain custody, update state estimates, or reconstruct the maneuver, which makes dynamics anomalies a primary concern. In contrast, missed-detection from other causes (e.g. cloud cover) does not provide any definitive evidence of anomaly and does not necessarily constitute re-tasking sensors for re-acquisition.

This section develops the evidence-based framework for this problem and applies the ignorance-reduction criterion using SSA sensors to resolve the custody and anomaly hypotheses. This is intended to be an illustrative application for the ignorance-reduction tasking technique, demonstrating methods for developing belief functions and evaluating sensor actions. In a real-world application, other hypotheses should be considered, including detection of clutter and generation or decay of other objects or debris, but for illustration the number of hypotheses is limited. Continuing research [82] focuses on further developing evidential reasoning sensor tasking methods in increasingly-realistic scenarios, including addressing computational complexity concerns with many more hypotheses.

4.4.1 Custody and Anomaly Discrimination Framework

Given M sensors tasked with keeping custody of N space objects, the sensors must verify the orbit state of each object. Each sensor observation has an associated measurement noise that contributes to state uncertainty. The state uncertainty then grows as the orbit state is propagated and, as the state uncertainty exceeds the limits of the sensor's field of view, the probability of successful detection decreases. This problem is compounded when considering the possibility of maneuvers or other anomalies between observations. Therefore, from an uncertainty standpoint, the ideal tasking would be constant observation of each space object, which is impractical considering observation constraints imposed by orbital mechanics and the vast number of space objects (typically, $N >> M$). Additionally, observation conditions (e.g. sky brightness, local weather) can lead to unsuccessful observation attempts.

When a sensor has been tasked to a particular space object, the result of that tasking can be described as either a successful detection, $\mathcal{D}$, or missed-detection, $\mathcal{M}$. Additionally, at any particular time the anomaly-state of the space object can be described as either nominal, $\mathcal{N}$, or anomalous $\mathcal{A}$. The cartesian product of these two binary hypothesis spaces represents the possible results of a tasking decision, leading to the following frames of discernment:

$$\Omega_T = \{\mathcal{D}\,,\,\mathcal{M}\}\,, \quad \Omega_A = \{\mathcal{N}\,,\,\mathcal{A}\} \tag{4.27}$$

$$\Omega = \{\mathcal{D}\cap\mathcal{N}\,,\,\mathcal{D}\cap\mathcal{A}\,,\,\mathcal{M}\cap\mathcal{N}\,,\,\mathcal{M}\cap\mathcal{A}\} \tag{4.28}$$

where $\mathcal{D}\cap\mathcal{N}$ means the space object has been detected and its state is nominal, $\mathcal{D}\cap\mathcal{A}$ means detected and anomalous, $\mathcal{M}\cap\mathcal{N}$ means missed-detection and nominal, and $\mathcal{M}\cap\mathcal{A}$ means missed-detection and anomalous. In each binary frame of Eqn. (4.27), the elements are mutually exclusive and collectively exhaustive, so the combined frame in Eqn. (4.28) also satisfies these properties.

Given this combined hypothesis space, the proposed goal of the sensor network is

to ensure detection ($\mathcal{D}$) of each space object so that the sensor network maintains custody regardless of the anomaly state. Note that the use of predicate logic confirms that $\Omega \supset \{\mathcal{D} \cap \mathcal{N}, \mathcal{D} \cap \mathcal{A}\} = \mathcal{D}$; "detection-and-nominal or detection-and-anomalous" is equivalent to "detection" because $\mathcal{N}$ and $\mathcal{A}$ are mutually exclusive and collectively exhaustive. Similarly, $\Omega \supset \{\mathcal{D} \cap \mathcal{N}, \mathcal{M} \cap \mathcal{N}\} = \mathcal{N}$, and $\Omega \supset \{\mathcal{D} \cap \mathcal{A}, \mathcal{M} \cap \mathcal{A}\} = \mathcal{A}$. Also, note that since Ω_T and Ω_A are both binary, $\mathrm{Ig}(\mathcal{D}) = \mathrm{Ig}(\mathcal{M})$ and $\mathrm{Ig}(\mathcal{N}) = \mathrm{Ig}(\mathcal{A})$.

The proposed goal can be stated in terms of evidence theory quantities as minimizing the ignorance associated with the hypothesis $\{\mathcal{D} \cap \mathcal{N}, \mathcal{M} \cap \mathcal{N}\} = \mathcal{N}$; in other words, the goal is to minimize the ignorance in the anomaly state of the space object. This is an intuitive interpretation of the goal of custody-based tasking: it cannot be controlled whether the space object is nominal or anomalous and the sensor tasking should not favor either result, but it is important to be able to confirm detection in either case. While this seems to emphasize only the Ω_A hypothesis space and ignore the Ω_T hypothesis space, its influence on the resolution of the ignorance in Ω_A is made clear in the following section on incorporating SSA information as Dempster-Shafer belief functions.

4.4.2 SSA Dempster-Shafer Experts

To formulate SSA Dempster-Shafer experts, the SSA sensor data available must be encoded as bpas, carefully considering the information that each sensor provides for each hypothesis; in particular, does the evidence provide direct support for a proposition (belief), or does it simply not directly disprove a proposition (plausibility). In order to arrive at accurate hypothesis resolution, the bpas must be rigorously formed and appropriately combined using evidence combination rules such as Dempster's conjunctive rule.

In this work, most of the available evidence only provides levels of plausibility on certain propositions. Recall that bpa m can also be represented as a belief function $\mathrm{Bel}(\cdot)$, or plausibility function $\mathrm{Pl}(\cdot)$, and each one can be converted into the other two. The conversions between each representation are linear functions, so they can all be represented

using linear algebra [76]. The determined plausibility from each sensor is converted to a bpa, committing the least belief mass to each relevant proposition [83].

For instance, given the hypothesis space in Eqn. (4.28), plausibility in the hypothesis $\mathcal{D} \cap \mathcal{N}$ is computed as the sum of component belief masses: $\mathrm{Pl}(\mathcal{D} \cap \mathcal{N}) = m(\mathcal{D}) + m(\mathcal{N}) + m(\mathcal{D} \cap \mathcal{N})$. The linear combinations of belief masses to compute plausibilities can be represented using a transformation matrix between a vector of belief masses $\mathbf{m}$ and a vector of plausibilities $\mathbf{p}$.

$$\mathbf{p} = \mathbf{T}_{m2p}\mathbf{m} \tag{4.29}$$

$$\begin{bmatrix} \mathrm{Pl}(\mathcal{D} \cap \mathcal{N}) \\ \mathrm{Pl}(\mathcal{D} \cap \mathcal{A}) \\ \mathrm{Pl}(\mathcal{M} \cap \mathcal{N}) \\ \mathrm{Pl}(\mathcal{M} \cap \mathcal{A}) \\ \mathrm{Pl}((\mathcal{D} \cap \mathcal{N}) \cup (\mathcal{D} \cap \mathcal{A})) \\ \mathrm{Pl}((\mathcal{M} \cap \mathcal{N}) \cup (\mathcal{M} \cap \mathcal{A})) \\ \mathrm{Pl}((\mathcal{D} \cap \mathcal{N}) \cup (\mathcal{M} \cap \mathcal{N})) \\ \mathrm{Pl}((\mathcal{D} \cap \mathcal{A}) \cup (\mathcal{M} \cap \mathcal{A})) \\ \mathrm{Pl}(\Omega) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} m(\mathcal{D} \cap \mathcal{N}) \\ m(\mathcal{D} \cap \mathcal{A}) \\ m(\mathcal{M} \cap \mathcal{N}) \\ m(\mathcal{M} \cap \mathcal{A}) \\ m(\mathcal{D}) \\ m(\mathcal{M}) \\ m(\mathcal{N}) \\ m(\mathcal{A}) \\ m(\Omega) \end{bmatrix} \tag{4.30}$$

Since the sensors provide evidence in the form of plausibilities, the inverse operation finds the belief mass assignments: $\mathbf{m} = \mathbf{T}_{m2p}^{-1}\mathbf{p}$. The transformation matrix is an identity matrix augmented with non-zero terms in the off-diagonal, and since each mass-to-plausibility conversion is unique, the matrix is full rank. Therefore, the matrix is invertible, and belief masses can be found from plausibilities.

Each sensor's evidence must be carefully considered to create valid plausibility functions that lead to valid belief mass functions. For instance, if the probability of detection (based on sky brightness and cloud cover measurements) is one, that does not guarantee a successful detection (the space object may have maneuvered); however, if the radiometric

probability of detection is zero, that does guarantee a missed detection but does not resolve the anomaly hypothesis. Therefore, a Dempster-Shafer expert based on the probability of detection can only discount plausibility in the detection hypotheses. Similarly, the uncertainty and reachability probabilities can be used to determine the plausibility of detecting the space object in its nominal or anomalous state.

4.4.3 Tasking

At each time step, each sensor has a set of potential actions as denoted by Eqn. (4.17). In this example application, the actions include nominal observations or searching observations, depending on the estimated anomaly state of the space object. Nominal observation tasks against the space object's uncertainty volume, while the anomalous object search investigates the space object's reachable volume. The objective in the tasking phase is to execute the actions that minimize the estimated anomaly-state ignorance after a proposed tasking decision. The contributing factors to consider are the growth of uncertainty and the observation conditions. If the orbit state uncertainty is allowed to grow unbounded,missed detection is more likely, so frequently observing the object to reduce covariance is useful in maintaining custody. If the covariance is too high, a missed-detection may be reasonably attributed to the nominal cause of state uncertainty-growth. If observation conditions for a particular sensor are poor at the relevant sky position for a space object, tasking that sensor to observe the space object does not resolve custody; a missed-detection may be reasonably attributed to poor detection probability.

In terms of the frame of discernment in Eqn. (4.28), the compounded effect of state uncertainty and observation conditions on detection in a nominal scenario is captured using

the following plausibility function:

$$\begin{aligned}
\mathrm{Pl}_{m_N}(\mathcal{D} \cap \mathcal{N}) &= \mathbb{P}_{unc}(\mathbf{P}, f, \gamma, \epsilon) \cdot \mathbb{P}_{det}(\tau_{atm}, I_{sky}, \gamma, \epsilon) \\
\mathrm{Pl}_{m_N}(\mathcal{M} \cap \mathcal{N}) &= 1 - \mathbb{P}_{unc}(\mathbf{P}, f, \gamma, \epsilon) \cdot \mathbb{P}_{det}(\tau_{atm}, I_{sky}, \gamma, \epsilon) \\
\mathrm{Pl}_{m_N}((\mathcal{D} \cap \mathcal{N}) \cup (\mathcal{D} \cap \mathcal{A})) &= \mathbb{P}_{unc}(\mathbf{P}, f, \gamma, \epsilon) \cdot \mathbb{P}_{det}(\tau_{atm}, I_{sky}, \gamma, \epsilon) \\
\mathrm{Pl}_{m_N}((\mathcal{M} \cap \mathcal{N}) \cup (\mathcal{M} \cap \mathcal{A})) &= 1 - \mathbb{P}_{unc}(\mathbf{P}, f, \gamma, \epsilon) \cdot \mathbb{P}_{det}(\tau_{atm}, I_{sky}, \gamma, \epsilon)
\end{aligned} \tag{4.31}$$

where $\mathbf{P}$ is the state covariance, f is the sensor field of view, τ_{atm} is the atmospheric transmittance, I_{sky} is the background sky irradiance, and γ, ϵ are the sensor's pointing in azimuth and elevation angles. The specific models for the probabilities in the above equations are described in detail in Section 4.5, using a radiometric model for $\mathbb{P}_{det}$ [15] and uncertainty and reachability results for $\mathbb{P}_{unc}$ and $\mathbb{P}_{reach}$ [84, 85]. The propositions that are not explicitly enumerated in Eqn. (4.31) are assigned $\mathrm{Pl}(\cdot) = 1$ since the available evidence cannot directly discount these propositions. The plausibility function can be converted to a bpa using known one-to-one linear mappings [76]. If observation conditions are poor, a missed-detection does not confirm or refute an anomalous state so the belief mass is assigned as ignorance. Similarly, if the state uncertainty volume exceeds the sensor field of view, a missed-detection does not confirm an anomalous state even in perfect observation conditions; therefore, the remainder of the belief mass is assigned as ignorance.

A similar analysis can be performed considering maneuvers by examining the reachability volume of state-space [84, 85]. The formulation is nearly identical, but this time the belief is applied to the "detected-anomalous" hypothesis:

$$\begin{aligned}
\mathrm{Pl}_{m_A}(\mathcal{D} \cap \mathcal{A}) &= \mathbb{P}_{reach}(P, u_{max}, f, \gamma, \epsilon) \cdot \mathbb{P}_{det}(\tau_{atm}, I_{sky}, \gamma, \epsilon) \\
\mathrm{Pl}_{m_A}(\mathcal{M} \cap \mathcal{A}) &= 1 - \mathbb{P}_{reach}(P, u_{max}, f, \gamma, \epsilon) \cdot \mathbb{P}_{det}(\tau_{atm}, I_{sky}, \gamma, \epsilon) \\
\mathrm{Pl}_{m_A}((\mathcal{D} \cap \mathcal{N}) \cup (\mathcal{D} \cap \mathcal{A})) &= \mathbb{P}_{reach}(P, u_{max}, f, \gamma, \epsilon) \cdot \mathbb{P}_{det}(\tau_{atm}, I_{sky}, \gamma, \epsilon) \\
\mathrm{Pl}_{m_A}((\mathcal{M} \cap \mathcal{N}) \cup (\mathcal{M} \cap \mathcal{A})) &= 1 - \mathbb{P}_{reach}(P, u_{max}, f, \gamma, \epsilon) \cdot \mathbb{P}_{det}(\tau_{atm}, I_{sky}, \gamma, \epsilon)
\end{aligned} \tag{4.32}$$

where u_{max} is the maximum control authority of the space object, a quantity that is assumed known in this work (e.g. obtained from operators) but may also require estimation. The remainder of the variables are as defined in Eqn. (4.31).

Both plausibility functions are converted to bpas, m_N and m_A respectively. Fusing these two bpas estimates the resolution of the hypotheses in Eqn. (4.28) based on a candidate tasking option. The ignorance in the anomaly state as a result of this tasking can then be evaluated.

$$\mathrm{Ig}\left(\mathcal{N}; m_N \oplus m_A\right) = \mathrm{Ig}\left(\{\mathcal{D} \cap \mathcal{N}, \mathcal{M} \cap \mathcal{N}\}; m_N \oplus m_A\right) \tag{4.33}$$

The optimal set of tasking decisions is then computed following Eqn. (4.26) to minimize the estimated weighted total ignorance in space object anomaly state following the candidate tasking, subject to weightings induced by relative priority in the anomaly states of the objects.

The look-angles (γ, ϵ) for each tasking option are determined by the estimated anomaly state of the space object. In the event that a space object is believed to be nominal, the mean state estimate is used to compute the look-angles. However, if the space object is thought to be anomalous, the look-angles are modified to progressively seek through the reachable space until it is once again detected. The process of flagging a space object as anomalous is discussed in the next section, followed by description of the custody reacquisition process.

4.4.4 Anomaly Discrimination

Once the optimal tasking has been determined, this tasking is implemented and the object is either detected or missed. In the event that a space object is detected, the proposition $\mathcal{D} \subset \Omega_T$ is confirmed and there is direct evidence that can be applied to interrogating the anomaly hypothesis. Numerous methods exist for the detection of anomalies from full state estimates. One such method is the Mahalanobis distance: a statistical method

that computes a multi-dimensional standard deviation between two state distributions [63]. This is governed by Eqn. (4.34):

$$d_{MD}\left(\mathbf{x}_A, \mathbf{x}_B, \mathbf{P}\right) = \sqrt{\left(\mathbf{x}_A - \mathbf{x}_B\right)^{\mathrm{T}} \mathbf{P}^{-1} \left(\mathbf{x}_A - \mathbf{x}_B\right)} \tag{4.34}$$

where $\mathbf{x}_A$ and $\mathbf{x}_B$ are the mean states of the distributions and $\mathbf{P}$ is the distance matrix [63]. Jaunzemis et al. [32] applied this metric in a binary hypothesis test using the following formulations for the null and alternative propositions ($\mathcal{H}_0$ and $\mathcal{H}_1$ respectively). In the event that Gaussian boundary conditions are assumed, the Mahalanobis distance distributions for each hypothesis are generated using Eqns. (4.35)-(4.36):

$$d_{MD,\mathcal{H}_0}\left(\boldsymbol{\mathcal{X}}_{k-1}, \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}\right) = \sqrt{\left(\boldsymbol{\mathcal{X}}_{k-1} - \hat{\mathbf{x}}_{k|k-1}\right)^{\mathrm{T}} \mathbf{P}_{k|k-1}^{-1} \left(\boldsymbol{\mathcal{X}}_{k-1} - \hat{\mathbf{x}}_{k|k-1}\right)} \tag{4.35}$$

$$d_{MD,\mathcal{H}_1}\left(\boldsymbol{\mathcal{X}}_{k}, \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k} + \mathbf{P}_{k|k-1}\right) = \sqrt{\left(\boldsymbol{\mathcal{X}}_{k} - \hat{\mathbf{x}}_{k|k-1}\right)^{\mathrm{T}} \left(\mathbf{P}_{k|k} + \mathbf{P}_{k|k-1}\right)^{-1} \left(\boldsymbol{\mathcal{X}}_{k} - \hat{\mathbf{x}}_{k|k-1}\right)} \tag{4.36}$$

where $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ are the a priori state and covariance estimates propagated to time t_k, $\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ are the a posteriori state and covariance estimates incorporating the successful detection at time t_k, and random vectors are drawn from the Gaussian state distributions as $\boldsymbol{\mathcal{X}}_{k-1} \sim \mathcal{N}\left(\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}\right)$, $\boldsymbol{\mathcal{X}}_{k} \sim \mathcal{N}\left(\hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}\right)$ [32]. The use of the combined covariance as the distance matrix in the alternative hypothesis is a modified form of the Mahalnobis distance to describe the distance between maneuvered and quiescent (non-maneuvered) distributions [30].

An anomaly bpa can be constructed to resolve the anomaly state of the space object:

$$\begin{aligned} m_{A,k|k}(\mathcal{D} \cap \mathcal{A}) &= \mathbb{P}_{A,MD}\left(d_{MD,\mathcal{H}_0}, d_{MD,\mathcal{H}_1}\right) \\ m_{A,k|k}(\mathcal{D} \cap \mathcal{N}) &= 1 - \mathbb{P}_{A,MD}\left(d_{MD,\mathcal{H}_0}, d_{MD,\mathcal{H}_1}\right) \end{aligned} \tag{4.37}$$

using belief mass directly since direct evidence is available through detection.

Alternate distance metrics may be employed, such as the control-distance method [32]. It should be noted that, since the probabilities of anomaly for Mahanalobis distance and control distance operate on the same information, and they are not independent evidence sources. Therefore if both are used, they must be combined into one bpa to be used in evaluating the resolution of the anomaly hypothesis, as shown in Eqn. (4.37). In this paper, only the Mahalanobis distance is implemented for anomaly detection. For more details on computing the anomaly probability, refer to Jaunzemis et al. [32]. Once this final piece of evidence has been gathered and fused with the existing evidence from the observation conditions and state uncertainty, a decision must be made on the anomaly state of the object. In Dempster-Shafer theory, the pignistic belief induces a probability function that is used to inform decisions from belief structures [57]. Therefore, the pignistic belief in the anomaly hypothesis is computed, and if it exceeds a threshold, the space object is flagged as anomalous.

In the event of a missed detection, the missed-detection hypothesis $\mathcal{M}$ is confirmed, but there is no further evidence to apply. The existing evidence is assessed to determine anomaly state using pignistic belief, and once again if a threshold is exceeded, the space object is flagged as anomalous.

4.4.5 Custody Reacquisition

If custody has been lost for any reason, a procedure for recovering the space object must be implemented. If the space object was missed due to poor observation conditions, there is not necessarily a need to believe an anomaly has occurred, so the preferred response may be to simply wait until conditions change to attempt another observation. If, however, the space object should have been detected based on the available SSA sensor data, the network should flag the space object as anomalous and react accordingly. In Dempster-Shafer terms, an anomaly is declared when the pignistic belief in anomaly exceeds some threshold.

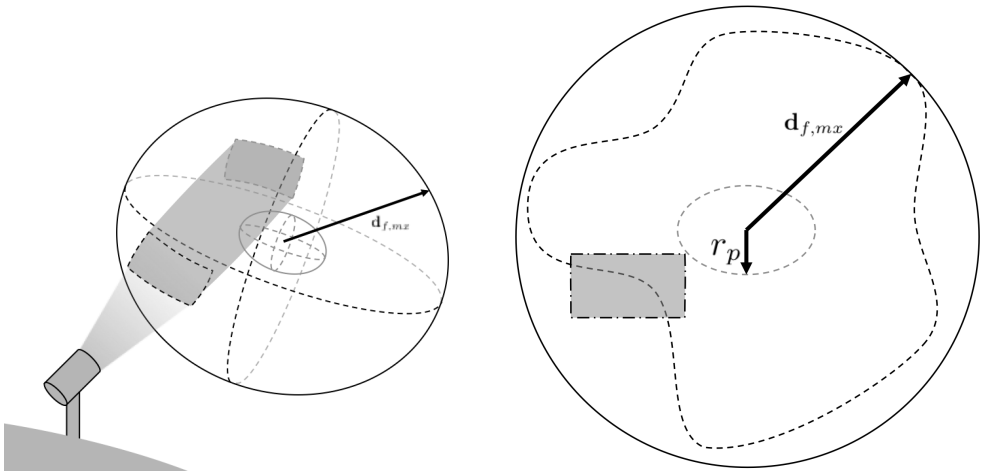
After this point, the tasking algorithm must change to re-acquire a potentially maneuvered object. If the observation conditions for the nominal location were exceedingly good but the observation was still missed, the anomaly hypothesis will be resolved to indicate an anomaly has occurred. This means the anomaly ignorance will already be minimized, so ignorance reduction will not help in re-acquiring the space object. A solution employed here is to enter a search mode that assigns a sensor to the anomalous space object based on the highest likelihood for re-acquisition. This entails interrogating the largest possible volume of reachable space (i.e. use the largest projected field of view) while ensuring the attempted observation resolves the custody hypothesis to "detection." Once the space object is re-acquired and custody is confirmed, the anomaly can be characterized and the estimator updated to ensure follow-on observations account for the anomaly.

4.5 SSA Sensor Evidence

In order to apply Dempster-Shafer theory, available SSA sensors must be cast as BPAs, contributing belief mass to the available hypotheses based on that sensor's data. In this paper, the primary sensors are electro-optical (EO) sensors, such as telescopes. The radiometric model in Eqn. (4.38) determines the probability of detection of an RSO by an EO sensor [15].

$$\mathbb{P}_{det}\left(\tau_{atm}, I_{sky}, \gamma, \epsilon\right) = \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{SNR_{alg}\sigma_n - \mu_{so}}{\sqrt{2}\sigma_{so}}\right)\right] \tag{4.38}$$

The algorithm-required SNR (SNR_{alg}) is a sensor-specific quantity dependent upon the algorithm used for RSO detection. The mean and standard deviation of the RSO signal received by the sensor (μ_{SO} and σ_{SO} respectively) and the standard deviation of background noise sources (σ_{SO}) can be estimated from data regarding the RSO state and observation conditions [15]. In this paper, SNR_{alg} is assumed to be given for each sensor and remaining quantities are estimated using auxiliary sensors, namely atmospheric transmittance (τ_{atm})



(a) 3D visualization (b) Projection onto image plane

Figure 4.1: Uncertainty and reachability volumes interrogated by field of view

and background sky irradiance (I_{sky}).

4.5.1 Uncertainty and Reachability

A challenge posed by non-deterministic state measurements is the growth of uncertainty. After propagation, uncertainty in the estimated space object state alone can cause a missed observation because the uncertainty exceeds the sensor field of view. The probability of the space object being in the sensor field of view given its uncertainty $\mathbb{P}_{unc}(\mathbf{P}, f, \rho, \gamma, \epsilon)$ can be computed given the state estimate and covariance, the sensor look-angles (γ, ϵ), the slant-range to the space object ρ, and the field of view of the sensor f. The probability is computed as the intersection of the sensor field of view with the uncertainty volume. In the case of Gaussian uncertainties, the uncertainty volume is an ellipsoid, so the probability density contained within the intersection of this ellipsoid with the field of view is the probability of detection due to uncertainty.

Related to uncertainty is computing the intersection of the sensor field of view with the reachable volume. This computation uses all the same values and adds the maximum control authority (u_{max}) to determine the reachable volume, and the probability of the space object being in the field of view given its reachable volume is denoted $\mathbb{P}_{reach}(\mathbf{P}, u_{max}, f, \rho, \gamma, \epsilon)$.

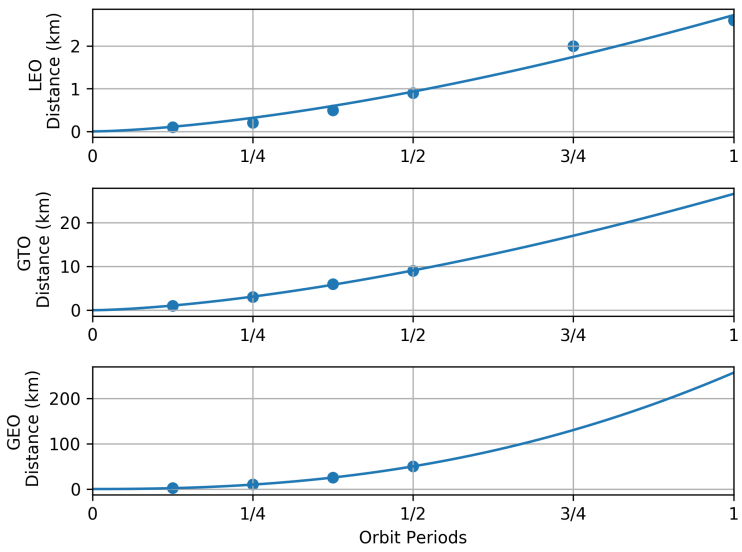


Figure 4.2: Reachability trendlines fit to reference data [84, 85]

To compute the reachable distance, various optimal control methods may be employed. For instance, Holzinger et al. perform this calculation under low-thrust assumptions [84], and Brew and Holzinger extended this work using continuation methods [85]. Since reachability is not the focus of this work, the results from these papers were simply fit to a polynomial of the form $d_{max} = at^b$, where t is the time since last successful observation and a and b are the trend-fitting parameters. A trend was developed for each the LEO, GTO, and GEO regimes, with trends shown in Fig. 4.2. To ensure applicability of this approximation, the same maximum control and initial uncertainties are used in this paper as were used in the reference studies.

4.5.2 Cloud Detection

One potential nominal (non-anomalous) cause for missed detection is occultation by clouds. This can be represented in the Coder et al. radiometric model through the use of atmospheric transmittance (τ_{atm}).

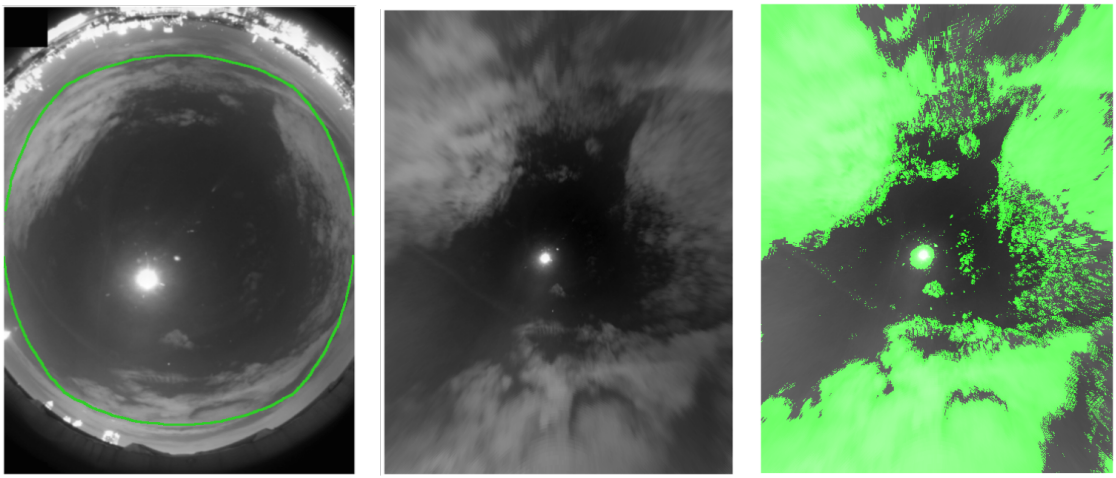
In order to estimate cloud cover, a hybrid thresholding and optical-flow algorithm was written for cloud detection using hemispherical All-Sky camera imagery. The thresholding

captures bright pixels attributed to light reflecting from the clouds back to the All-Sky Sensor, while the optical-flow algorithm is able to capture the more transient, whisp-like layers of cloud cover. The Horn-Schunk optical flow algorithm operates by assuming brightness constancy and relatively short motion of objects in the frame, both valid for night-time cloud detection with relatively short, constant-length exposures. This optical flow algorithm formulates the energy function in Eq. (4.39) subject to a candidate flow field (u, v), seeking to minimize energy by modifying the flow field through a gradient descent [86]. Pixel-wise gradients (I_x and I_y) are computed using central differencing, and frame-to-frame derivatives (I_t) are computed using backward differencing. After a user-defined convergence criterion is met, the magnitudes of the flow field (u, v) can be evaluated to determine which pixels contain cloud and which contain empty sky.

$$E(x,y,t) = \int\int ([I_x(x,y,t)u(x,y,t) + I_y(x,y,t)v(x,y,t) + I_t(x,y,t)]^2 + \|\nabla u(x,y,t)\|^2 + \|\nabla v(x,y,t)\|^2)\mathrm{d}x\mathrm{d}y \tag{4.39}$$

The optical flow algorithm excels at detecting faint, whispy clouds, but does not detect the flat, bright areas of heavy cloud with either pixel-wise or frame-wise derivatives. Therefore, in this hybrid implementation, if a particular pixel exceeds either of the brightness or flow magnitude thresholds, that pixel is deemed to be cloud, constructing a cloud-cover mask from the image.

Figure 4.3 shows a sample result of this implementation. For practical implementation, the barrel distortion of the wide field-of-view sensor, highlighted in these sample images, needs to be corrected. The All-Sky camera used exhibits well-known *f-theta* barrel distortion, which allows for simple conversions between distorted (fish-eye) and undistorted images. In the undistorted image, straight-line motion (such as a cloud moving across the sky) is rendered correctly in a straight line from frame to frame, which is important for correct computation of the flow magnitude using optical flow. This necessarily excludes



(a) Distorted (b) Undistorted (c) Cloud cover overlay

Figure 4.3: All-Sky cloud detection

the corners of the distorted image, but this is acceptable in this application since this corresponds to low-elevation sky positions which are not candidates for tasking.

Since the resultant cloud cover mask is binary, it is blurred by convolving with a 5×5 Gaussian kernel to produce smoother transitions from cloud to clear sky. The atmospheric transmittance τ_{atm} is then estimated given the azimuth-elevation pair (γ, ϵ) for the target sky position.

4.5.3 Sky Brightness

Another potential nominal (non-anomalous) cause for missed detection is background sky brightness. In the Coder et al. radiometric model [15], this is computed as a photon flux due to the sky background irradiance and considered as a noise source when computing SNR and detection probability. The sky irradiance can be measured using a sky brightness monitor. In this work, the sky brightness monitor measures the irradiance at the zenith (directly overhead) in units of magnitudes per square arc-second. Garstang provides a unit conversion useful in sky brightness computations as seen in Eqn. (4.40):

$$B = 34.08 \exp\left(20.7233 - 0.92104V\right) \tag{4.40}$$

where B is the sky brightness in nanoLamberts and V is the sky brightness in magnitudes per square arc-second [87]. The sky brightness at a target elevation can then be computed from the sky brightness at zenith using Eqns. (4.41)-(4.42):

$$d_{opt} = \left(1 - 0.96 \sin^2 d_{zen}\right)^{-0.5} \tag{4.41}$$

$$B(d_{zen}) = B_{zen} 10^{-0.4k(d_{opt}-1)} d_{opt} \tag{4.42}$$

where d_{zen} is the angular distance from zenith to the target object (i.e. the complement of elevation), k is the extinction coefficient in units of magnitudes per air mass, and d_{opt} is the optical pathlength along the line of sight in units of air masses [88]. The sky brightness at the target elevation, converted to units of magnitudes per square arc-second using Eqn. (4.40), is used as the I_{sky} variable in the Coder et al. radiometric model, computing the photon flux due to background sky brightness ($q_{p,sky}$, in electrons per second per pixel) as:

$$L_{sky} = \Phi_0 10^{-0.4 I_{sky}} \left(\frac{180}{\pi}\right)^2 3600^2 \tag{4.43}$$

$$q_{p,sky} = \frac{L_{sky} \tau_{opt} \tau_s \pi \eta_{qe} p^2}{1 + 4N^2} \tag{4.44}$$

where Φ_0 is the photon flux density of a zeroth-magnitude object in photons per second per square meter, τ_{opt} is the transmittance due to the primary optics, τ_s is transmittance due to secondary optics, η_{qe} is the quantum efficiency, p is the pixel size in meters, and N is the binning of the sensor.

4.6 Simulation Results

Using the anomaly and custody problem described above, a multi-sensor, multi-target simulation is presented to evaluate the validity and applicability of the theoretical results to realistic SSA scenarios. In this simulation, three electro-optical sensors are tasked with maintaining custody of five space objects: three in Geostationary Earth orbits (GEO) and

two in Geostationary transfer orbits (GTO).

The sensors are modeled after existing or under-development SSA sensors, distributed throughout the western hemisphere with location and optical parameters listed in Table 4.1. The Georgia Tech (GT) sensor is based on the Georgia Tech Space Object Research Telescope (GT-SORT), a Raven-class telescope located in downtown Atlanta. GT-SORT features half-meter optics (Officina Stellare Pro RC 500 with focal reducer for $f/6$ configuration) and a 6-megapixel CMOS sensor (Point Grey Grasshopper3) for high-resolution, small field-of-view (approximately 0.23×0.18 degrees), high frame-rate (up to 20 Hz) SSA imagery [89]. The Utah State University (USU) sensor is based on the USU Space Situational Awareness Telescope for Astrodynamics Research (USU-STAR), located at the Bear Lake Observatory in Utah. USU-STAR is a Raven-class telescope with 10-inch optics (AG Optical Systems Imaging Dall-Kirkham astrograph) and a 4K CCD sensor (Finger Lakes Imaging Proline PL 16801), resulting in a larger field-of-view (approximately 1.25×1.25 degrees). The third sensor is a fictitious one-meter Raven-class optical sensor placed at the Air Force Maui Optical and Supercomputing site (AMOS) on Haleakala, Maui, with parameters similar to USU-STAR (e.g. field of view approximately 1.17×1.17 degrees). The quantum efficiency, dark current, and optics transmittance for all three sensors are approximated using known values from GT-SORT. The required signal-to-noise ratio (SNR) for the detection algorithm is set at a conservative value of 10 for all three sensors. The nominal atmospheric transmittance and sky brightness are worst for GT-SORT, due to its location in downtown Atlanta, and best for the AMOS sensor since it is located on top of Haleakala.

The space objects are initialized from two-line elements (TLEs) by generating TLEs for 3 GEO space objects and 2 GTO space objects. The orbit parameters for these simulated objects are shown in Table 4.2. The orbit shape parameters, semi-major axis and eccentricity, are selected using representative space objects from each class, GEO or GTO, using the publicly available space object catalog published by *space-track.org*. The orbit

Table 4.1: Simulated sensor locations and optical parameters

	1	**2**	**3**
City	Atlanta, GA	Garden City, UT	Maui, HI
Abbreviation	GT	USU	AMOS
Latitude (°N)	33.7756	41.93	20.7083
Longitude (°W)	84.3963	111.42	156.2571
Altitude (m)	200	1981	3052
Diameter (m)	0.5	0.254	1
Focal Length (m)	3	1.7018	2
Resolution (h × v)	2736 × 2192	4096 × 4096	4096 × 4096
Pixel Size (μm)	4.54	9	10
Quantum Efficiency, η_q (%)	74	74	74
Dark Current, q_d (e^-/s)	10.88	10.88	10.88
Optics Transmittance, $\tau_p\tau_s$ (%)	76	76	76
Minimum Elevation Req. (°)	20	20	20
SNR Req. for Detection	10	10	10
Nom. Sky Irrad., I_{sky} $\left(\frac{mag}{arcsec^2}\right)$	16.5	18	20
Nom. Atm. Trans., τ_{atm} (%)	50	70	70

orientation parameters for the GEO objects are chosen from random uniform distributions on the following ranges: inclination between zero- and one-degrees, and all others (argument of perigee, right ascension of ascending node, and mean anomaly) between zero- and 360-degrees. Note that the only constraint enforced in GEO orbit orientation is that the longitude must be between between 90- and 130-degrees West to ensure visibility by the simulated sensors. Similarly, orbit orientation parameters for the GTO objects are chosen from random uniform distributions on the following ranges: inclination ranges between zero- and five-degrees, and all others (argument of perigee, right ascension of ascending node, and mean anomaly) between zero- and 360-degrees. The generated TLEs are used to initialize unscented Kalman filters (UKFs) to track the space object states and incorporate measurement updates. The estimator utilizes a 2×2 geopotential model for orbit propagation dynamics, primarily to capture the perturbing effects of J_2 on the GTO space objects. Additionally, truth-states are generated for each test case using the TLE initial conditions and a 6×6 geopotential model. The resultant nominal orbit configurations can be seen in Fig. 4.4. Each simulation test case uses the same pseudo-random number generator seed to

Table 4.2: Simulated space object parameters

	1	**2**	**3**	**4**	**5**
Name	GEOSat1	GEOSat2	GEOSat3	GTOSat1	GTOSat2
TLE Date, t_0 (UTC)	5/20/2016	5/20/2016	5/20/2016	5/20/2016	5/20/2016
TLE Time, t_0 (UTC)	18:00	18:00	18:00	18:00	18:00
Semi-major Axis (km)	42000	42000	42000	25000	25000
Eccentricity	0.001	0.001	0.001	0.7	0.7
Inclination (°)	0.9195	0.6261	0.8925	3.9307	0.5717
Arg. of Perigee (°)	199.3991	218.1703	46.3874	294.9229	333.9905
RAAN (°)	291.0735	251.8371	347.0544	3.8344	66.4339
Mean Anomaly (°)	248.0162	303.3310	8.9211	206.4718	14.1699
Longitude (°E)	-129.03	-94.4	-105.14	n/a	n/a
Radar area (m^2)	32.4	32.4	32.4	21.4	21.4
Pos. Unc. $\sigma_{r,0}$ (km)	10	10	10	1	1
Vel. Unc. $\sigma_{v,0}$ (km/s)	1	1	1	.1	.1

ensure that the only difference between simulation cases are the true hypothesis resolutions and the observation conditions.

For the given M sensors and N space objects, there are N^M possible network tasking configurations at each time instant. The inclusion of the option "none of the above" or NOTA for each sensor increases the number of possible network tasking configurations to $(N+1)^M$. NOTA is employed whenever a sensor is inactive due to daytime or adverse weather, but is also employed when "Other" tasking objectives may be addressed. This addresses real-world sensor tasking resource demands, where the number of potential hypotheses and actions far exceeds the number of sensors. Since this example application is restricted in scope to remain illustrative, the "Other" category is employed to represent addressing other hypotheses. If two potential courses of action result in the same hypothesis resolution, the course of action with the least active sensors is chosen to free sensors for use in addressing other hypotheses or objectives. Considering a receding time horizon of H steps, the total number of tasking configurations for this sensor network over that horizon is $\left((N+1)^M\right)^H = (N+1)^{MH}$ As any of these parameters grows, evaluating each possible tasking configuration to find the best option, known as the "brute-force" approach, becomes intractable. For evaluation purposes in this study, the number of objects and sen-

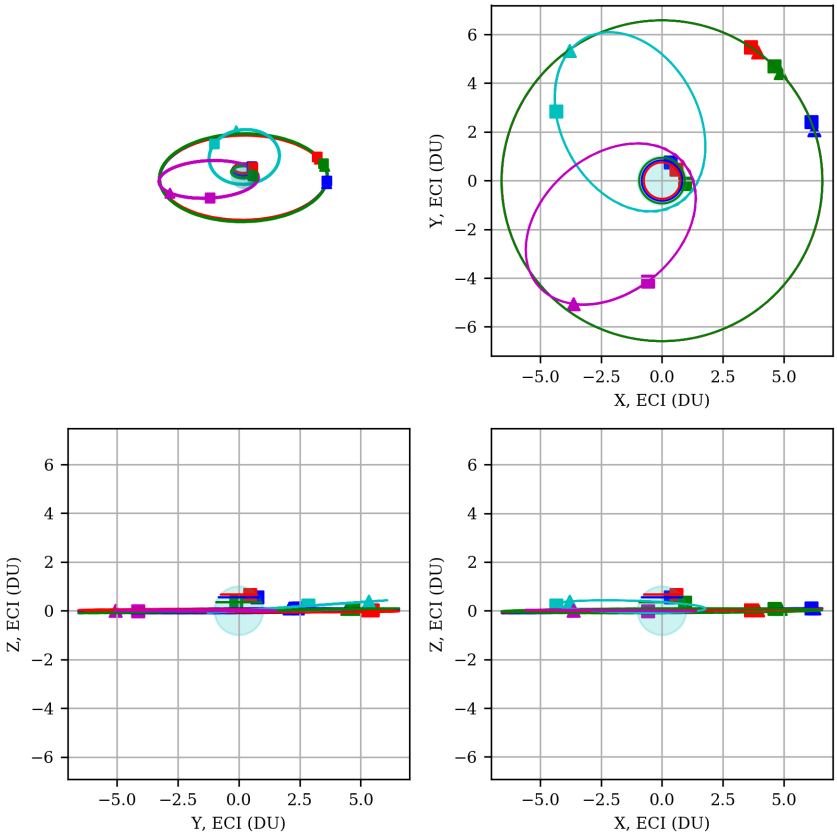


Figure 4.4: Propagated orbit and sensor geometries (triangles and boxes indicate simulation initial and final positions, respectively)

sors under consideration are low ($N = 5$, $M = 3$) and the greedy-optimization approach is used ($H = 1$), so the brute-force approach is feasible. Additionally, the evaluation of the minimum-ignorance objective function in this application is fast, so the evaluation for all $(N + 1)^{MH}$ configurations is completed well within the simulated tasking time-step ($t_{k+1} - t_k = 10$ minutes). Continuing research seeks to efficiently explore this candidate tasking configuration space to address problems with higher dimensionality and added realism, using mixed-integer optimization methods to solve the combinatorial optimization problem [82].

4.6.1 Case 1: Clear Observation, Nominal Dynamics

In the first test case, sensor tasking is simulated for 24 hours, with 10 minute tasking intervals. An equal weighting is applied to each space object: $w_i = \frac{1}{5} \forall\, i \in 1, \ldots, 5$. No maneuvers are prescribed, and observation conditions are clear and dark during the night hours, with nominal values shown in Table 4.1. The values for the GT sensor indicate deteriorated observation conditions due to city lights and pollution [15], using average values obtained from measured GT-SORT sky brightness. The USU and AMOS sensors are placed in less populated areas, so nominal sky brightness and atmospheric transmittance are adjusted accordingly. Between sunrise and sunset for each sensor, the background sky irradiance is set to $1.5 \frac{mag}{arcsec^2}$, significantly brighter than even the brightest measured sky brightness data. This dataset provides the baseline test for the sensor network operating in all-around nominal conditions and provides an illustrative example for data analysis that must be understood before examining more complex scenario results.

The sensor tasking schedule is shown in Fig. 4.5(a), formatted similar to a Gantt chart to indicate how much time is being spent by each sensor on each action (including the "none of the above" option, or NOTA). The hash-marked sections of the schedule indicate the times when a space object is unobservable to a particular sensor, which might be due to excessive sky brightness (e.g. daylight), low atmospheric transmittance (e.g. cloud cover),

or no line-of-sight (space object below the minimum elevation threshold for observation). Figure 4.5(b) shows ignorance in the anomaly hypothesis for each space object, as well as the weighted total sum. Recall that the goal is to minimize the weighted total anomaly ignorance, so the sensor network selects the actions that minimize this value at each time step. Figure 4.6 shows the resultant Dempster-Shafer belief structure for each space object. This includes belief, plausibility, and ignorance in both custody and anomaly, as well as the pignistic belief ("bet") associated with each individual proposition: detected and nominal ($\mathcal{D} \cap \mathcal{N}$), detected and anomalous ($\mathcal{D} \cap \mathcal{A}$), missed and nominal ($\mathcal{M} \cap \mathcal{N}$), missed and anomalous ($\mathcal{M} \cap \mathcal{A}$).

During the first few hours of the tasking window, all three sensors are in daylight, so the NOTA option is selected since any observations would not provide usable information (a missed detection is most likely due to excessive sky brightness). During this period, state uncertainty and reachable distance grows, increasing anomaly ignorance as the 3-σ uncertainty bounds exceed each sensor's field of view (as seen in Fig. 4.5(b) and Fig. 4.6). When the space objects become observable, the sensors begin to task against them and attempt observations to detect and confirm their states. Being the first sensor to exit daylight, the GT sensor quickly scans all available targets, significantly restraining the increase in ignorance. When a space object is detected, the associated anomaly hypothesis ignorance drops as it is resolved to either nominal or anomalous (in this case, always being resolved to nominal). Additionally, the custody hypothesis is resolved to indicate that custody is kept, so the pignistic belief in the "detected-nominal" state for that object jumps to one while the others all drop to zero (see Fig. 4.6). Therefore, successful detection can be deduced from the Dempster-Shafer belief structures through this change in pignistic belief.

The GT sensor also revisits each GEO space object to ensure the uncertainty and reachability volumes stay within the field of view. Soon-after, the USU sensor becomes usable and also aids in re-checking some of the space objects. The use of geometrically-diverse

electro-optical sensors allows reduction of uncertainty in slightly different axes, further controlling the growth of anomaly ignorance. Finally, the AMOS sensor becomes usable and quickly checks GTOSat1, since it is the only sensor that is able to successfully observe this object. Multiple observations at different points in the GTO object's orbit also allow further reduction in state uncertainty and therefore anomaly ignorance.

The ignorance-reduction methods take a relatively limited number of actions against the relevant space objects because further actions would not lead to improved hypothesis resolution. After the initial detection of each space object, follow-on observations are not required immediately, allowing state uncertainty and reachability volumes to grow. Once the volumes are large enough to start affecting anomaly hypothesis resolution, follow-on observations are executed. This aspect is a strength of hypothesis-based methods, employing the minimal sensor resources required to reach the best possible (weighted) hypothesis resolution and allowing reassignment of extra sensors to other tasks or priorities. This is important as operational scenarios are significantly more complicated and involve many more priorities that place a strain on sensor network resources.

It is worth noting that ignorance for each GTO space object grows significantly faster than ignorance for the GEO space objects due to quicker-growing state uncertainty near perigee. Ignorance also fluctuates and can reduce naturally (without being tasked by a sensor) as the space object approaches apogee since the projected field-of-view of each sensor is wider with an increased slant range.

Toward the end of the simulation, each sensor enters daylight once again and the anomaly ignorance continues to grow with increased uncertainty. The GEO objects all show increased anomaly ignorance at this time, which cannot be reduced due to daytime sky brightness; however, had the sensors taken observations of each space object just before entering daylight this could have been mitigated. Since the greedy optimization approach was tested here, the sensor network could not look ahead to see this eventuality. This indicates a limitation of the greedy optimization approach that the receding horizon approach

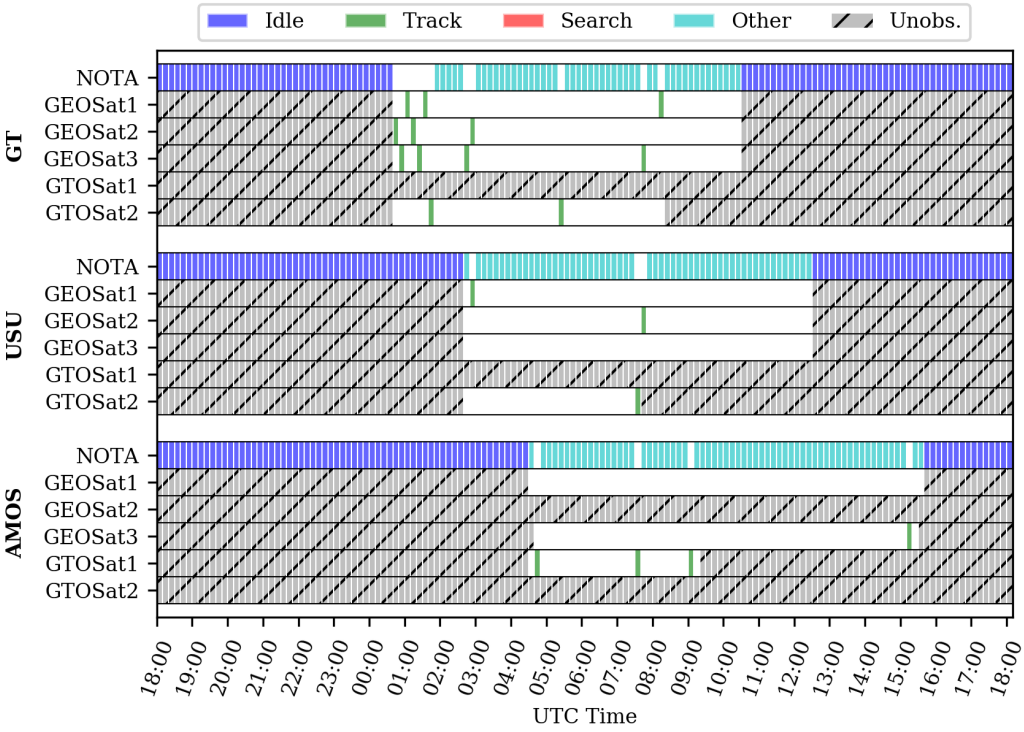
can counteract, targeting minimum ignorance at a future time.

In total, each GEO object is tasked and successfully detected a total of four times, and each GTO object is tasked and successfully detected three times. Overall, the sensor network is relatively passive since the FOV thresholds are not surpassed quickly, so the majority of sensor time is spent on the $NOTA$ option in the "other" tasking mode; tasking to a different target would not cause a loss in ignorance-reduction in any of the hypotheses. This indicates that this simulated sensor network could handle a significant increase in the number of tracked targets or hypotheses considered under nominal conditions, or handle tighter tolerances on allowable uncertainty for the current hypotheses. However, nominal conditions are too idealistic for application in real-world scenarios, so the following cases evaluate performance in the case of off-nominal conditions.

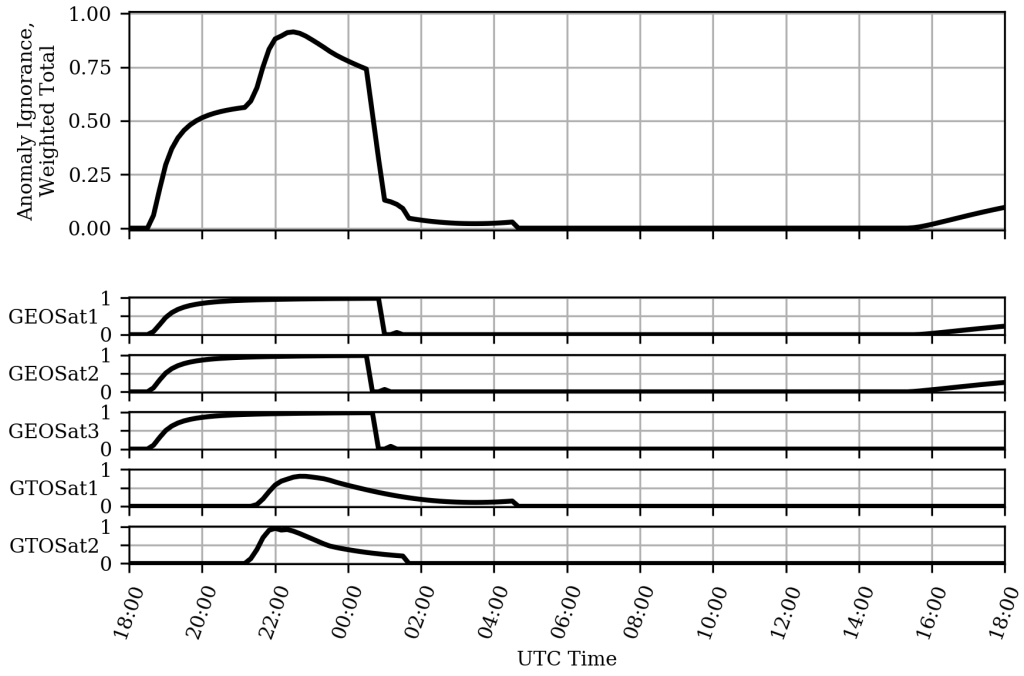
4.6.2 Case 2: Unclear Observation, Nominal Dynamics

In the second test case, most parameters are identical to the first case: sensor and space object parameters (see Tables 4.1 and 4.2), simulation timing, space object weighting, and nominal observation conditions. However, poor observation conditions are induced at specific times in the simulation to cause the sensor network to compensate and retask. To simulate the presence of cloud cover, atmospheric transmittance τ_{atm} is reduced to zero between 18:00-3:00 UTC for the GT sensor and between 7:00-10:00 UTC for the USU sensor (compare the "unobservable" times between Fig. 4.5(a) and Fig. 4.7(a)).

The result is that the USU sensor is the first to enter clear observation conditions, so it takes the first pass at detecting all the GEO objects and the one observable GTO object. As with Case 1, once targets are observable to the GT sensor observations are conducted successfully. Since the GT sensor has a narrow instantaneous field of view, its observations reduce state uncertainty significantly. The performance of the AMOS sensor does not change significantly: it is still the only sensor able to observe GTOSat1, and its detection of GEOSat3 occurs slightly later due to the delay in initial acquisition caused by weather



(a) Sensor tasking schedule



(b) Anomaly ignorance

Figure 4.5: Case 1 - Sensor tasking schedule and anomaly ignorance

conditions in Atlanta. The added cloud cover in Utah later in the simulation also causes the AMOS sensor to reacquire GEOSat1.

The ignorance results in Fig. 4.7(b) and Dempster-Shafer belief structure results in Fig. 4.8 are very similar to the results from Case 1: anomaly and custody ignorance are both kept low once tasking begins, and the pignistic belief in the detected-nominal proposition $\mathcal{D} \cap \mathcal{N}$ remains near one. The total number of detections decreases slightly, with one less detection of GEOSat3 simply due to impact of weather on detection timing, but the effect on anomaly ignorance is negligible. The sensor network successfully accounts for a predicted deterioration in observation conditions and maintains performance similar to the fully nominal case.

4.6.3 Case 3: Clear Observation, Anomalous Dynamics

In test Case 3, the same simulation parameters from Case 1 are used, but a maneuver is imposed on GEOSat2 to test the ability for the sensor network to identify the anomaly. The truth data for GEOSat2 is modified to include a roughly 0.5 degree change in mean anomaly at the TLE epoch, similar to an East-West station-keeping maneuver. On the first missed detection, the custody hypothesis shows the miss and the sensor network determines the observation conditions should not have caused the missed detection. The sensor network flags the space object as anomalous and initiates a search to reacquire GEOSat2.

Since this occurs early in the simulation, the GT sensor is the only one available for searching. The GT sensor has a very small field of view, so it does not interrogate much of the reachable volume with each attempt and continues to miss observations. Meanwhile, the uncertainty and reachability volumes continue to grow.

Once the anomalous space object is detectable to the wider-field-of-view USU sensor, reacquisition responsibilities shift and the Utah sensor detects it on the first try. At this point the anomaly is resolved, and the space object estimator is updated. The anomaly is confirmed in a follow-up observation of GEOSat2: the detected-anomalous, $\mathcal{D} \cap \mathcal{A}$, pignistic

belief in Fig. 4.10(b) where it jumps to one.

Through the rest of the simulation, anomaly ignorance is kept in check with follow-on observations for all the objects. The schedule of observations in Fig. 4.9(a) changes slightly as a result, but the end-result anomaly ignorance in Fig. 4.9(b) is roughly the same as in Case 1.

4.6.4 Case 4: Unclear Observation, Anomalous Dynamics

In Case 4, the same simulation parameters from Case 2 are used, and the same maneuver from Case 3 is applied: roughly 0.5 degree change in mean anomaly for GEOSat2. In this case, the GT sensor cannot observe it first due to cloudy conditions. The USU sensor is able to acquire it on the first attempt due to its wider field of view, but the estimator update from this observation is not sufficient to fully capture the anomaly. This can be seen by the attempted follow-up observation of GEOSat2 after 9:00 UTC, where the GT sensor enters a search mode to find it. At this time, the USU sensor is covered by poor observation conditions, and the AMOS sensor can not observe it due to low elevation. After 10:00 UTC, the USU sensor is available once more, and it reacquires GEOSat2 and captures the anomaly as seen in Fig. 4.12(b).

As with the other test cases, these changes in observation conditions affects the order of observations as shown in the schedule (Fig. 4.11(a)). However, the performance as measured by ignorance (Fig. 4.11(b)) does not change significantly.

4.6.5 Comparison to Covariance-Minimization

A covariance-minimization scheduler was also implemented and tested in the same cases as above. This scheduler selects the actions that minimize the weighted-sum of the traces of the covariances at each time step. A subset of these results are shown here for comparison to ignorance-reduction.

Figure 4.14 shows the covariance-minimizing schedule. Notably, this tasking scheme

takes far more actions than ignorance-reduction since every successful detection prevents uncertainty from growing through state propagation. The only times when "other" actions are taken are when all visible objects are already being tracked by other sensors. However, Fig. 4.13 shows that the hypothesis resolution is similar to the ignorance-reduction results in Fig. 4.5. This highlights an important distinction between the methods. Covariance-minimization will continuously observe objects to reduce covariance, but that reduced covariance does not necessarily lead to improved hypothesis resolution. Ignorance-reduction takes only the actions necessary to resolve the hypotheses, which may be preferred if there are other secondary objectives or hypotheses to address ("other" actions). Throughout the sensor-tasking portion of the simulation, covariance-minimization makes 176 tasking actions compared to 19 actions for ignorance reduction, while achieving the same hypothesis resolution.

The hypothesis resolutions only differ at the end of the simulation, during the unobservable post-sunrise portion. Anomaly ignorance begins to increase again for ignorance-reduction because state uncertainty increases throughout the unobservable portions of the simulation. The covariance-minimization approach has lower state covariance at the end of the observable portion of the simulation, and therefore covariance is lower at the end of the simulation as well leading to lower anomaly ignorance. This discrepancy can be rectified for ignorance-reduction using the receding-horizon approach with a sufficiently long time horizon to capture the ignorance increase. A longer-horizon ignorance-reduction approach would schedule final pre-sunrise observations to reduce covariance, and thereby ignorance, at the end of the horizon, as observed in preliminary work on ignorance-based tasking [90]. Similarly, a receding-horizon covariance-minimization approach may delay some actions in order to allow covariance to increase while focusing on other priorities. The delayed actions result in larger single-action reductions in covariance, useful in scenarios where the number of objects far exceeds the number of available sensors. The drawback of longer-horizon approaches is the exponential increase in computational complexity, and

continuing research focuses on addressing computational complexity requirements for sensor tasking problems [82].

The previous comparison is uneven in the amount of information available for hypothesis resolution, since the covariance-minimizing scheme takes many more actions against the hypotheses. For an information-equivalent comparison, the same case was run while limiting the covariance-minimization technique to actions only at the same times that the ignorance-reduction scheme took actions. This simulates that the "other" actions taken were required to address other objectives, and only a limited number of observations are allowed for these anomaly hypotheses. This added constraint emulates the presence of other priorities that limit sensor resources without requiring a vast expansion of the simulated search space. Figure 4.16 shows the resulting sensor schedule, and Fig. 4.15 shows the resulting covariances and hypothesis resolution. In this case, the covariance-minimization scheme chooses actions which do not resolve the hypotheses with the same clarity as ignorance-reduction, leading to increased anomaly ignorance throughout the simulation. This shows that the minimum-uncertainty actions are not necessarily the actions which lead to the best hypothesis resolution, a result related to the coupling of covariance and hypothesis evidence.

This trend repeats for all four test cases, so only the baseline (Case 1) and most challenging (Case 4) scenario results are shown for comparison. For instance, the anomalous unclear (Case 4) schedule in Fig. 4.18 and results in Fig. 4.17 show identical characterization of the anomaly event, using 149 actions for covariance-minimization as compared to the 18 actions used in ignorance-reduction. Once again, few "other" actions are taken as each successful detection reduces the covariance, but once the actions are limited the anomaly discrimination suffers. The limited schedule in Fig. 4.20 and results in Fig. 4.19 also show that, with limited actions, tasking to minimize covariance did not result in better hypothesis resolution than tasking to minimize ignorance in these hypotheses. The covariance-minimization results for Cases 2 and 3 are omitted since the observed trends

are the same.

The comparison results, summarized in Fig. 4.21, show that ignorance-reduction performed as well as covariance-minimization in resolving the anomaly hypotheses during the tasking window, even though the covariance minimization approach used many more actions. In many cases during these simulations, ignorance cannot be further reduced by follow-on observations (until the uncertainty and reachability volumes have a chance to grow again through propagation), so in ignorance-based tasking the sensors are allowed to be assigned to other priorities. In contrast, the state estimate covariance can always be reduced with follow-on observations, so covariance-based tasking will always take any available actions to make incremental improvements. The ignorance-reduction method takes strong actions toward resolving the hypotheses, and when covariance-minimization is limited to taking actions at the same times as ignorance-reduction, the covariance-minimizing actions often resulted in weaker hypothesis resolution. Taking far fewer actions may be considered a benefit of the ignorance-reduction approach if secondary objectives or hypotheses exist that can be investigated through "other" actions, as is likely the case in more complicated operational contexts.

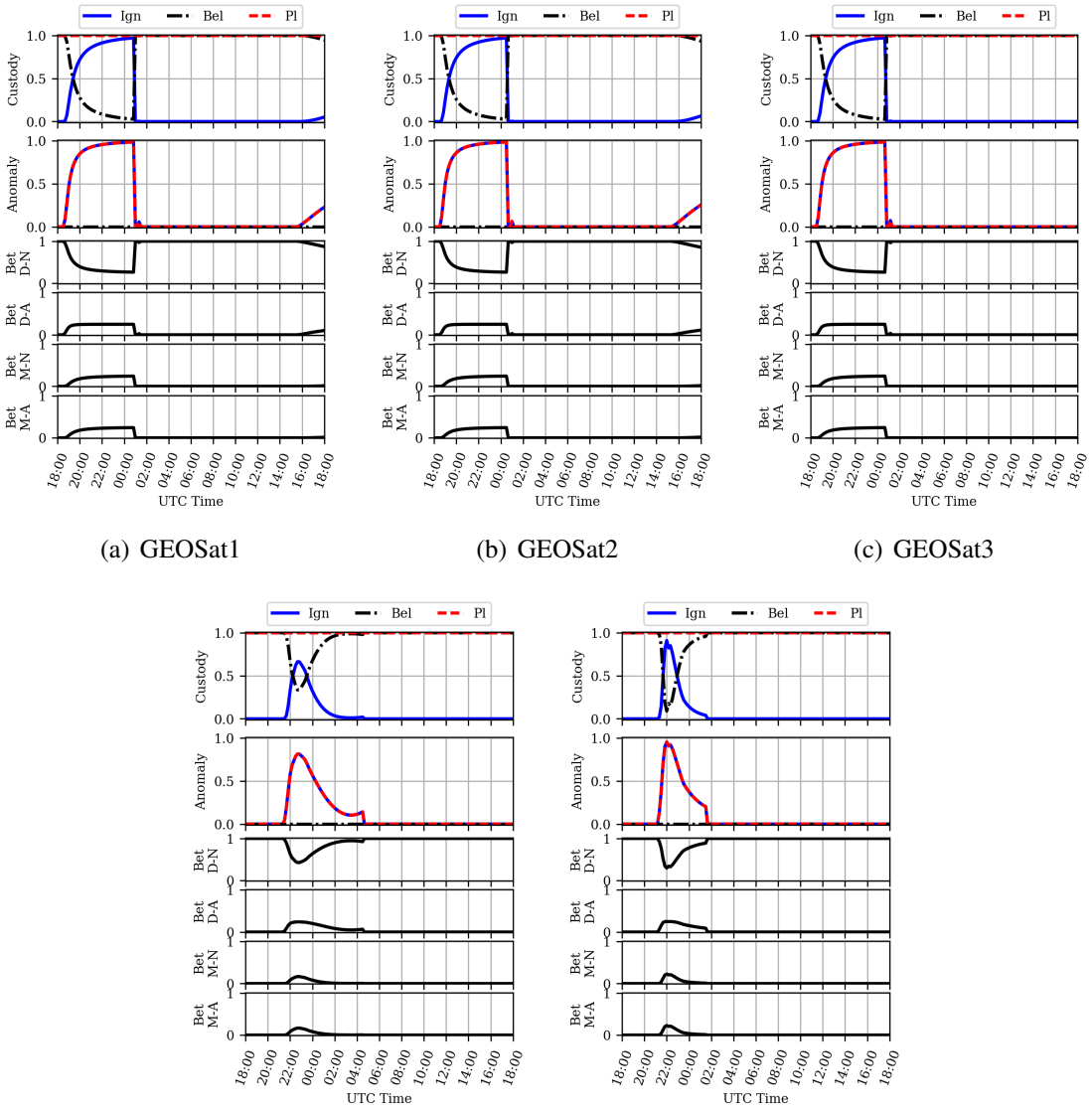
4.7 Conclusions

Current state-of-the-art sensor tasking for SSA focuses on maximizing information or minimizing uncertainty and is largely reactionary. This work proposes a formulation of the sensor tasking problem that addresses specific hypotheses using Dempster-Shafer theory to interrogate these hypotheses with sensor evidence. This hypothesis-based approach allows incorporation of a wide variety of SSA sensors as evidence-based experts to predict hypothesis resolution quality and select the set of actions that minimizes hypothesis ambiguity.

Applied to a space object custody and anomaly discrimination scenario, this framework is able to generate sensor tasking schedules based on predicted observation conditions to ensure custody is maintained. This is accomplished through maximizing ignorance-

reduction in the hypothesis of interest, which selects only tasking combinations that provide high-quality evidence. Additionally, in the case of missed observations, the framework initiates a search that incorporates the same SSA sensor data to quickly reacquire the lost space object and characterize the anomaly. The sensor network maintains custody of the objects and minimizes hypothesis ignorance with limited sensor effort, freeing up the sensors to perform other activities in-between custody measurements. In comparison to a covariance-minimization approach, the ignorance-reduction approach performs just as well in hypothesis resolution with far fewer actions required. When covariance-minimization actions are limited to the same times as ignorance-reduction, the ignorance-reduction actions result in better hypothesis resolution. This emphasizes that the minimum-covariance actions do not always result in optimal hypothesis resolution, depending on the coupling between hypothesis evidence and covariance.

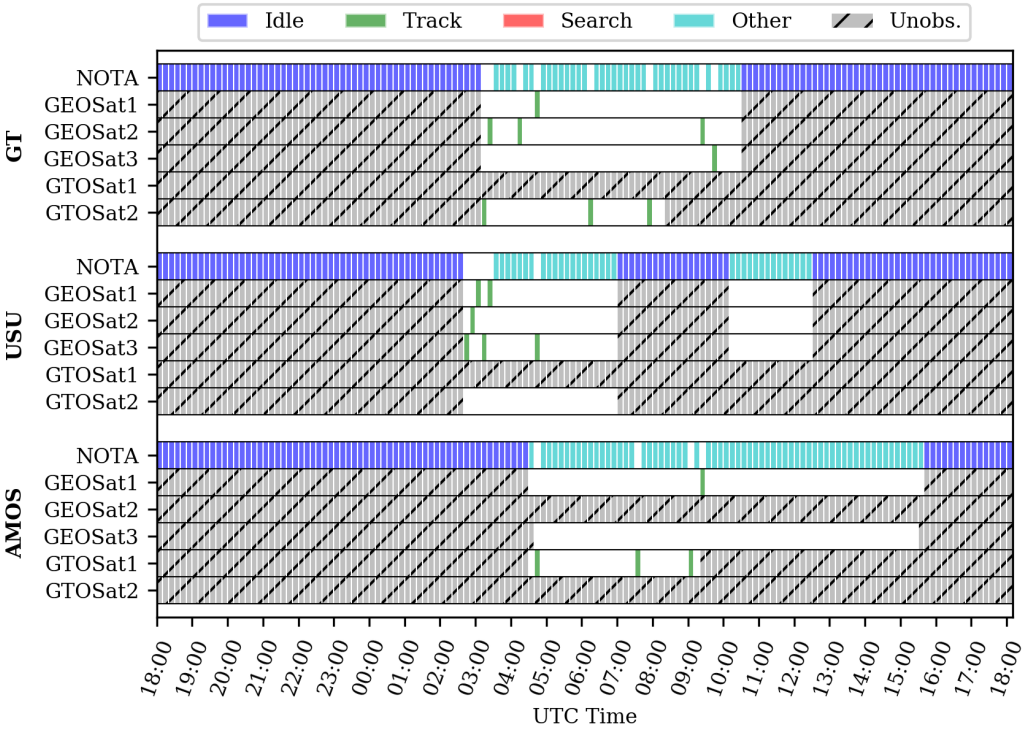
Taking far fewer actions to achieve similar hypothesis resolution may be considered a benefit of the ignorance-reduction approach if secondary objectives or hypotheses exist that can be investigated through "other" actions. In a more complex operational scenario, there are likely many more space objects and associated hypotheses, placing additional strain on the sensor network to gather strong evidence in limited actions. The results of the comparisons show that the ignorance-based approach selects actions that contribute strong evidence based on hypothesis priorities. In the event that the number of objects (i.e. evidence) far exceeds the sensor network tasking capabilities, the covariance-minimizing actions are not guaranteed to result in the same quality of hypothesis resolution. Increased numbers of objects and potential actions also magnify the computational complexity concerns associated with brute-force resource allocation evaluations, which must be addressed for operational implementation. The following chapter expands on the use of evidential reasoning for evidence-gathering to address these and other practical implementation concerns for hypothesis resolutions tasks.



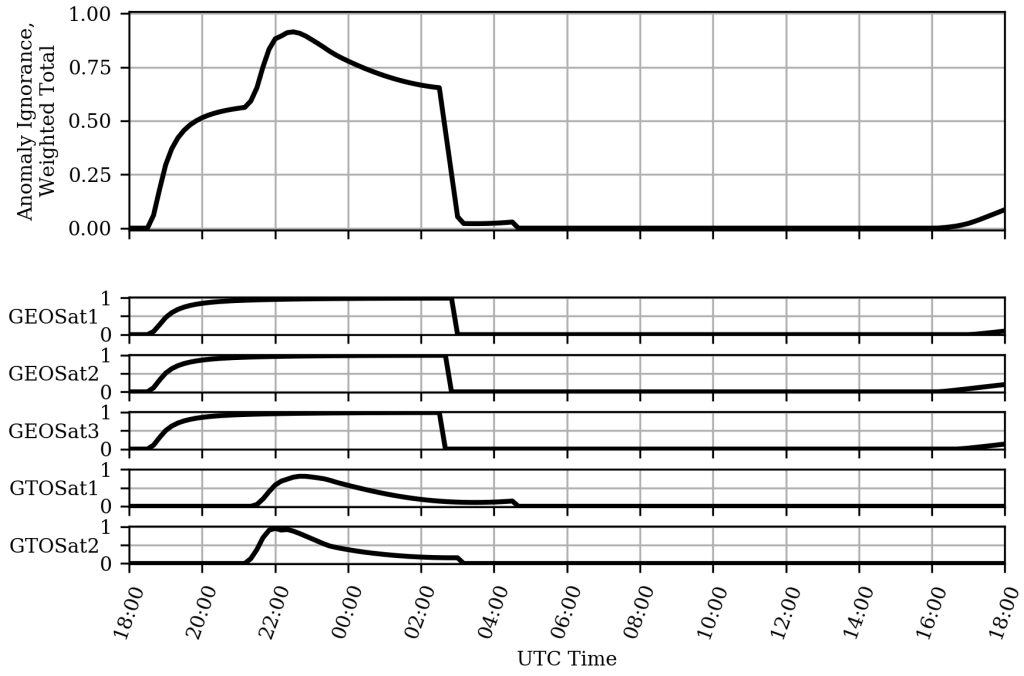
(a) GEOSat1 (b) GEOSat2 (c) GEOSat3

(d) GTOSat1 (e) GTOSat2

Figure 4.6: Case 1 - Belief assignment

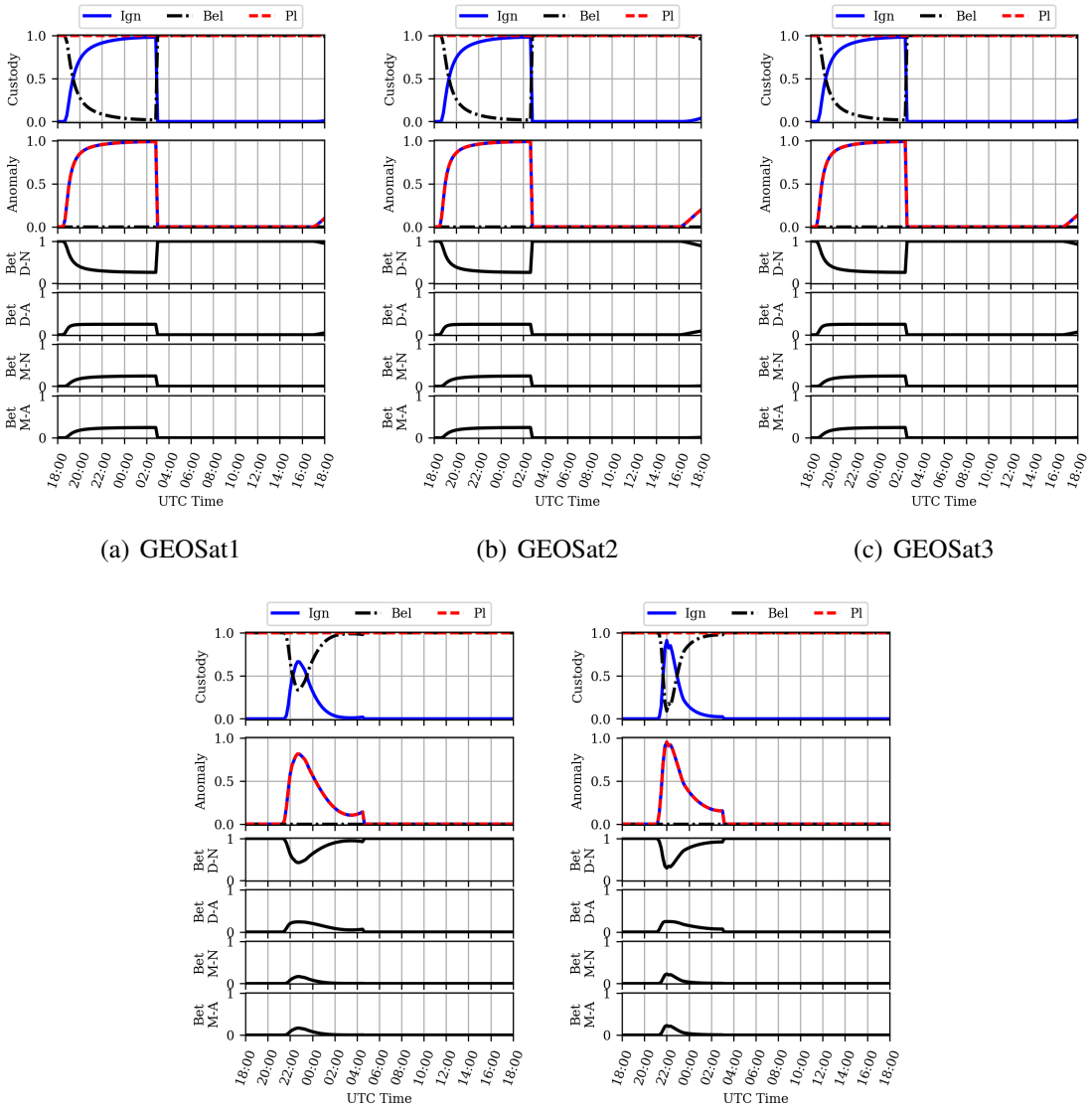


(a) Sensor tasking schedule



(b) Anomaly ignorance

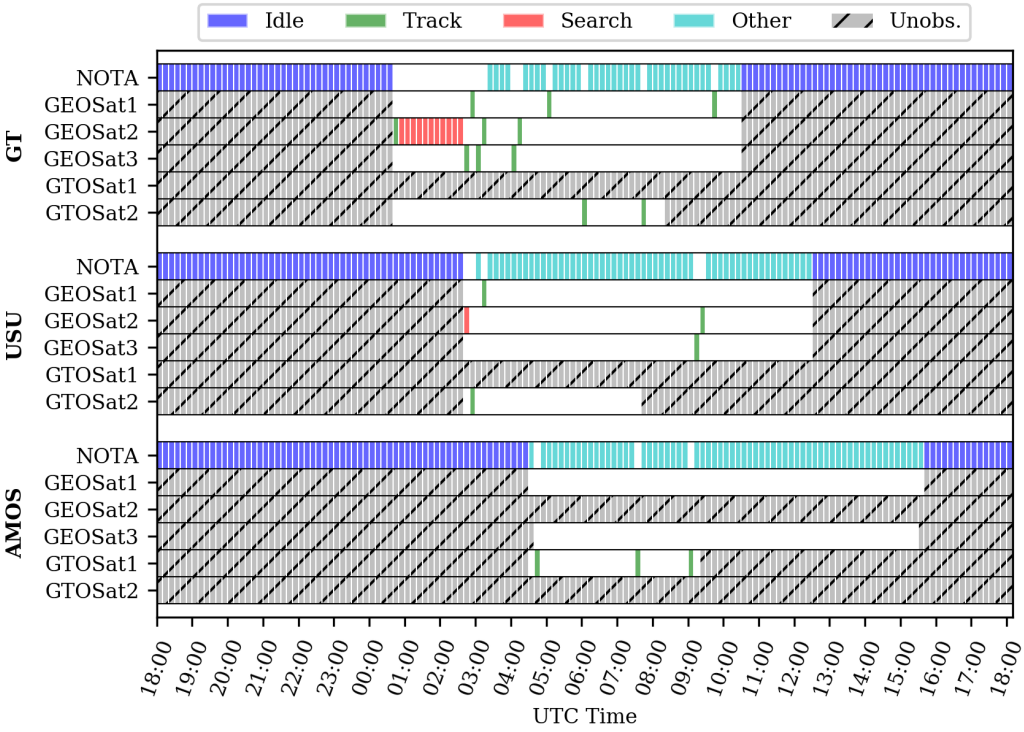
Figure 4.7: Case 2 - Sensor tasking schedule and anomaly ignorance



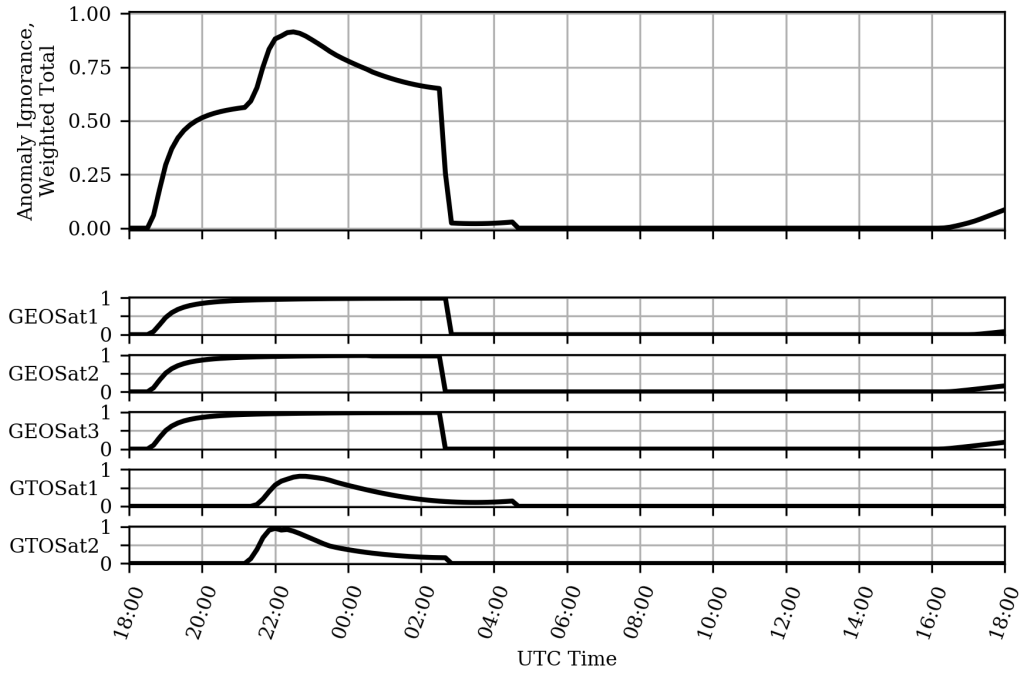
(a) GEOSat1 (b) GEOSat2 (c) GEOSat3

(d) GTOSat1 (e) GTOSat2

Figure 4.8: Case 2 - Belief assignment results

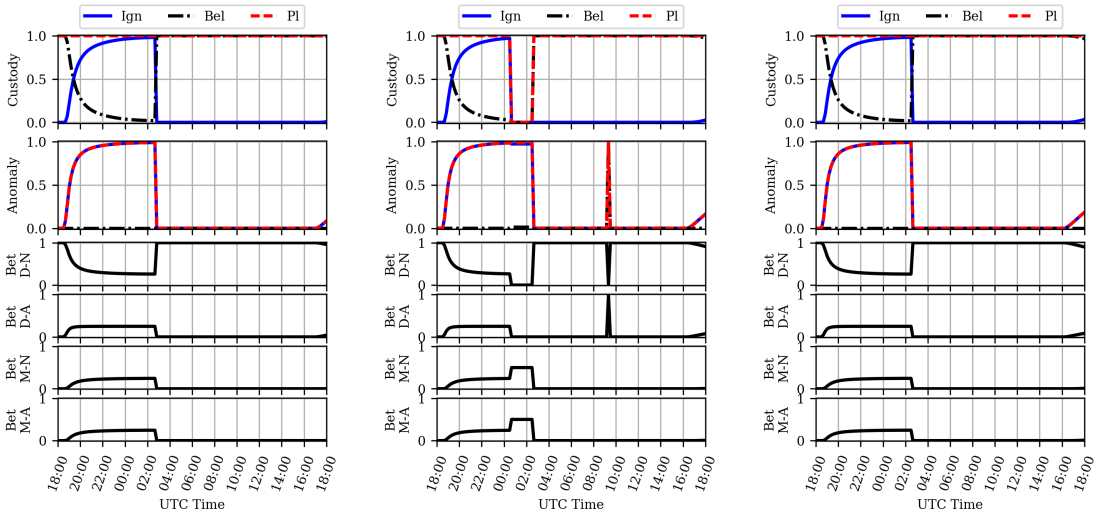


(a) Sensor tasking schedule

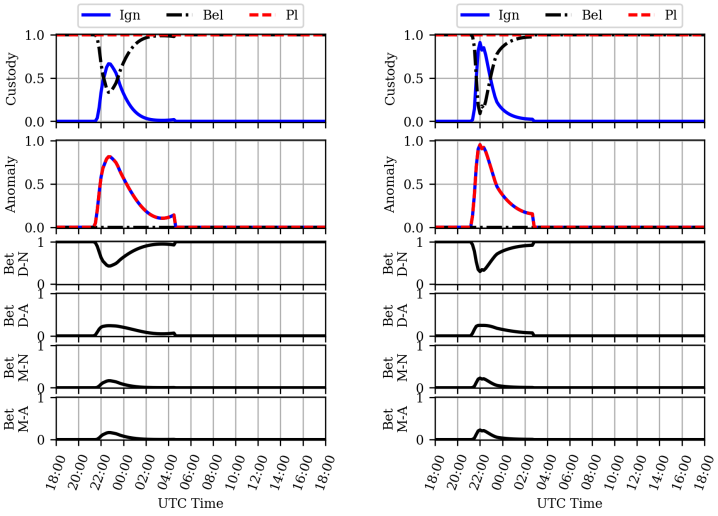


(b) Anomaly ignorance

Figure 4.9: Case 3 - Sensor tasking schedule and anomaly ignorance

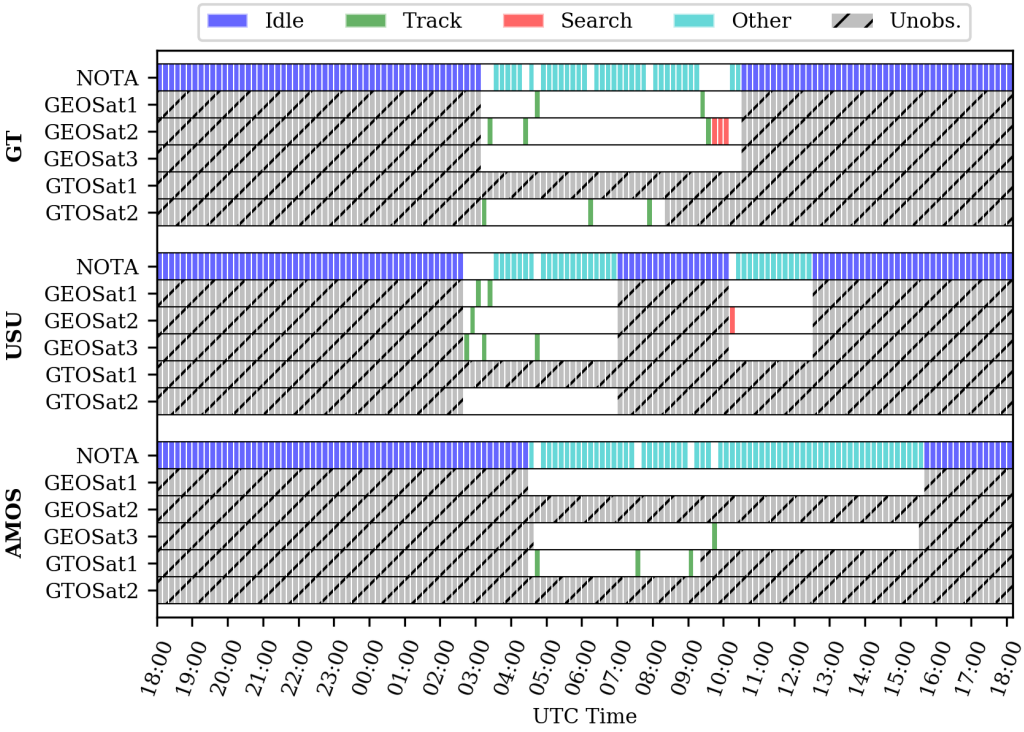


(a) GEOSat1 (b) GEOSat2 (c) GEOSat3

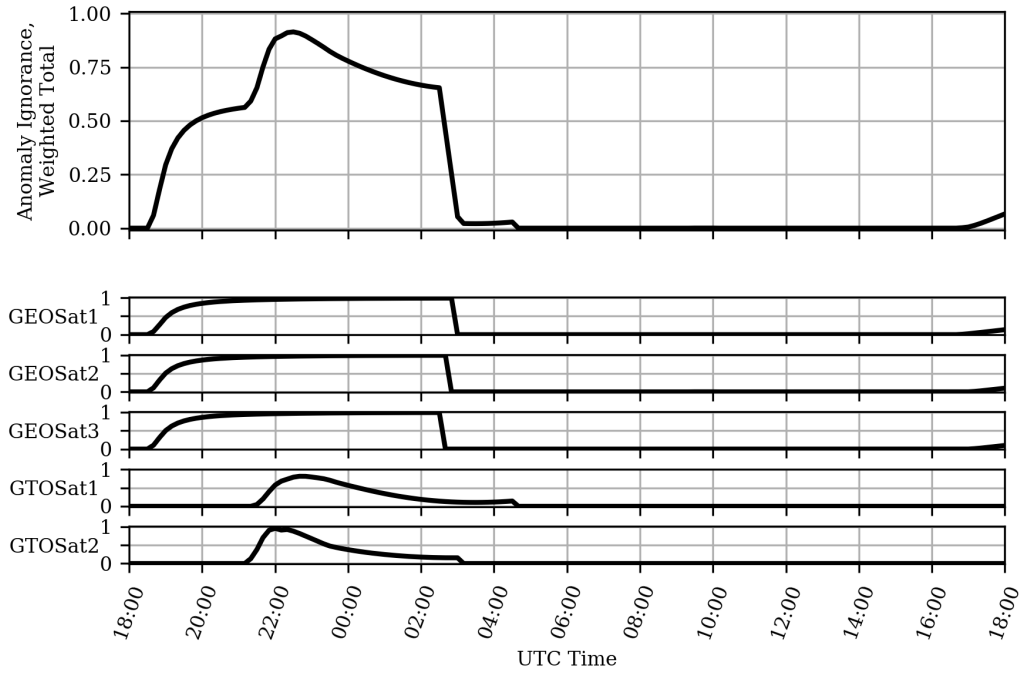


(d) GTOSat1 (e) GTOSat2

Figure 4.10: Case 3 - Belief assignment results

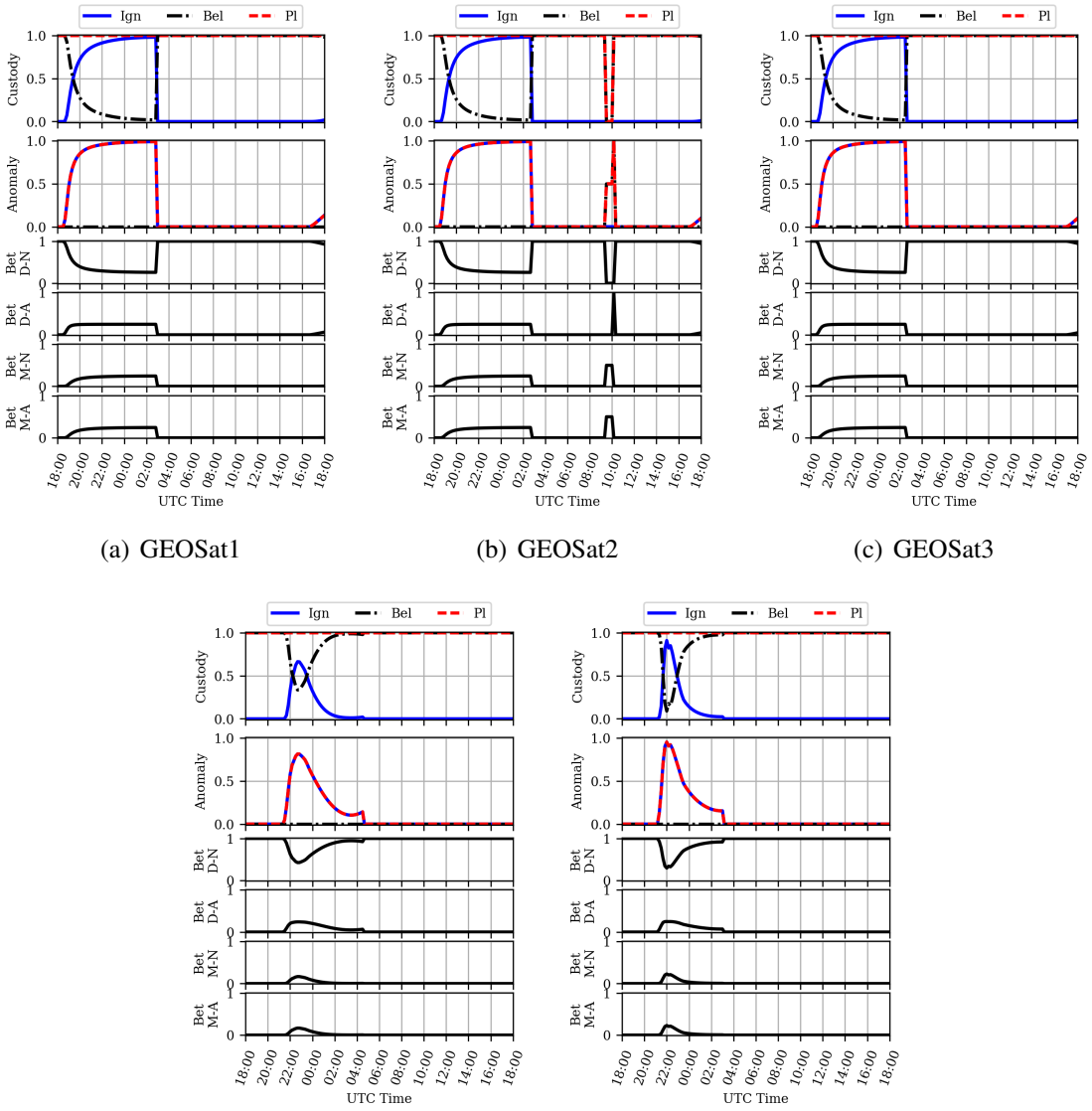


(a) Sensor tasking schedule



(b) Anomaly ignorance

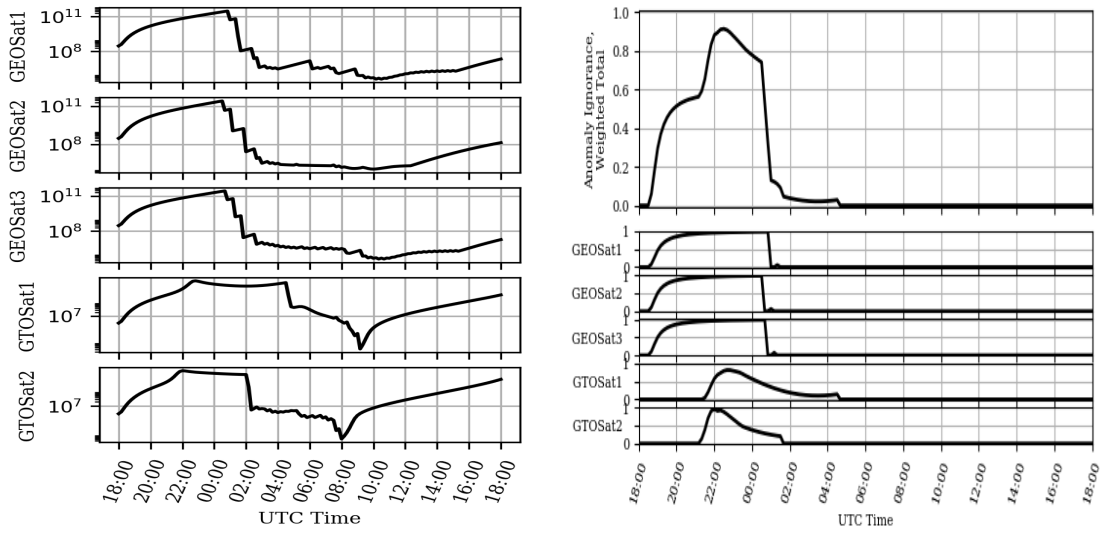
Figure 4.11: Case 4 - Sensor tasking schedule and anomaly ignorance



(a) GEOSat1 (b) GEOSat2 (c) GEOSat3

(d) GTOSat1 (e) GTOSat2

Figure 4.12: Case 4 - Belief assignment results



(a) Covariance traces (b) Anomaly hypothesis ignorance

Figure 4.13: Case 1: Covariance-minimization

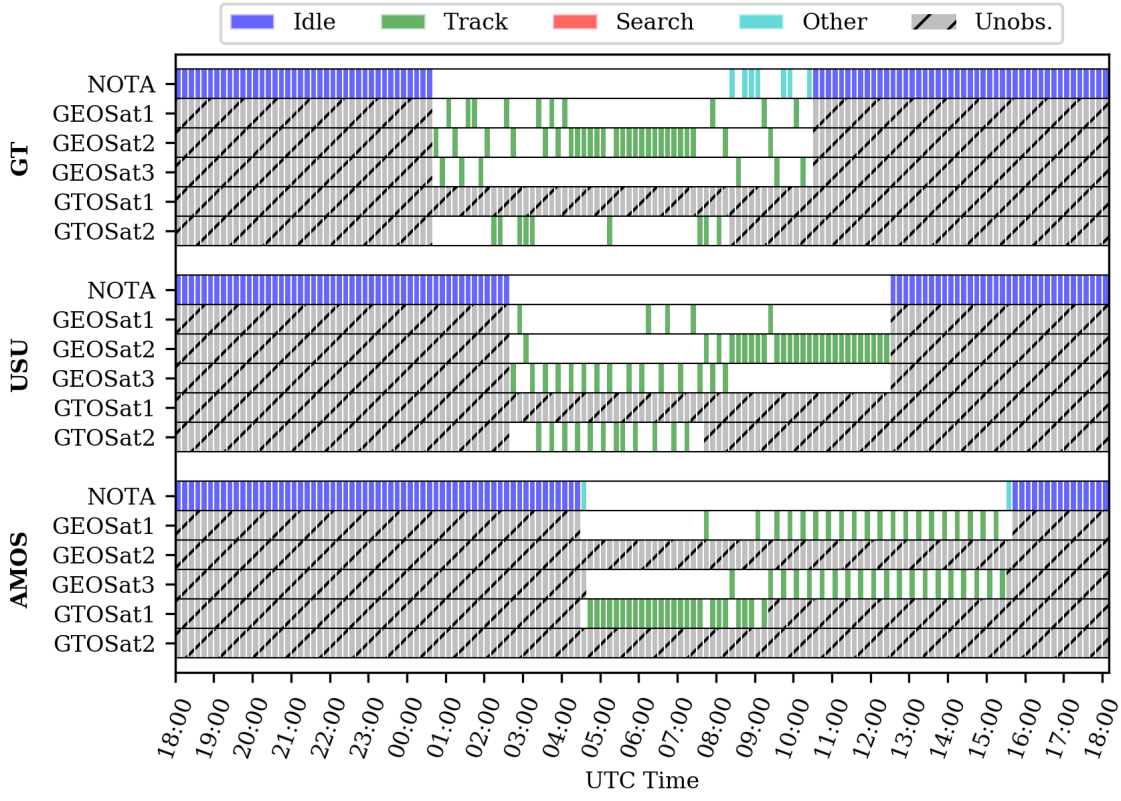
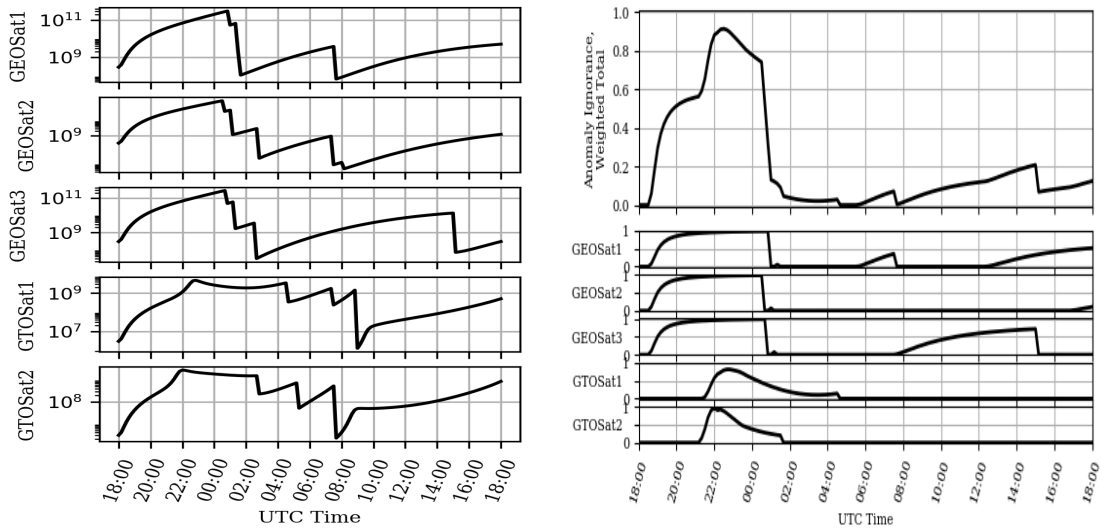


Figure 4.14: Case 1: Covariance-minimization sensor tasking schedule and anomaly ignorance



(a) Covariance traces (b) Anomaly hypothesis ignorance

Figure 4.15: Case 1: Covariance-minimization (limited)

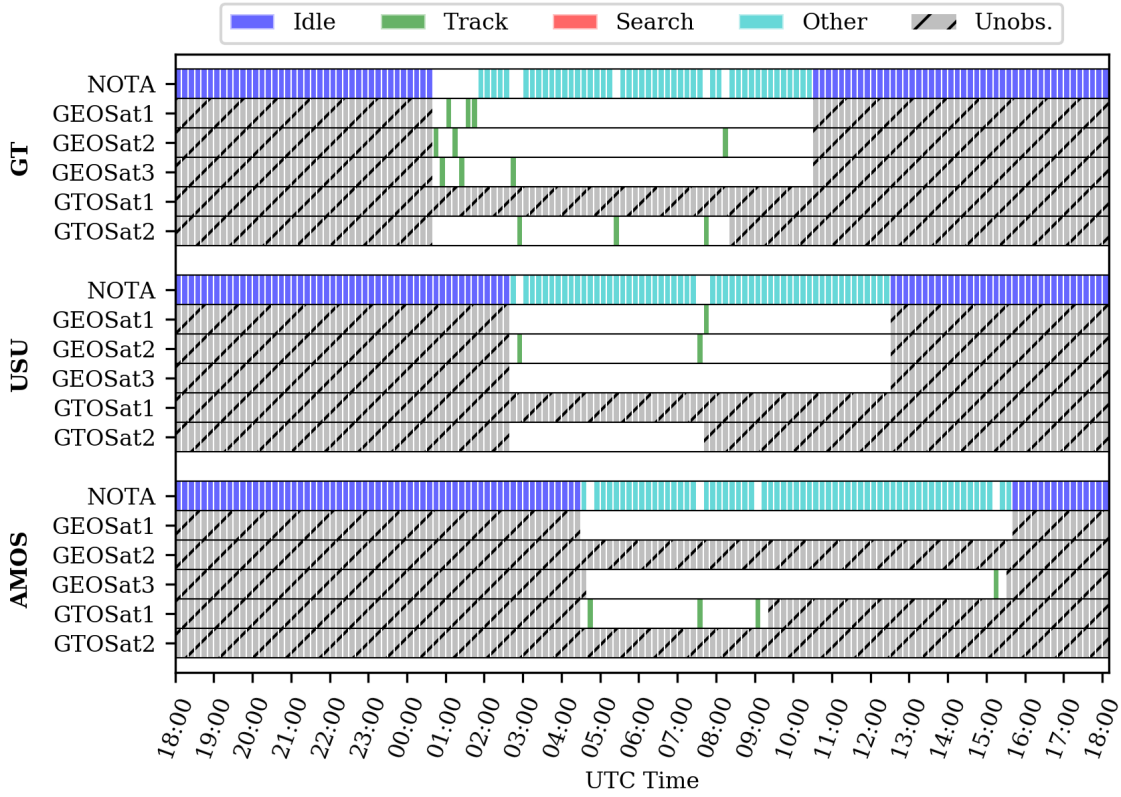
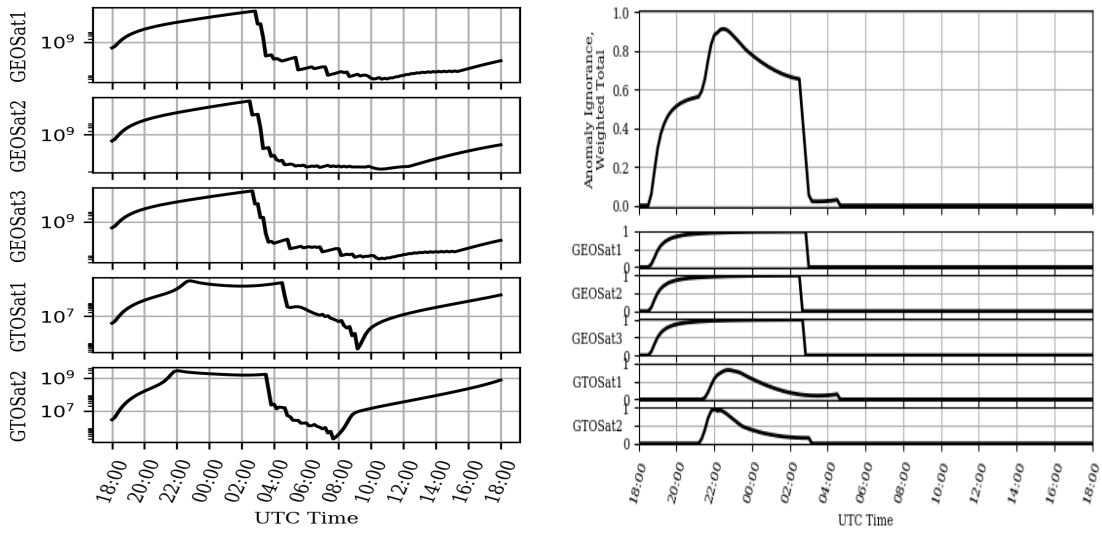


Figure 4.16: Case 1: Covariance-minimization (limited) schedule and anomaly ignorance



(a) Covariance traces (b) Anomaly hypothesis ignorance

Figure 4.17: Case 4: Covariance-minimization

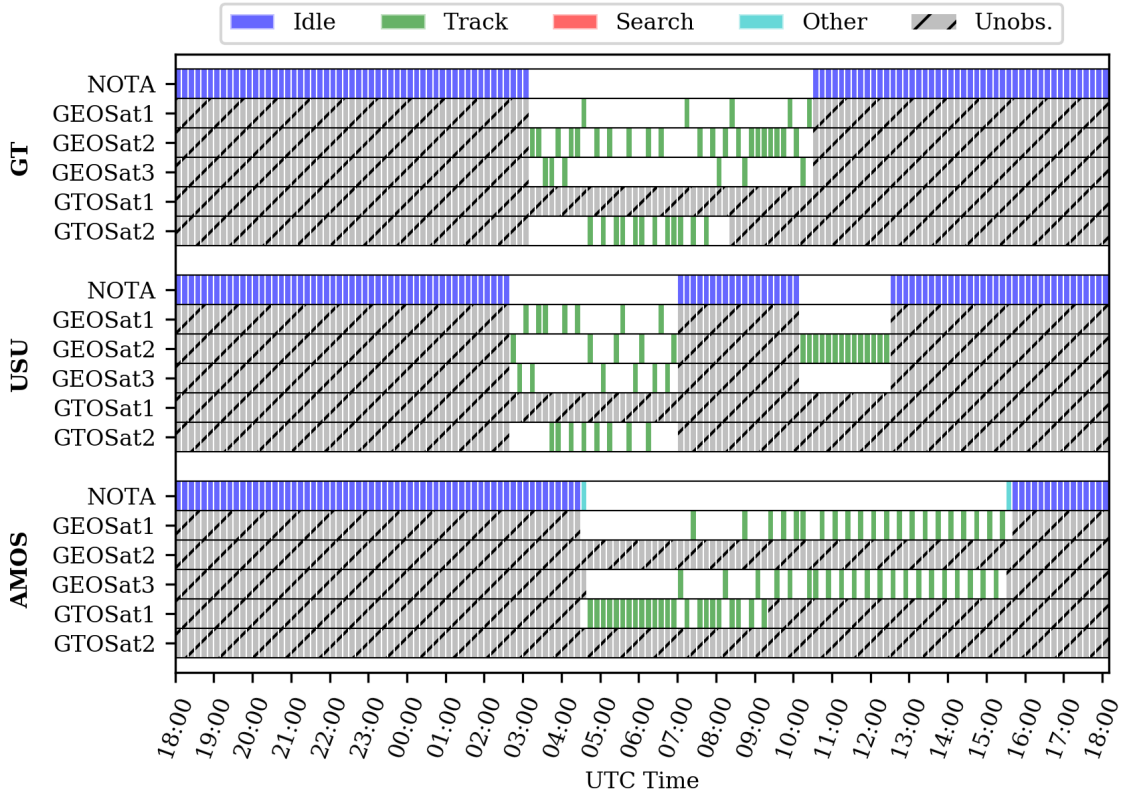
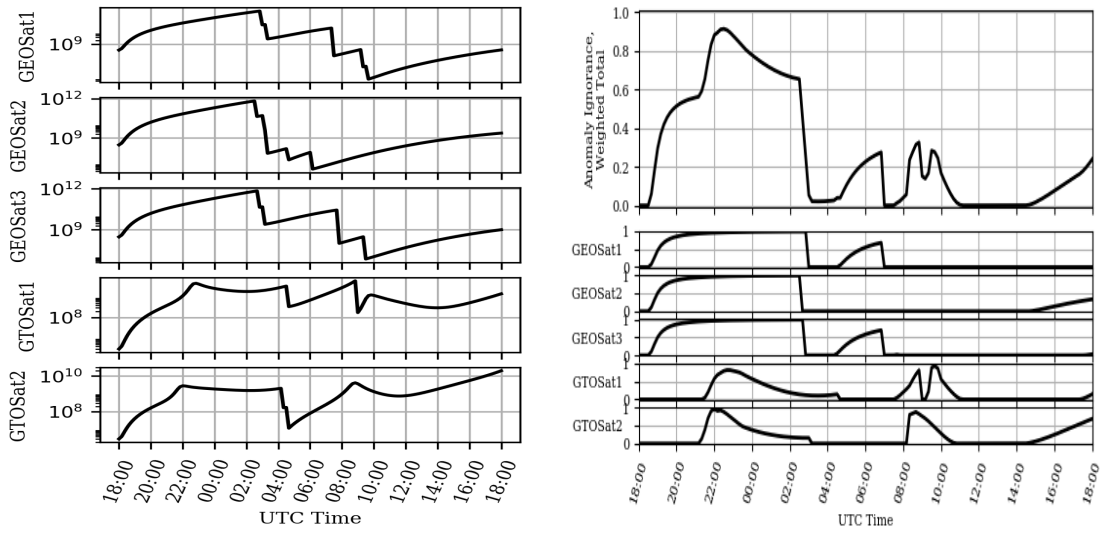


Figure 4.18: Case 4: Covariance-minimization sensor tasking schedule and anomaly ignorance



(a) Covariance traces (b) Anomaly hypothesis ignorance

Figure 4.19: Case 4: Covariance-minimization (limited)

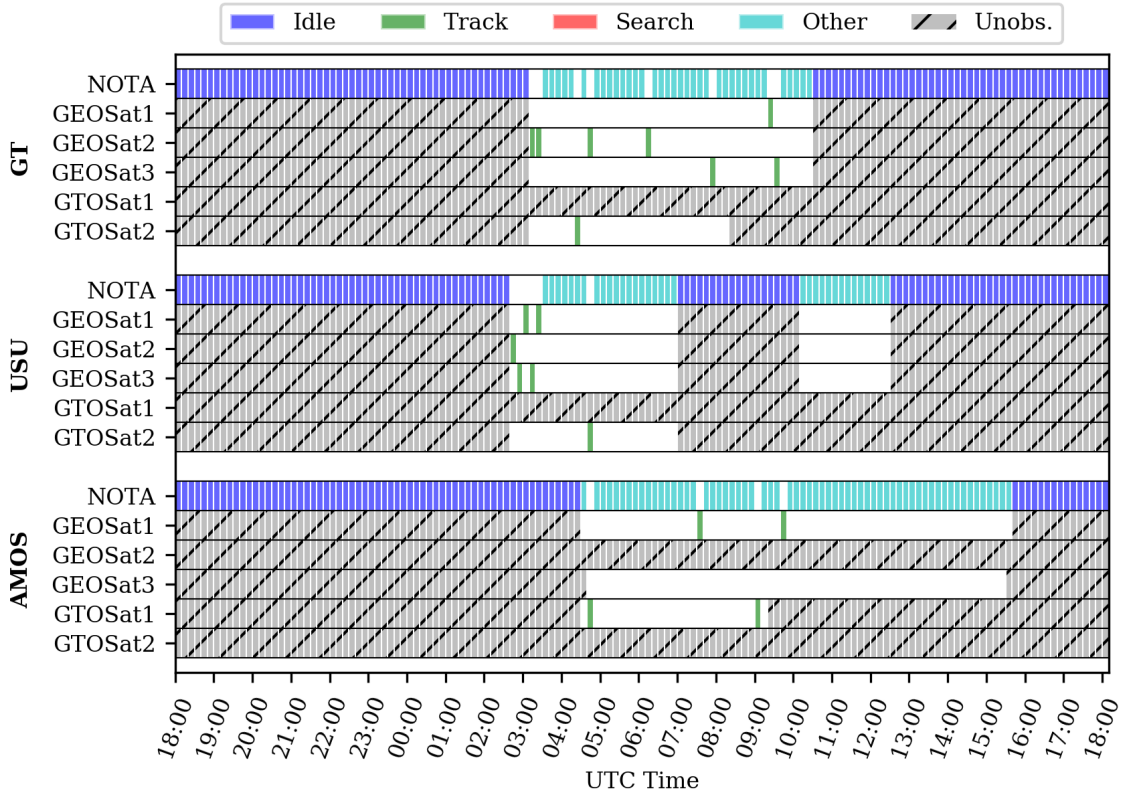
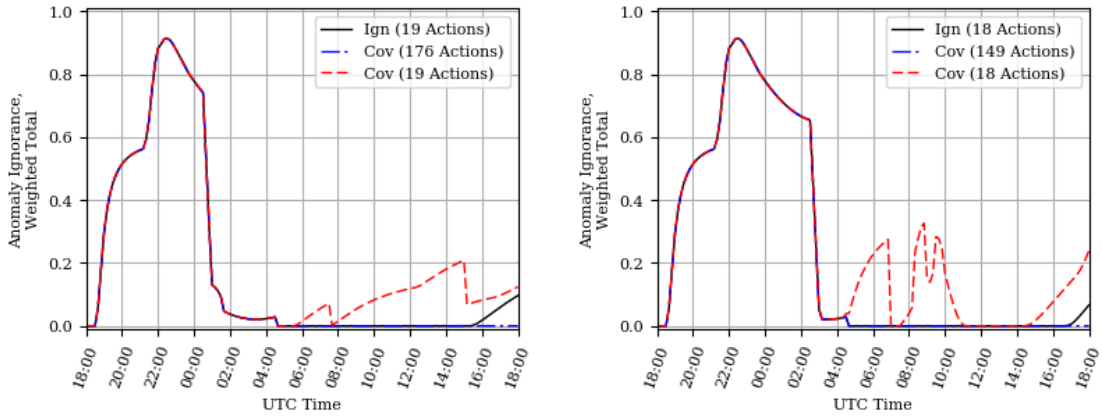


Figure 4.20: Case 4: Covariance-minimization (limited) sensor tasking schedule and anomaly ignorance



(a) Case 1: Nominal, clear

(b) Case 4: Anomalous, unclear

Figure 4.21: Anomaly hypothesis ignorance comparison for ignorance-reduction and covariance-minimization methods.

CHAPTER 5
EVIDENCE GATHERING FOR HYPOTHESIS RESOLUTION USING JUDICIAL EVIDENTIAL REASONING

Frequently, decision-making occurs in environments where there is insufficient time to gather all available evidence before a decision must be rendered, requiring efficient processes for prioritizing between candidate action sequences. As with the sensor tasking problem, evidence-gathering is a high-dimensional, multi-objective, mixed-integer, nonlinear optimization problem; as such, many approaches focus on tractable sub-problems (e.g. single objectives, limited targets, limited actors).

In addition to the concern of ambiguous evidence, another concern in evidence-gathering is confirmation bias: a preferential tendency to gather evidence that confirms prior beliefs [91]. In regimes with uncertainty and ambiguity, this effect also applies by interpreting ambiguous evidence in favor of prior beliefs. Appropriate hypothesis resolution should efficiently and conclusively confirm or refute each proposition while avoiding fixation based on prior information, which may be plagued with uncertainty or ambiguity. Studies have shown that decision-makers demonstrate several biases in decisions involving probability judgment [91, 92, 93, 94], so decision-makers must be confident that the evidence-based hypothesis resolution is strong and impartial before declaring a resolution.

This chapter generalizes the application of evidential reasoning from the previous chapter to a broader evidence-gathering framework for multi-hypothesis resolution, allowing quantified ambiguity representation. The generalized framework utilizes Dempster-Shafer theory to address the well-documented decision-making phenomenon of ambiguity aversion. Adversarial optimization techniques are applied to mitigate confirmation bias and improve computational tractability of the multi-hypothesis resolution problem. The developed framework, called Judicial Evidential Reasoning, is demonstrated in several simpli-

fied example cases as well as a more nuanced application to SSA sensor tasking.

5.1 Ambiguity Aversion

It has been shown that human decision-makers overwhelmingly prefer known risks to unknown risks, making ambiguity a major concern in modeling knowledge states [94]. Ellsberg's paradox, re-stated here, is a well-known example that violates Savage's theory of subjective expected utility [94]. Consider two urns, each filled with 100 red or yellow balls. The first urn contains an unknown distribution of red and yellow balls. The second urn contains an equal distribution of red and yellow balls, 50 of each. The goal is to draw a red ball from one of the urns, and the human decision-maker is allowed to choose which urn they draw from. The results of Ellsberg's study show that humans overwhelmingly chose to draw from the second urn, which has a known probability distribution, even though the first urn may contain a favorable distribution of red balls. This is a phenomenon known as "ambiguity aversion" and is a predictable characteristic of human decision-making in the face of uncertainty.

The first urn in Ellsberg's paradox represents a vacuous knowledge state, while the second urn represents the equal-probability knowledge state. Using Bayesian probability, both knowledge states would be represented with the same probability mass function, meaning the information presented to the decision-maker would not adequately convey information on the presence or lack of ambiguity that would impact the decision. This highlights a deficiency in Bayesian probability theory that has a significant impact in human decision-making contexts, which motivates the use of alternative methodologies such as evidential reasoning.

5.1.1 Plausibility Transformation for Decision-Making

While the ability to represent ambiguity in belief functions is useful for accurately representing knowledge states, a key criticism is that the theory of belief functions lacks a

coherent decision theory [55]. Multiple methods exist for translating between Dempster-Shafer belief functions and probability models, allowing the use of Bayesian decision theory. Smets suggested the use of the pignistic transformation [57], but it has been argued that the pignistic transformation may not be consistent with Dempster's rule of combination [55]. An alternative method, the plausibility transformation, is defined in Eqn (5.1) [55]:

$$\mathrm{Pr}_{\mathrm{pl_m}}(x) = K^{-1}\mathrm{Pl}_m(\{x\}), \tag{5.1}$$

$$\text{where } K = \sum_{x\in\Omega} \mathrm{Pl}_m(\{x\}) \tag{5.2}$$

Note that the normalization constant K in (5.2) is different from the normalization constant for Dempster's conjunctive rule in Eqn. (4.6). The plausibility transformation is consistent with Dempster's rule, particularly in situations where pignistic probability is inconsistent [55].

5.1.2 Entropy for Decision-Making

Another important concept in both probabilistic and evidential reasoning is entropy as an information content measure. For Dempster-Shafer theory, multiple definitions of entropy have been proposed, many of which are summarized by Jirousek and Shenoy [74]. Conflict in the belief structure is measured through Shannon entropy using the plausibility transform, where low conflict means a significant belief mass attributed to a singleton proposition. Non-specificity captures ambiguity as the entropy associated with non-singleton focal sets of the bpa using the Dubois-Prade entropy. The Jirousek-Shenoy (J-S) definition of entropy combines Shannon and Dubois-Prade entropy to capture both conflict and non-specificity. Minimizing both conflict and non-specificity ensures that the resulting belief structure is internally consistent (i.e. prefers strong hypothesis resolution over an equally-probable result) and is non-ambiguous.

Table 5.1: Ellsberg's paradox belief structures and entropy

Urn	$m(\{\text{red}\})$	$m(\{\text{yellow}\})$	$m(\{\text{red, yellow}\})$	$H_S(m)$	$H_{DP}(m)$	$H_{JS}(m)$
1	0.5	0.5	0	1	0	1
2	0	0	1	1	1	2

One useful property of J-S entropy is that maximum entropy is only attained by a vacuous bpa, which is the bpa where all belief mass is assigned to the entire frame: $m(\Omega) = 1$. Including both conflict and non-specificity (or ambiguity) in the entropy calculation allows for appropriate modeling of the ambiguity aversion phenomenon [74]. Recalling Ellsberg's paradox, the first urn is an equally-likely belief structure and the second urn is a vacuous belief structure:

$$m_1(\{\text{red}\}) = m_1(\{\text{yellow}\}) = 0.5\,, \qquad m_1(\{\text{red, yellow}\}) = 0$$
$$m_2(\{\text{red}\}) = m_2(\{\text{yellow}\}) = 0\,, \qquad m_2(\{\text{red, yellow}\}) = 1$$

The Shannon entropy, Dubois-Prade entropy, and J-S entropy for these belief structures are shown in Table 5.1. As expected, Shannon entropy shows high conflict for both belief structures, but Dubois-Prade entropy is only non-zero for the ambiguous distribution, so the second urn has a higher J-S entropy. The decision-maker wants to minimize conflict and non-specificity, so selecting urn 1 with the lower J-S entropy is consistent with the result from Ellsberg's paradox. Therefore, minimizing J-S entropy can be used as a reliable and consistent metric for a strong hypothesis resolution.

5.2 Hypothesis Abstraction

Many evidence-gathering approaches (e.g. sensor network tasking) operate on maintaining a low overall uncertainty (e.g. information-maximum); however, it may not be readily apparent to a decision-maker how reducing state uncertainty by a certain amount affects situation awareness or answers decision-making questions. This motivates an approach

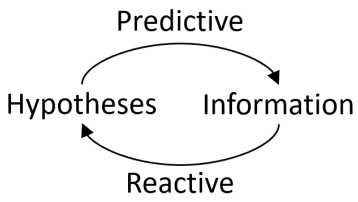


Figure 5.1: Predictive and reactive evidence-gathering.

that encodes decision-making priorities as hypotheses that can be interrogated by evidence-gathering actions.

Hypothesis-driven approaches enable a predictive mode of evidence-gathering designed to answer specific questions, using prior knowledge of relevant hypotheses to estimate information-gain from potential courses of action and propose actions that are predicted to resolve the hypotheses. This is fundamentally different from reactive approaches, where the gathered information is used to form hypotheses *a posteriori* about what caused the observed behavior. This relationship between hypotheses and information in predictive and reactive evidence-gathering is illustrated in Fig. 5.1.

Re-framing evidence-gathering in terms of hypotheses supports human decision-making strengths in abstract-level cognitive tasks required for objective prioritization and goal-adjustment [11]. Forcing an operator to switch between different levels of the abstraction, effectively approaching the problem at multiple different levels of detail, leads to increased frustration and workload and decreased situation awareness [82]. Designing a decision-support system that directly conveys hypothesis resolution information ensures that the human decision-maker spends more time on strategic cognitive tasks.

5.3 General Evidence-Gathering Problem Definition

Consider a set of hypotheses and a set of actors tasked with gathering evidence to resolve these hypotheses over a given T-step time horizon, from t_k to t_{k+T}. The finite set of hypotheses under consideration can be represented as $\Omega = \{\Omega_1, \ldots, \Omega_n\}$, where Ω_i is the frame of discernment for the i^{th} hypothesis and $|\Omega| = n \in \mathbb{Z}^+$ is the number of hypothe-

ses. Recall that each hypothesis frame of discernment $\Omega_i \in \Omega$ contains a set of mutually exclusive and collectively exhaustive propositions for resolution of that hypothesis.

At time t_k, define the actions available to the s^{th} actor as the finite set $\mathbb{A}_{s,k}$. Under the assumption that each actor can only perform one action at a given time t_k, the available action sets for all m actors in the network at time t_k are described through the Cartesian product:

$$\mathbb{A}_k = \mathbb{A}_{1,k} \times \ldots \times \mathbb{A}_{m,k} \tag{5.3}$$

where an action set $\mathcal{A}_k \in \mathbb{A}_k$ denotes a valid set of m actions at time t_k and $A_{s,k} \in \mathcal{A}_k$ is a valid action for actor s from that action set.

Define an actor's sequence of actions over the time horizon t_{k+1} to t_{k+T} as the following ordered list (or T-tuple):

$$\mathcal{A}_{s,1:T} = (A_{s,1}, \ldots, A_{s,T}) \ , \quad s = 1, \ldots, m \tag{5.4}$$

Similarly, define a set of action sequences for all actors as the finite set:

$$\mathcal{A}_{1:T} = \{\mathcal{A}_{1,1:T}, \ldots, \mathcal{A}_{m,1:T}\} \tag{5.5}$$

This set contains an action sequence for each of the m actors in the network and thus fully defines the actions taken by the network over the time horizon. Furthermore, the set of all valid sets of action sequences (all valid combinations of action sequences) is also represented by a Cartesian product:

$$\mathbb{A}_{1:T} = \mathcal{A}_{1:T} \times \ldots \times \mathcal{A}_{1:T} \tag{5.6}$$

The goal is to select the set of action sequences that minimizes a to-be-defined cost

function at the end of the T-step receding time horizon. Generically, this cost function may be represented as follows:

$$J_T : (\Omega, \mathcal{W}; \mathcal{A}_{1:T}) \mapsto \mathbb{R} \tag{5.7}$$

where $\mathcal{W}$ is a user-defined set of weights such that $w_i \in \mathcal{W}$ quantifies the priority of hypothesis Ω_i relative to the other hypotheses in Ω, and T indicates that the cost function is evaluated at the end of the time horizon, time t_{k+T}. It stands to reason that some hypotheses will be more important to decision-makers than others, so this weighting is considered a user-defined (potentially time-varying) parameter. It is not subject to optimization in this study but is instead treated as a tunable parameter.

Therefore, the generic hypothesis-based evidence-gathering optimization problem is:

$$\mathcal{A}^*_{1:T} = \underset{\mathcal{A}_{1:T} \in \mathbb{A}_{1:T}}{\arg\min} \; J_T \left(\Omega, \mathcal{W}; \mathcal{A}_{1:T}\right) \tag{5.8}$$

In other words, the optimal set of action sequences minimizes the cost function J_T, evaluated at time t_{k+T} subject to the evidence from each action $A_{s,\cdot} \in \mathcal{A}_{s,1:T}$ in each action sequence $\mathcal{A}_{s,1:T} \in \mathcal{A}_{1:T}$ for each actor $s = 1, \ldots, m$. In the following sections, a specific cost function is developed based on reaching strong (unambiguous and unbiased) hypothesis resolutions.

5.4 Evidence-Gathering for Hypothesis Entropy Reduction

Hypothesis resolution refers to the goal of determining which proposition is true from the set of propositions in the frame of discernment. Recall that Jirousek-Shenoy entropy [74] quantifies both conflict and non-specificity in hypothesis knowledge, providing an apt minimization objective for strong hypothesis resolution.

At a given time t_k, each candidate action $A \in \mathcal{A}_k$ gathers evidence that may be used to resolve hypotheses. Denote the total amount of evidence gathered through action set

$\mathcal{A}_k$ as p, noting that a single action may gather multiple distinct pieces of evidence or may gather no evidence, restricting p to the non-negative integers. The hypothesis-resolution contribution of a given piece of evidence is represented by the bpa:

$$m_{i,j,k} : 2^{\Omega_i} \mapsto [0, 1] \tag{5.9}$$

where the subscript i indicates that this bpa is related to hypothesis Ω_i, the subscript $j = 1, \ldots, p$ refers to the piece of evidence relevant to this bpa, and the subscript k indicates the evidence is gathered at time t_k. The bpas for all p pieces of evidence can be fused using Dempster's rule to arrive at a hypothesis update bpa:

$$\tilde{m}_{i,k} = \bigoplus_{j=1}^{p} m_{i,j,k} \tag{5.10}$$

Recall that Dempster's rule is associative and commutative, meaning the combination can be done sequentially and order doesn't matter [95]. However, Dempster's rule is not idempotent so the pieces of evidence being combined must be independent to avoid artificially inflating the effect of a particular piece of evidence. If a particular piece of evidence j gathered at time t_k does not contribute to hypothesis Ω_i, then $m_{i,j,k}$ is simply the vacuous bpa, ensuring that each term in the summation is defined.

Therefore, the resulting knowledge state for hypothesis Ω_i, incorporating all evidence from time from t_0 to t_k is denoted as:

$$m_{i,k}^{+} = m_{i,k}^{-} \oplus \tilde{m}_{i,k} \tag{5.11}$$

where $m_{i,k}^{+}$ is the a posteriori knowledge state and $m_{i,k}^{-}$ is the a priori knowledge state for hypothesis Ω_i at time t_k based on all evidence gathered prior to t_k.

5.4.1 Normalized Jirousek-Shenoy Entropy

The resolution of hypothesis Ω_i based on bpa m_i, as measured through Jirousek-Shenoy entropy [74], is defined as:

$$H_{JS}\left(m_i\right) = \left(\sum_{x \in \Omega_i} \mathrm{Pl}_{\mathrm{P_{m_i}}}\left(x\right) \log_2 \left(\frac{1}{\mathrm{Pl}_{\mathrm{P_{m_i}}}}\right)\right) + \left(\sum_{A \in 2^{\Omega_i}} m_i(A) \log_2\left(|A|\right)\right) \tag{5.12}$$

where the first summation term, related to Shannon entropy, quantifies conflict and the second summation term, called Dubois-Prade entropy, quantifies non-specificity. Theorem 5.4.1 shows that Jirousek-Shenoy entropy is on the scale $[0, 2\log_2\left(|\Omega_i|\right)]$ [74].

Theorem 5.4.1 (Maximum Entropy)**.** *Consider a bpa m for discrete random variable X with frame of discernment $\Omega_X = \{x_1, \ldots, x_n\}$. The maximum value of Jirousek-Shenoy entropy for m is*

$$H_{JS}(m) = 2\log_2\left(n\right) \tag{5.13}$$

Proof. The maximum entropy principle [74] states that the maximum value of entropy is attained by the vacuous bpa. Therefore, assume m is vacuous:

$$m(\Omega_x) = 1,\ m(x) = 0 \,\forall\, x \in 2^{\Omega_X} \setminus \Omega_X \tag{5.14}$$

The plausibility transformation yields the following plausibility probability function:

$$\mathrm{Pl}_{\mathrm{P_m}}\left(x\right) = \frac{1}{n} \,\forall\, x \in \Omega_X \tag{5.15}$$

Computing the Shannon entropy:

$$\sum_{x \in \Omega_i} \mathrm{Pl}_{\mathrm{P_{m_i}}}\left(x\right) \log_2 \left(\frac{1}{\mathrm{Pl}_{\mathrm{P_{m_i}}}}\right) = n\left(\frac{1}{n}\log_2(n)\right) = \log_2(n) \tag{5.16}$$

Similarly, computing the Dubois-Prade entropy for the vacuous bpa m:

$$\sum_{A \in 2^{\Omega_i}} m_i(A) \log_2 (|A|) = \log_2(n) \tag{5.17}$$

because $m(A)$ is non-zero only for $A = \Omega_X$, in which case the cardinality $|A| = |\Omega_X| = n$. Maximum Jirousek-Shenoy entropy is therefore the sum: $H_{JS}(m) = 2\log_2 (n)$. □

To use entropy as a cost function while accounting for hypotheses with different numbers of propositions, the normalized Jirousek-Shenoy entropy is defined as follows:

$$\tilde{H}_{JS}(m_i) = \frac{H_{JS}(m_i)}{2\log_2 (|\Omega_i|)} \tag{5.18}$$

where m_i is the bpa representing knowledge of hypothesis Ω_i.

5.4.2 Optimization Formulation

To accomplish the goal of minimizing hypothesis conflict and non-specificity, the normalized entropy defined in Eqn. (5.18) is employed as the cost function to further specify the optimization problem in Eqn. (5.8).

$$\mathcal{A}^*_{1:T} = \underset{\mathcal{A}_{1:T} \in \mathbb{A}_{1:T}}{\arg\min} \sum_{i=1}^{|\Omega|} w_i \tilde{H}_{JS} (\hat{m}_{i,T}) \tag{5.19}$$

where $w_i \in \mathcal{W}$ are the hypothesis weights used to denote relative priorities such that $\sum_i w_i = 1$, and $\hat{m}_{i,T}$ is the estimated bpa for hypothesis Ω_i at the end of the time horizon t_{k+T}. The optimal set of action sequences, $\mathcal{A}^*_{1:T}$, are actions estimated to gather evidence that minimizes conflict and non-specificity in user-prioritized hypotheses.

5.4.3 Computational Complexity

The general formulation in Eqn. (5.19) suffers from a number of practical issues in implementation. Most notably, the number of action sequences to evaluate can quickly preclude

brute-force evaluation of all possible action sequences over the time horizon. Computational complexity of a brute-force approach to this optimization problem scales with the number of hypotheses $|\Omega|$, the number of sensors m, the number of valid actions for each sensor n_m, and the time horizon T:

$$\mathcal{O}\left(\prod_{t=1}^{T}\left(\prod_{j=1}^{m} n_{s,t}\right)\right) \tag{5.20}$$

where m is the number of actors and $n_{s,t}$ is the number of valid actions for the s^{th} actor at time t_{k+t}. The upper bound on this complexity is found by defining the worst-case number of valid actions as $n = \max\left(n_{s,t} \,|\, s = 1, \ldots, m\,;\, t = 1, \ldots, T\right)$, yielding a worst-case computational complexity:

$$\mathcal{O}\left(\left(n^m\right)^T\right) \tag{5.21}$$

As expected, computational complexity for a brute-force approach scales exponentially with the number of valid actions, the number of actors, and the length of the time horizon. Depending on the resources required to estimate the hypothesis resolution after a set of action sequences, this algorithm can become computationally restrictive, motivating several complexity mitigations.

5.5 Implementation Considerations

This section modifies the general optimization approach in Eqn. (5.19) to arrive at a computationally tractable solution by decomposing the problem into individual sub-problems and applying adversarial optimization techniques to reduce the number of action sequence evaluations. An additional concern with the entropy-reduction algorithm is the effect of evidence ambiguity and evidence-gathering bias induced by prior information. Adversarial optimization is applied to reduce the number of action sequence evaluations and combat

confirmation bias.

5.5.1 Unbiased Hypothesis Resolution

Confirmation bias is a cognitive phenomenon where prior belief causes fixation on a particular proposition, causing the human to favor evidence that confirms prior beliefs and overlook conflicting evidence [91]. In regimes with uncertainty and ambiguity, this effect also applies by interpreting ambiguous evidence in favor of prior beliefs. Similar to human cognitive fixation, socio-technical systems might also exhibit confirmation bias. For instance, a most-probable-first evidence-gathering approach would prioritize actions estimated to gather further evidence to confirm prior knowledge. However, spurious detections or false alarms may lead to increased belief in the incorrect proposition. In this way, prior information has the potential to skew future evidence-gathering actions, so technological fixation may be induced by measurement noise, sensor bias, or other sources of uncertainty.

For illustration, consider a binary frame $\Omega = \{x_1, x_2\}$ and a prior that places slight belief in the x_1 proposition: $m(\{x_1\}) = 0.1$, $m(\{x_2\}) = 0$, $m(\{x_1, x_2\}) = 0.9$. A most-probable-first approach would focus future actions on confirming $\{x_1\}$, while ignoring the (much larger) ignorance in the estimated proposition. If the true resolution of this hypothesis is actually $\{x_2\}$, evidence gathered from tasking on the incorrect proposition ($\{x_1\}$) may be vacuous, causing the knowledge state stagnate. In this case, the most-probable-first approach stalls as no further evidence is admitted to increase belief in $\{x_2\}$ and change the proposed tasking.

It is important to avoid fixating on any particular proposition where incorrect priors or evidence ambiguity may be the cause of any bias, adding a competing objective to the requirement of minimizing hypothesis entropy. Just as fixation should not be ignored in favor of time optimality, fixation should not be the only focus at the cost of resolving hypotheses within time constraints. Quantifying confirmation bias is an active area of research, with cognitive sciences researchers using various measures comparing selection of sup-

porting versus refuting evidence [91, 93]. One such measurement is the difference between numbers of selected supporting and refuting evidence elements [92], meaning an unbiased sequence of actions selects equal numbers of supporting and refuting elements.

The proposed approach employs a related heuristic, a principle of equal effort that distributes resources (e.g. actions, time, money) evenly amongst propositions. An apt analogy for this heuristic is the fair trial system, wherein the defense and prosecution are given equal opportunity to present the strongest evidence to confirm or refute a hypothesis. Similarly, the proposed framework employs a pair of agents for each proposition, advocate and critic, which alternate action turns to allow equal opportunity for gathering supporting or refuting evidence, respectively. Due to strong parallels to the fair trial system, the proposed framework is called Judicial Evidential Reasoning (JER).

Application of this alternating-turns heuristic encourages resolution guided by evidence, not prior beliefs, biases, or ambiguity. In the event of multiple competing resources, the principle of equal effort creates an additional multi-objective optimization and uniqueness of the solution using this heuristic is not guaranteed. However, improved measures for confirmation bias are an area for future research and could extend the JER approach by altering the agent-pair action ordering.

5.5.2 Sub-Problem Definition

The primary intuition that allows decomposition of the entropy-reduction approach in Eqn. (5.19) is that not all sensor actions contribute evidence related to all hypotheses. The sub-problems can be solved independently (and in parallel), resulting in $|\Omega|$ sub-problem action sequence sets that must be combined into a single optimal set of action sequences.

Consider one of the hypotheses $\Omega_i \in \Omega$ and the subset of valid actions relevant to that hypothesis as $\mathbb{A}_{s,k,i}$:

$$\mathbb{A}_{s,k,i} \subseteq \mathbb{A}_{s,k} \tag{5.22}$$

where $\mathbb{A}_{s,k}$ are all the valid actions for sensor $s = 1, \ldots, m$. Similarly, the action sequences relevant to hypothesis Ω_i over the time horizon t_k to t_{k+T_i} are denoted

$$\mathbb{A}_{1:T_i,i} \subseteq \mathbb{A}_{1:T_i} \tag{5.23}$$

By definition, $|\mathbb{A}_{s,k,i}| \leq |\mathbb{A}_{s,k}|$ and $|\mathbb{A}_{1:T_i,i}| \leq |\mathbb{A}_{1:T_i}|$. Note that the time horizon T_i is allowed to be different for each hypothesis since, in operation, not all hypotheses need to have the same optimization horizon.

The sub-problem optimization objective is first represented using a generic cost function specific to each hypothesis:

$$J_{T_i,i} : (\Omega_i; \mathcal{A}_{1:T_i,i}) \mapsto \mathbb{R} \tag{5.24}$$

Note that w_i is not relevant to this portion of the optimization as the sub-problems are being solved independently, but it will play a role in the combination of the sub-problem sequences. The sub-problem optimization problem is defined as:

$$\mathcal{A}^*_{1:T_i,i} = \underset{\mathcal{A}_{1:T_i,i} \in \mathbb{A}_{1:T_i,i}}{\arg\min} \; J_{T_i,i} \left(\Omega_i; \mathcal{A}_{1:T_i,i}\right) \tag{5.25}$$

This sub-problem decomposition approach allows for parallel computation of action sequence sets for each agent-pair. However, if the entropy-reduction cost function is employed as in Eqn. (5.19), the same concerns related to confirmation bias will arise: an incorrect prior induces actions against the incorrect proposition, leading to weak or vacuous evidence and weak hypothesis resolution. Therefore, a different optimization approach is employed for the sub-problems while entropy-minimization is reserved for the combination of the sub-problem solutions.

5.5.3 Combating Confirmation Bias

Adversarial optimization techniques are employed to reduce confirmation bias, similar to the opposing counsel in the judicial system. Approaches such as minimax optimization have been heavily applied in game theory for turn-based, zero-sum games such as Chess and GO. In minimax optimization, an agent plans its actions with the knowledge that the opposing agent will select actions toward the opposite goal. In light of this conflict, both agents attempt to minimize potential loss in a worst-case scenario. Conversely, for a maximizing objective, maximin optimization represents agents maximizing the minimum gain from a sequence of actions.

Consider a single hypothesis from the set of considered hypotheses at time t_k: $\Omega_i \in \Omega$. Each proposition must be either conclusively confirmed or refuted with evidence, so each proposition is assigned a pair of JER agents. Therefore, for hypothesis Ω_i there are $|\Omega_i|$ alternating JER agent-pairs. When the advocate agent is active, its goal is to maximize belief in the proposition $\{\theta\}$, accomplished using maximin optimization with the plausibility probability transformation:

$$\mathcal{A}^*_{1:T_i|\{\theta\}} = \arg \max_{\mathcal{A}_{1:T_i,i} \in \mathbb{A}_{1:T_i,i}} \min \quad \mathrm{Pr}_{\mathrm{pl}}\left(\theta; m_{i|\mathcal{A}_{1:T_i,i}}\right) \tag{5.26}$$

where $m_{i|\mathcal{A}_{1:T_i,i}}$ is the estimated bpa resulting from the proposed action sequence $\mathcal{A}_{1,H}$. The plausibility transformation is applied here because of its relationship and consistency with decision-making. The maximum attainable value for this objective is 1 when proposition $\{\theta\}$ has full belief, and the minimum attainable value for this objective is 0 when proposition $\{\neg\theta\}$ has full belief. When the critic agent is active, its goal is to maximize belief in the alternative proposition ($\{\neg\theta\}$) or equivalently minimize belief in the null proposition ($\{\theta\}$). Therefore, the formulation simply flips to a minimax optimization:

$$\mathcal{A}^*_{1:T_i|\{\neg\theta\}} = \arg \min_{\mathcal{A}_{1:T_i,i} \in \mathbb{A}_{1:T_i,i}} \max \quad \mathrm{Pr}_{\mathrm{pl}}\left(\theta; m_{i|\mathcal{A}_{1:T_i,i}}\right) \tag{5.27}$$

The result of the JER agent-pair schedule optimization is a minimax-optimal action sequence for each agent-pair. In the next section, these sub-problem action sequences are combined to arrive at a single optimal schedule. If an agent-pair's action is selected in the final schedule for this iteration, that agent-pair flips its active agent for the next time step.

5.5.4 Resolving Combined Schedule Incongruity

After determining optimal schedules for each agent-pair, the schedules must be combined into a single schedule. Depending on the hypotheses, it is possible or even likely that two or more agent-pairs will require the same actor for different actions. These incongruities are resolved by choosing the actions that lead to the strongest hypothesis resolution as measured by entropy.

Using the set of actions from all sub-problem optimal sequences $\mathbb{A}^*_{1:T,\cdot}$, all possible combinations of these actions are used to form candidate congruous action sequences. The combination schedules are evaluated up to the longest time horizon. The evaluation criterion for selecting the optimal combined schedule is the weighted-sum of entropy:

$$\mathcal{A}^*_{1,T} = \underset{\mathcal{A}_{1:T} \in \mathbb{A}^*_{1:T,\cdot}}{\arg\min} \sum_{i=1}^{|\Omega|} w_i \tilde{H}_{JS}\left(m_{i|\mathcal{A}_{1:T}}\right) \tag{5.28}$$

where w_i is the weighting for the i^{th} hypothesis, and $\tilde{H}_{JS}$ is the normalized J-S entropy as defined in Eqn. (5.18). Since Jirousek-Shenoy entropy quantifies both conflict and non-specificity, and the weighting parameters encode decision-maker priorities, the resulting action sequence $\mathcal{A}^*_{1:T}$ is the action sequence with the strongest priority-weighted resolution.

At worst case, this is the same as a brute-force re-evaluation, but this would require all hypotheses to have the same applicable action subsets and all possible actions produce an optimal result for at least one hypothesis. This implies an extreme interdependence between the hypotheses that is unlikely to occur in operation. In more realistic cases, where at least some hypotheses are distinct enough to have different applicable actions, this re-evaluation

is much less computationally complex than brute-force.

5.5.5 Efficient Minimax Optimization

To further reduce the number of action sequences evaluated, the alternating-agent formulation of the sub-problems can be further exploited using adversarial optimization techniques. Combinatorial optimization techniques often employ methods for intelligently exploring or pruning expansive decision trees (e.g. branch and bound) to quickly eliminate costly or infeasible options.

In naive minimax (or maximin) optimization, the number of sequences evaluated grows exponentially with the number of valid actions and the search depth, as in Eqns. (5.20) and (5.21). However, depending on the order in which the tree of action sequences is traversed, some sequences do not need to be evaluated if they are known never to lead to the optimal solution. A popular technique to accomplish this is called alpha-beta pruning [96]. Determined by previously evaluated sequences, alpha represents the minimum score that the maximizing player is already guaranteed, while beta represents the maximum score that the minimizing player is guaranteed. These values function as thresholds to prune branches of the search tree that cannot possibly result in the optimal sequence.

The effect of pruning the known sub-optimal branches early is to reduce the number of required sequence evaluations while still arriving at the same optimal solution as naive minimax. In an ideal case, the computational complexity reduces to Eqn. (5.29), a significant improvement over the brute-force complexity in Eqn. (5.21).

$$\mathcal{O}\left(\sqrt{(n^m)^T}\right) \tag{5.29}$$

While this idealized complexity may not be fully realized in application, alpha-beta pruning is still likely to eliminated unnecessary searches to provide a more efficient minimax search.

Figure 5.2 illustrates the internal logic of an alpha-beta pruning search. Each node

represents an action and a sample objective-function return resulting from that action in the minimax approach. The maximizing agent at the middle-depth recognizes that, if it were to take the second available action (from the left), the minimizing agent has an opportunity to choose an action to reach an objective function value of 3, possibly less. Since this is less than the guaranteed 7 from the maximizing agent's first action, the remainder of that branch is pruned. Similarly, the minimizing agent recognizes that, at the top-node, taking the second available action allows the maximizing agent to attain an objective value of 8, possibly greater. Therefore, the remainder of the right-side of the evaluation tree is eliminated. This reduces the number of sequence evaluations required from a maximum of 8 (in naive minimax) to just 5.

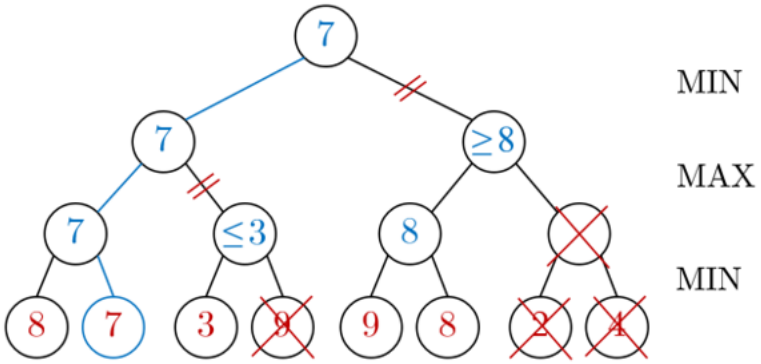


Figure 5.2: Sample alpha-beta pruning

Section 5.6.1 further illustrates the benefits of alpha-beta pruning in example case 1, specifically Fig. 5.5. The implemented JER algorithm for this paper uses alpha-beta pruning for efficient minimax optimization.

5.5.6 Hypothesis Pruning via Entropy Stopping Condition

A final computational consideration is the pruning of resolved hypotheses. Once sufficient evidence has been gathered to resolve a hypothesis, it is beneficial to remove that hypothesis from consideration for future tasking evaluations. Decision-makers should be able to indicate an acceptable level of conflict and ambiguity, manifesting as J-S entropy thresh-

olds $\tilde{H}_{th}(m_i)$ for each hypothesis Ω_i. If the entropy for a given hypothesis falls below this threshold, that hypothesis is considered adequately resolved and action sequences related to that hypothesis no longer need to be considered. This improves computational complexity further by removing entire sub-problems from consideration.

5.5.7 Judicial Evidential Reasoning Summary

The three primary considerations of the JER framework, as described in the preceding sections, are: hypothesis abstraction, ambiguity aversion, and unbiased hypothesis resolution. Employing a hypothesis abstraction enables predictive tasking and supports human cognition at a strategic and planning level. The use of evidential reasoning, specifically Dempster-Shafer theory, to model hypothesis knowledge allows quantified conflict and ambiguity together in the entropy measurement. Applying a principle of equal effort through the alternating-turns heuristic, inspired by the fair trial system, provides impartial or unbiased hypothesis resolution to guard against confirmation bias while also prioritizing time-efficient hypothesis resolution. The inclusion of efficient minimax algorithms and a hypothesis resolution pruning condition further improve computational tractability.

The JER framework developed in the previous sections is summarized graphically in Fig. 5.3. Algorithm 1 outlines the JER algorithm outer-loop process, termed the JER schedule manager. The manager starts by calling each agent-pair inner-loop process in parallel. Each agent-pair solves its sub-problem using the alternating-agent minimax optimization on the plausibility probability transformation, outlined in Alg. 2 using a naive minimax algorithm for ease of description. Recall that the alpha-beta pruning enhancement is simply an efficient minimax that reduces the search space, so both the naive and alpha-beta implementations reach the same result. The JER manager combines the sub-problem schedules and determines the optimal combined action sequence $\mathcal{A}^{**}$ using the total weighted entropy objective. Once the optimal action sequence is determined, the active agent-pairs (those whose actions are chosen in the optimal sequence) are flipped for the next iteration.

Algorithm 1 JER manager, subproblem schedule combination

t_k: current time
t_T: horizon time
w_j: weight for hypothesis j
$\mathcal{A}_i^*$: sub-problem solution for i^{th} JER agent-pair
procedure OPTIMIZEACTIONSEQUENCE
Solve sub-problem schedules
    for each agent-pair i **do**
        $m_{i,k}^- \leftarrow$ a priori bpa for relevant hypothesis of agent-pair i at time t_k
        $isMax_i \leftarrow$ flag, true if advocate agent is active for agent-pair i at time t_k
        $A_i \leftarrow []$: initialize empty sequence
        $(s_i^*\ ,\ A_i^*) \leftarrow$ **EvaluateAgentPair**$\left(t_k, t_T, \hat{m}_{i,k}^-, isMax_i, A_i\right)$
    end for
Resolve combined schedule incongruity
    $\mathbb{A}^* \leftarrow A_1^* \times \ldots \times A_N^*$: Cartesian product of subproblem sequences
    $\mathbb{A}_u^* \leftarrow$ **unique**$(\mathbb{A}^*)$: unique combination sequences
    for each sequence A_i in $\mathbb{A}_u^*$ **do**
        for each hypothesis $\mathcal{H}_j$ **do**
            $\hat{m}_j^+ \leftarrow m_j^-$: initialize updated hypothesis estimates
        end for
        for each action a in A_i **do**
            for each hypothesis $\mathcal{H}_j$ **do**
                $\hat{m}_{j,a} \leftarrow$ estimated evidence for hypothesis $\mathcal{H}_j$ from action a
                $\hat{m}_j^+ \leftarrow \hat{m}_j^+ \oplus \hat{m}_{j,a}$
            end for
        end for
        $J_i \leftarrow \sum_j w_j \tilde{H}_{JS}\left(\hat{m}_j^+\right)$: total weighted entropy objective
    end for
    $A^{**} \leftarrow$ sequence corresponding to minimum total weighted entropy
Flip active agent-pairs
    for each sub-problem sequence A_i^* **do**
        if any action a in A_i^* is in A^{**} **then**
            $isMax_i \leftarrow !isMax_i$
        end if
    end for
    return A^{**}
end procedure

Algorithm 2 JER agent-pair evaluation, recursive naive minimax implementation

t_k: current time
t_T: horizon time
$\hat{m}_k^-$: estimated a priori bpa at time t_k
$isMax$: flag, true if advocate is active at t_k
$\mathcal{A}$: action sequence
s: evaluation score, plausibility probability transformation of proposition $\{\theta\}$
procedure EVALUATEAGENTPAIR(t_k, t_T, $\hat{m}_k^-$, $isMax$, $\mathcal{A}$)
    if $t_k >= t_T$ **then**
        $\{\theta\} \leftarrow$ relevant proposition for this agent-pair
        return $(\mathrm{Pr}_{\mathrm{pl}}(\{\theta\}), \mathcal{A})$
    else
        if $isMax$ **then**
            $s^* \leftarrow -\inf$
        else
            $s^* \leftarrow +\inf$
        end if
        $\mathcal{A}^* \leftarrow \mathcal{A}$
        $\mathbb{A}_k \leftarrow$ candidate actions relevant to this agent-pair at t_k
        $t_{kp} \leftarrow$ next time step
        for each action set A in $\mathbb{A}_k$ **do**
            $\mathcal{A}[t_k] \leftarrow A$
            $\hat{m}_A \leftarrow$ estimated bpa from action set A
            $\hat{m}^+ \leftarrow \hat{m}^- \oplus \hat{m}_A$
            $(s, \mathcal{A}^+) \leftarrow$ **EvaluateAgentPair**(t_{kp}, t_T, $\hat{m}_k^+$, $!isMax$, $\mathcal{A}$)
            if $isMax$ **then**
                if $s > s^*$ **then**
                    $s^* \leftarrow s$
                    $\mathcal{A}^*[t_k : t_T] \leftarrow \mathcal{A}^+[t_k : t_T]$
                end if
            else
                if $s < s^*$ **then**
                    $s^* \leftarrow s$
                    $\mathcal{A}^*[t_k : t_T] \leftarrow \mathcal{A}^+[t_k : t_T]$
                end if
            end if
        end for
        return $(s^*, \mathcal{A}^*)$
    end if
end procedure

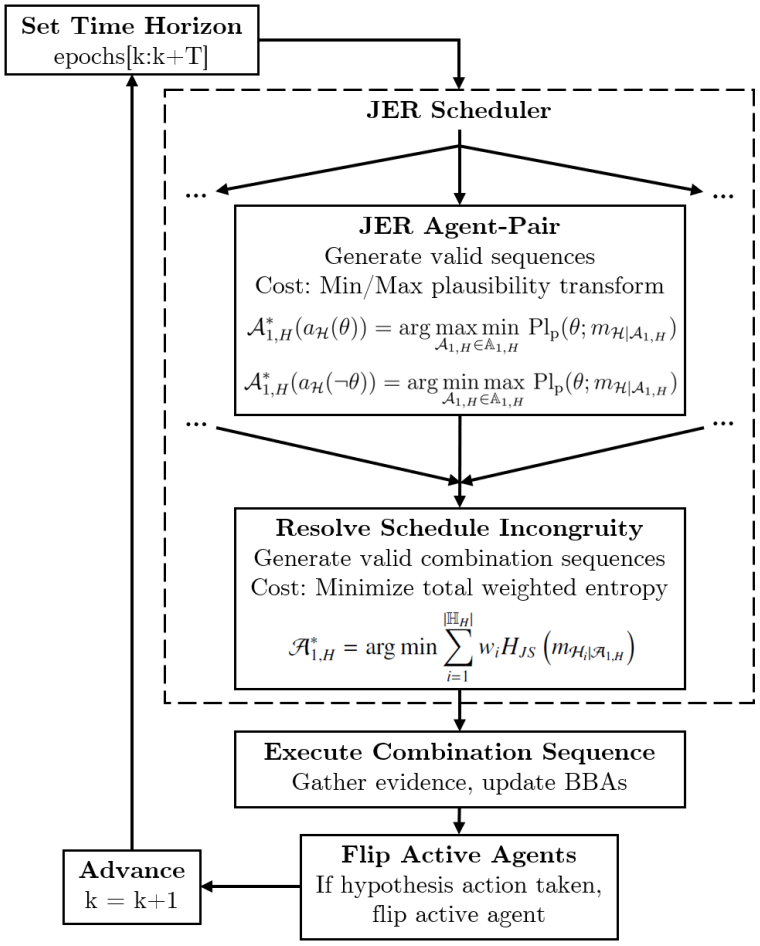


Figure 5.3: Judicial Evidential Reasoning algorithm

The following sections apply the JER algorithm described above to illustrative medical diagnosis examples as well as a sensor network tasking scenario.

5.6 Examples

This section contains illustrative examples of the JER approach using simplified medical diagnosis situations. The first example is intended to illustrate the JER agent-pair sub-problem optimization in detail, and the second example illustrates the combination of multiple JER agent-pairs to form a unified schedule. In each case, the relevant hypotheses and propositions are outlined, available tests are outlined as potential actions, and the JER approach is applied to determine the sequence of tests ordered. Since this is intended to be an illustrative example, the diagnosis details have been simplified and constraints have been enforced such that not all actions may be taken within the diagnosis window (i.e. time-critical decision-making). Section 5.7 contains a more detailed and nuanced application of JER to a real-world sensor network tasking problem.

5.6.1 Case 1: Single JER Agent-Pair

The first example involves a single hypothesis with two competing (mutually exclusive and collectively exhaustive) propositions, yielding the following frame of discernment: $\{\theta, \neg\theta\}$. To simplify notation, define the corresponding propositions as: $A = \{\theta\}$ and $\neg A = \{\neg\theta\}$. Since this is a single hypothesis problem with a binary frame of discernment, only one JER agent-pair is needed, and this example serves to illustrate the JER inner-loop: minimax optimization using plausibility probability.

Assuming no prior information on the correct resolution (i.e. full ignorance), the prior belief assignment is vacuous. Three tests are available to inform this diagnosis, but time and cost constraints limit the number of tests to two. Therefore, the goal is to determine which two tests result in a strong-but-unbiased resolution. Table 5.2 lists basic probability assignments (bpas) for each available test, functions of known statistics on the test such

as false alarm rate. Test 1 is a strong indicator of the confirmation of proposition A but does not carry information to negate A (or equivalently, confirm $\neg A$). Similarly, Test 2 is a strong indicator for confirming proposition $\neg A$. Both tests have non-zero probabilities of false alarm, meaning that neither can contribute complete belief to either proposition, resulting in non-zero belief mass attributed to the frame. Test 3 does operate as an indicator of both confirmation and negation of A, but provides weaker evidence toward both.

Table 5.2: Example Case 1: Basic probability assignments for diagnostic tests

Test #	A	$\neg A$	$A \cup \neg A$
1	0.7	0.0	0.3
2	0.0	0.7	0.3
3	0.3	0.3	0.4

Fig. 5.4 shows the tree of all possible evaluations, easily visualized due to the low dimensionality of this example. Each edge of this tree is a potential test (action), and each terminating node denotes the plausibility probability of A as a result of the two actions leading to it. For instance, the left cluster of three terminating nodes represents all possible action sequences beginning with Test 1. Traversing down the tree, each successive level alternates the active agent: supporting A or $\neg A$. This results in a two-step minimax optimization on $\mathrm{Pr}_{\mathrm{pl}}(A)$. The non-terminating nodes display the chosen node from below based on the active minimax mode at that step (max or min). Therefore, the minimizing agent (supporting $\neg A$) will select the lowest plausibility probability from each cluster of three terminating actions to populate the middle nodes, and the maximizing agent (supporting A) will select the highest plausibility probability from those three middle nodes to determine the selected action sequence (highlighted in blue in the figure).

Following test 1 with another test 1 yields an estimated bpa that may strongly indicate the diagnosis A but does not carry any unique evidence to confirm $\neg A$. Therefore, if $\neg A$ is the correct result, performing test 1 twice would not provide a strong result. Following test 1 with test 2 or test 3 results in estimated belief mass attributed to both A and $\neg A$, resulting in (at least partial) proposition confirmation regardless of the correct (true) result.

Fig. 5.4 shows that the unbiased solution in this minimax optimization scheme is Test 1 followed by Test 2, resulting in $\Pr_{\text{pl}}(A) = \Pr_{\text{pl}}(\neg A) = 0.5$. This indicates that both propositions are given equal opportunity since the prior information (vacuous) did not indicate a preference toward either proposition. This result matches intuition that, in the case of vacuous prior information, both strong indicator tests should be run to ensure the true diagnosis is confirmed.

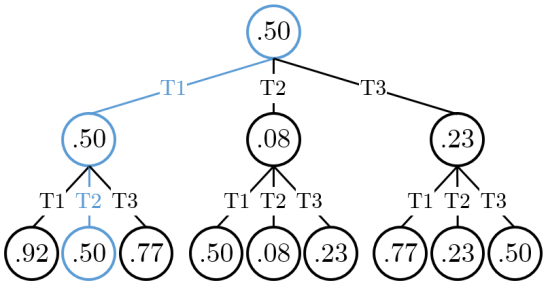


Figure 5.4: Example Case 1: Brute force evaluation tree

In Fig. 5.4, all nine possible test sequences are computed in a brute-force manner. However, this is not required as alpha-beta pruning provides a more efficient approach to minimax optimization by eliminating branches of the evaluation tree that need not be searched based on the previously-searched nodes. Figure 5.5 demonstrates this approach, reducing the number of sequence evaluations from nine to seven. As expected, alpha-beta pruning efficiently finds the same end-result as naive brute-force minimax using less evaluations.

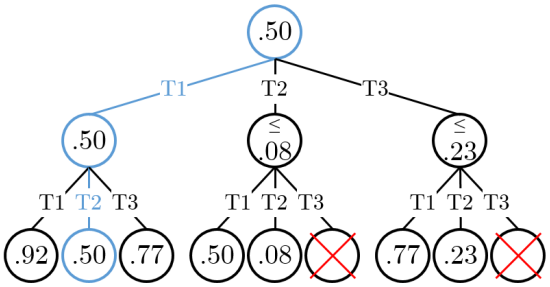


Figure 5.5: Example Case 1: Alpha-beta evaluation tree

5.6.2 Case 2: Multiple JER Agent-Pairs

The second example involves a single hypothesis with three competing (mutually exclusive and collectively exhaustive) propositions, yielding the following frame of discernment: $\{\theta_1, \theta_2, \theta_3\}$. Once again, for ease of notation, define the corresponding propositions as: $A = \{\theta_1\}$, $B = \{\theta_2\}$, $C = \{\theta_3\}$. Since this case contains a non-binary frame of discernment, three JER agent-pairs are needed (one as an advocate for each proposition). Each agent-pair solves its own minimax sub-problem (as in Case 1), and the sub-problem solutions are combined to resolve incongruity between the sub-problem schedules. Therefore, this example serves to illustrate the application of the entropy-minimization objective to resolve schedule incongruity, as well as updating the prior with test results iteratively in receding-horizon optimization.

As before, the prior belief assignment is assumed vacuous. Three tests are available to inform this diagnosis, but time and cost constraints limit the number of tests to two. Table 5.3 lists basic probability assignments (bpas) for each available test, functions of known statistics on the test such as false alarm rate. Each test has two possible outcomes, pass and fail, which affect the knowledge state based on the test's evidence and therefore affect the test bpas. For instance, test 1 is a strong indicator for proposition A, so a pass outcome gives strong belief for A whereas a fail outcome gives strong belief for $B \cup C = \neg A$. In both cases, though, there is still a non-zero chance of false alarm, so some belief is still given to the frame as ignorance. Tests 2 and 3 are similar for propositions B and C, respectively, though test 2 is only moderately strong and test 3 is even weaker.

Table 5.3: Example Case 2: Basic probability assignments for diagnostic tests

Test #	A	B	C	$A \cup B$	$A \cup C$	$B \cup C$	$A \cup B \cup C$
1 (Pass)	0.9	0.0	0.0	0.0	0.0	0.0	0.1
1 (Fail)	0.0	0.0	0.0	0.0	0.0	0.9	0.1
2 (Pass)	0.0	0.7	0.0	0.0	0.0	0.0	0.3
2 (Fail)	0.0	0.0	0.0	0.0	0.7	0.0	0.3
3 (Pass)	0.0	0.0	0.5	0.0	0.0	0.0	0.5
3 (Fail)	0.0	0.0	0.0	0.5	0.0	0.0	0.5

For each iteration, each JER agent-pair must solve its own sub-problem, selecting the two minimax-optimal tests in a similar manner to the previous example case, using the alpha-beta pruning improvement. Each agent operates under the supposition that its desired proposition is correct. In other words, the advocate agent for proposition A estimates that test 1 will return successful while tests 2 and 3 will return failed, attempting to contribute belief to proposition A. On the other hand, the critic agent for proposition A estimates the opposite test results, since it is attempting to reduce plausibility of proposition A.

After the sub-problem test schedules are optimized, all unique test sequences are found from the Cartesian product of these schedules. The estimated bpa result is computed for each under the assumption of a successful test return, and the sequences are ranked according to normalized Jirousek-Shenoy entropy. The test sequence resulting in the lowest entropy is selected, and its first test action implemented for that iteration. Finally, the agent-pairs that were active during this iteration (i.e. the agent-pairs that requested the selected action) are flipped so that, in the next iteration, the critic agent is active first.

When executing the test, the true hypothesis resolution is used to determine success or failure: in other words, if the true hypothesis resolution is assumed to be A, then test 1 would pass while tests 2 and 3 would fail. The prior bpa is updated through combination with the resulting test bpa to give the result after one iteration. The procedure above is repeated for the second iteration using this updated bpa as the prior.

Table 5.4 shows the resulting test sequence and probabilities for each realization of the hypothesis (using the plausibility transformation) after each iteration. Since the initial prior is vacuous, the chosen first test in each realization is test 1, matching intuition because test 1 is the strongest result and therefore minimizes entropy. If proposition A is true, test 1 passes and significant belief is already attributed to A. In the second iteration, test 1 is repeated because, in this case, the entropy will not be significantly decreased by the contribution of other (weaker) test results. Even though the critic agent is active for A, test 1 is still the strongest potential belief contribution for $\neg A$, providing the strongest entropy

reduction. Instead, test 1 passes again, confirming proposition A and minimizing entropy since A is the probable correct resolution. If either proposition B or C are true, test 1 fails, resulting in significant ambiguity after the first test. In both cases, the next test selected is test 2, because it is the strongest remaining test and proposition A has (nearly) been eliminated. If proposition B is true, test 2 passes and now a significant belief is attributed to B. If proposition C is true, test 2 fails as well, resulting in a less-significant but still definitive belief attributed to C: the fail result in test one indicates B or C and the fail result in test 2 indicates A or C, leaving C as the only logically consistent option. In this case, the belief attributed to A slightly increases after the second test result, but evidence still overwhelmingly indicates proposition C.

Table 5.4: Example Case 2: Results based on true hypothesis realization

Truth	**Tests**	$\mathrm{Pr}_{\mathrm{pl}}(A)$	$\mathrm{Pr}_{\mathrm{pl}}(A)$	$\mathrm{Pr}_{\mathrm{pl}}(C)$	$\tilde{H}_{JS}$
A	1 - Pass	0.84	0.08	0.08	0.308
	1 - Pass	0.98	0.01	0.01	0.055
B	1 - Fail	0.04	0.48	0.48	0.721
	2 - Pass	0.02	0.75	0.02	0.389
C	1 - Fail	0.04	0.48	0.48	0.721
	2 - Fail	0.07	0.21	0.71	0.468

Once again, this result matches intuition. Since test 3 has the weakest belief contribution and only two tests may be executed, it is never selected. The correct resolution is determined with high probability through the use of only two tests in each case.

Note that, while this example case did require more test sequence evaluations than a brute-force implementation, this is only because each available test is relevant to each agent-pair (for simplicity). Computational complexity of this evaluation scales exponentially with the number of available actions, so in a low-dimensional scenario such as this example, more available actions may not be an issue. In real-life complex decision-making scenarios with multiple propositions or hypotheses, it is likely that this computational burden can be significantly reduced by only considering relevant actions for each agent-pair.

5.6.3 Example Summary

The two example cases presented illustrate both key components of the JER approach. Case one demonstrates the inner-loop sub-problem resolution using agent-pairs and efficient minimax with the plausibility probability transformation. Case two demonstrates the outer-loop combination of sub-problem schedules and resolution of incongruities using entropy.

While both cases illustrated use a single hypothesis, multiple hypotheses do not significantly change the implementation. New agent-pairs are introduced for the new hypotheses, The only additional mechanism required is the hypothesis weighting, which is applied in the schedule incongruity resolution step of the outer-loop. The following section of simulated results illustrates this through a more realistic decision-making scenario with multiple hypotheses.

5.7 Simulation Results

This section contains a more nuanced application of JER scheduling sensor network actions to resolve multiple space situational awareness (SSA) hypotheses.

5.7.1 Scenario Description

Operators in a SSA decision-support environment receive notice from a space launch entity that a planned geostationary transfer orbit (GTO) insertion maneuver has experienced an anomaly. The anomaly is estimated to have occurred 5 minutes prior to the notification during a critical orbit-raising maneuver. The objective is to re-acquire the space object and diagnose the anomaly to regain situation awareness.

Anomalous GTO objects are particularly difficult to characterize as the range prohibits use of radar, requiring a wide state-space search using electro-optical sensors. Timely re-acquisition is critical as the spacecraft was nominally bound for Geostationary Earth Orbit

(GEO), a densely populated orbit regime with many high-value defense and telecommunications assets. The nominal transfer time from LEO to GEO is just over five hours, placing additional time-pressure on resolving the anomaly to complete conjunction analyses and alert other satellite operators. If the anomaly resulted in a GEO-intersecting trajectory, it is crucial to characterize the new orbit to inform conjunction analyses. Similarly, if the resultant trajectory remains close to low-Earth orbit (LEO), it becomes a collision risk in a densely populated orbit regime.

The entire simulation occurs over a 15 minute time span, including the 5-minute delay between the anomaly event and the beginning of the sensor tasking window. The simulation time span is limited by observation constraints (e.g. short horizon-to-horizon times in LEO, eclipse, adverse weather), placing time pressure on the hypothesis resolution. At the end of this simulation the sensor positions will prohibit gathering further evidence, so the anomaly must be characterized within 15 minutes of the event.

5.7.2 Dynamics

The nominal transfer orbit geometry is shown in Fig. 5.6. The primary spacecraft begins in a 1000 km altitude circular parking orbit Space objects are propagated using Keplerian two-body dynamics to compute lines-of-sight to sensors. The sensor network is comprised of two 3-degree field-of-view electro-optical sensors, separated by 20 degrees in longitude for geometric diversity. Observations are simulated using a radiometric model, including simulated effects for background sky irradiance and atmospheric transmittance (e.g. cloud cover, atmospheric turbulence) [15] with illumination conditions estimated using a cannon-ball model.

The sensor-tasking time span is limited by observation constraints (e.g. short horizon-to-horizon times in LEO, eclipse, adverse weather), placing a 15-minute time limit on the hypothesis resolution. The sensors may change actions each minute, and a receding time-horizon of two minutes is used.

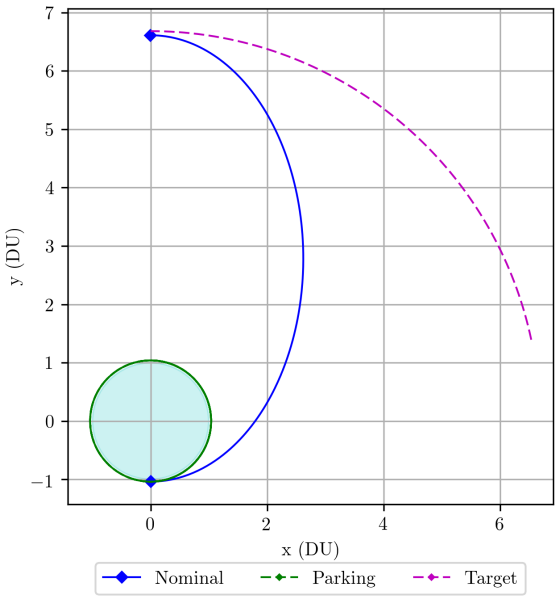


Figure 5.6: Nominal GTO transfer orbit and target GEO orbit.

5.7.3 Belief Function Models

A limited subset of potential failure modes is analyzed for illustrative purposes in this test case. As shown in Fig. 5.7, the anomaly is characterized at the subsystem level to determine root-cause. Since multiple point-of-failure events are exceedingly rare, an assumption is made that the anomaly results from a single point-of-failure, isolating the anomaly to one of these subsystems.

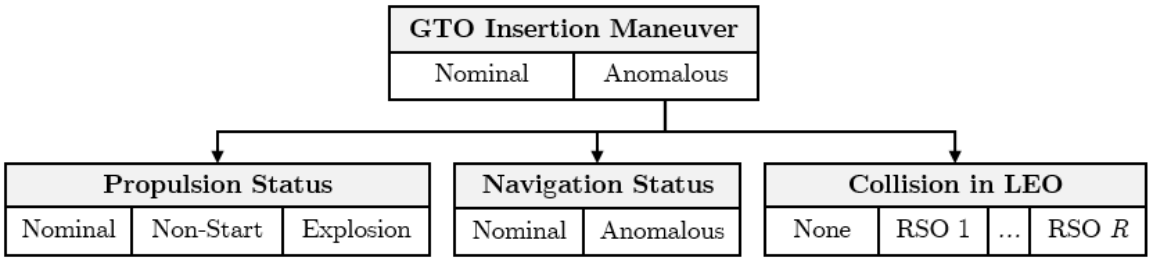


Figure 5.7: Possible causes for GTO insertion failure

The hypotheses considered for this GTO insertion maneuver anomaly include: propulsion status, navigation status, and collision in LEO. To construct JER agent-pairs, each

hypothesis is further decomposed into frames of discernment:

$$\Omega = \Omega_{man} \times \Omega_{prop} \times \Omega_{nav} \times \Omega_{coll} \tag{5.30}$$

Hypotheses are considered resolved if the normalized J-S entropy drops below the threshold value of $\tilde{H}_{JS,thr} = 0.05$.

The following sections describe each hypothesis, and the available evidence, in more detail.

Propulsion Status

The propulsion status hypothesis, Ω_{prop}, yields a three-element frame:

$$\Omega_{prop} = \{\omega_{prop,nom}, \omega_{prop,ns}, \omega_{prop,exp}\}$$

Nominal propulsion status, $\omega_{prop,nom}$, represents the case where the propulsion subsystem is not the cause of the anomaly. The non-start proposition, $\omega_{prop,ns}$, occurs when the propulsion system fails to fire, leaving the spacecraft in its LEO parking orbit. The explosion proposition, $\omega_{prop,exp}$, occurs when there is a catastrophic failure, resulting in debris in LEO near the spacecraft's parking orbit.

Navigation Status

The navigation status hypothesis, Ω_{nav}, yields a binary frame:

$$\Omega_{nav} = \{\omega_{nav,n}, \omega_{nav,a}\} \tag{5.31}$$

Nominal navigation status, $\omega_{nav,n}$, represents the case where the navigation subsystem is not the cause of the anomaly. Anomalous navigation, $\omega_{nav,a}$, results in an off-nominal transfer orbit due to pointing error, causing detection of the primary spacecraft off-track

near the nominal GTO orbit.

Collision in LEO

The collision in LEO hypothesis, Ω_{coll}, yields the following non-binary frame:

$$\Omega_{coll} = \{\omega_{coll,none}, \omega_{coll,1}, \ldots, \omega_{coll,R}\}$$

where R is the number of resident space objects (RSOs) considered. For this illustrative example, three RSOs ($R = 3$) will be considered. The "none" proposition, $\omega_{coll,none}$, represents the case where a collision has not occurred and therefore is not the cause of the anomaly. Collision with object j, $\omega_{Coll,j}$ where $j = 1, \ldots, R$, results in debris in both orbits as well as missing nominal tracks for both object j and the primary spacecraft. Recall that explosion also generates debris in the LEO parking orbit, so the missing LEO object j and debris in its orbit differentiate the hypotheses. Nominal detection of an RSO refutes that RSO's collision proposition.

JER Agent-Pairs

The full problem considers each frame described in the decomposition above to investigate the cause of a maneuver anomaly. Each frame binary frame contributes one JER agent-pair, while each non-binary frame contributes $|\Omega_\cdot|$ JER agent-pairs. Therefore, for this simulation, there are eight JER agent-pairs: three for propulsion status, one for navigation status, and four for collision in LEO.

5.7.4 Evidence to Belief Function Mappings

Each candidate action is evaluated for its estimated effect on the hypotheses to develop evidence-to-belief-function mappings. This process is highly problem-specific, requiring the modeler to consider what each potential successful or missed detection means with

respect to each hypothesis. For instance, a missed detection of the nominal GTO orbit may indicate anomaly, but if the estimated electro-optical probability of detection [15] predicted a low chance of success, that evidence is vacuous and belief mass should be attributed to ignorance instead.

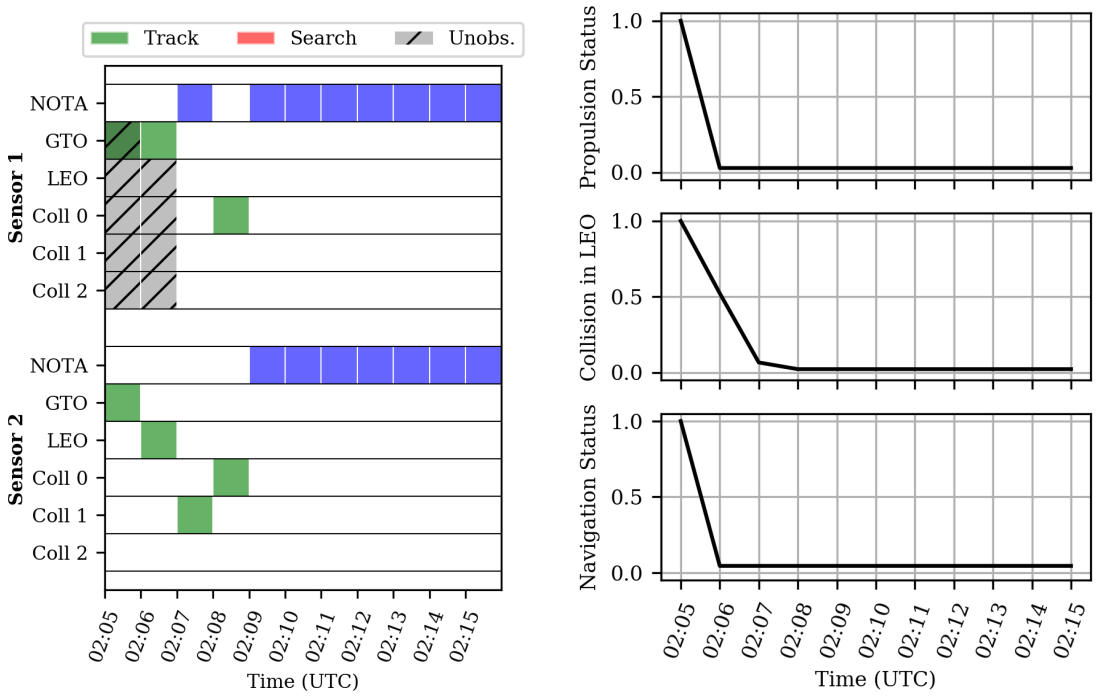
Additionally, implicit knowledge about relationships between these frames can be imposed through conditional bpas [98]. In particular, it is known that, if evidence confirms that none of the subsystems are nominal, the maneuver status is likely nominal. A small chance (0.01) is allowed that there may be other causes for maneuver anomaly even if the modeled causes are nominal to account for mis-modeling of the problem. Similarly, if any one of the other causes is anomalous, then the maneuver status is likely anomalous.

5.7.5 Case 1: Nominal Maneuver

As a baseline, the true proposition for this first test case is the nominal maneuver status. The resulting sensor tasking schedule is shown in Fig. 5.8(a), and Fig. 5.8(b) shows the normalized J-S entropy for each hypothesis. The resolutions of each proposition (belief and plausibility) are plotted in Figs. 5.9(a), 5.9(b), and 5.9(c).

The schedule in Fig. 5.8(a) indicates actions for each sensor at each time step, overlaid with target observability and tasking mode information. For sensor 1, only the GTO target is valid for the first two steps. Its first attempted observation is missed, but the radiometric model for probability of detection confirms that the observation conditions (target near the horizon) contributed to this miss.

However, sensor 2 makes several detections early in this simulation, including a task on the nominal GTO spacecraft that results in combined detection of the GTO object (supporting "nominal" propulsion and navigation statuses) and the collision objects "Coll 0" and "Coll 2". These detections result in strong resolution of the respective collision propositions (see Fig. 5.9(c) at 02:06), followed by successful detection of the "Coll 1" collision object (at 02:07). Sensor 2 also confirms the propulsion status "nominal" proposition by



(a) Sensor tasking schedule

(b) Hypothesis entropy

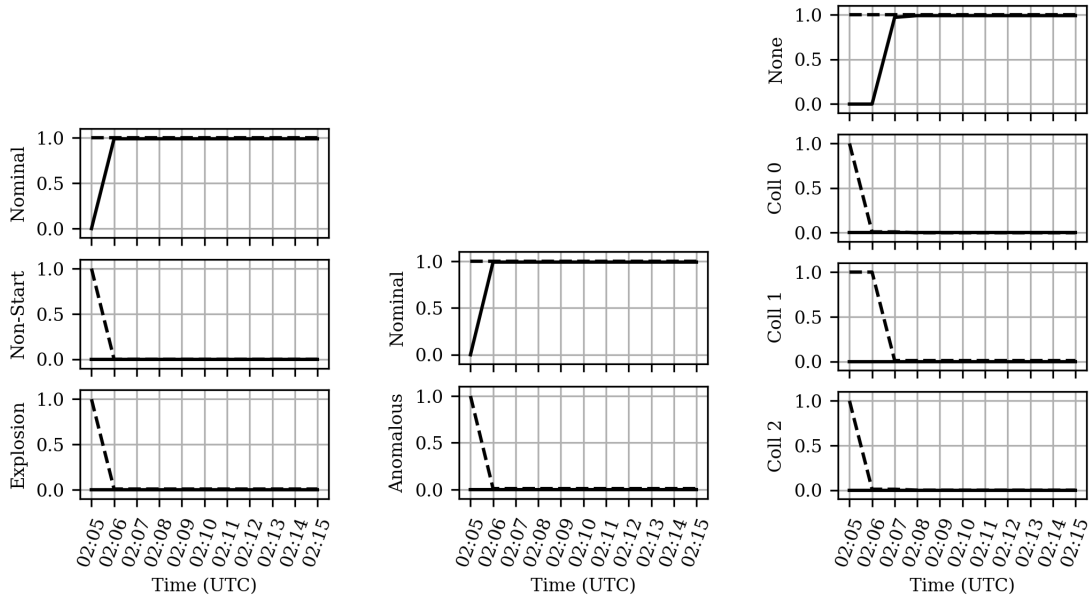
Figure 5.8: Case 1: nominal maneuver (baseline)

making a successful detection of the GTO object (at 02:05) and failing to detect the space object in its LEO parking orbit (at 02:06).

All hypotheses are resolved within the prescribed entropy tolerance within four steps (see Fig. 5.8(b)), so for the remainder of the simulation (from 02:09 onward) the sensors are free to perform other actions as necessary, as indicated by the none-of-the-above (NOTA) option and blue tasking mode. Using the sub-problem decomposition and efficient minimax search, JER only required a maximum of 271 sequence evaluations (including all agent-pairs and the combination schedule evaluations) in any iteration, far less than the theoretical brute-force maximum of 1024 evaluations.

5.7.6 Case 2: Propulsion Non-Start

In this test case, a propulsion anomaly occurs resulting in no maneuver and leaving the spacecraft in its LEO parking orbit. The resulting sensor tasking schedule and hypothesis



(a) Propulsion status (b) Navigation status (c) Collision in LEO

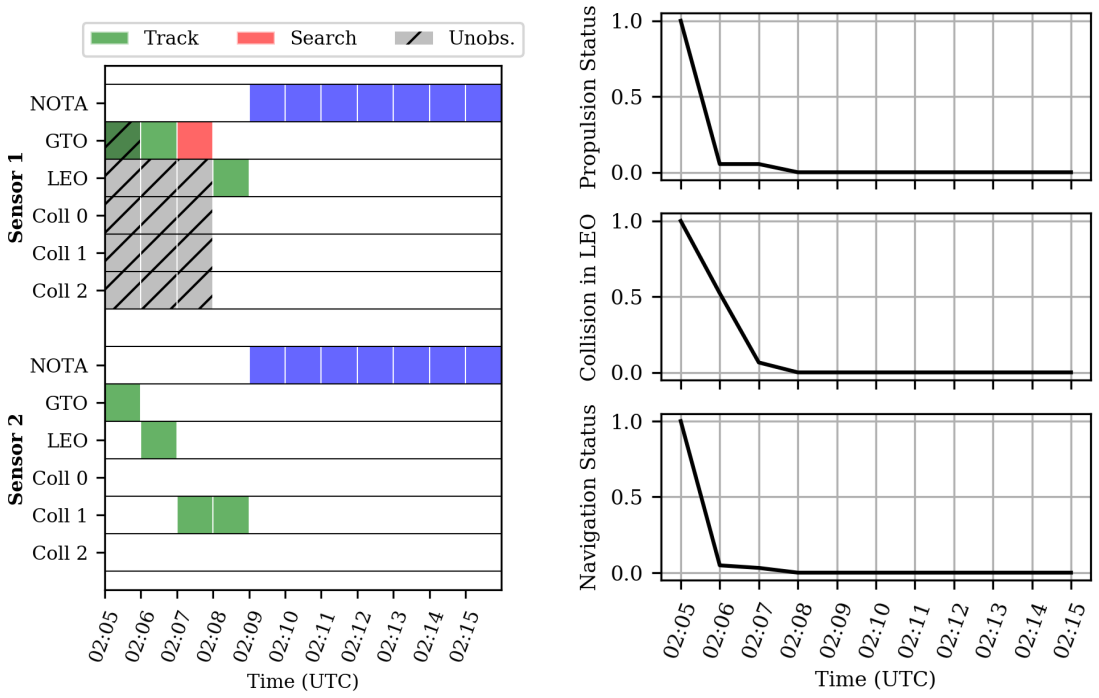
Figure 5.9: Case 1: nominal maneuver (baseline), hypothesis resolutions (solid line for belief, dashed line for plausibility)

entropies are shown in Figs. 5.10(a), and 5.10(b), respectively. The resolutions of each proposition are plotted in Figs. 5.11(a), 5.11(b), and 5.11(c).

Similar to the baseline test case, strong evidence is available to confirm hypothesis resolutions quickly despite observation conditions. The initial observation of GTO by sensor 2 simultaneously misses an expected observation of the GTO object (refuting the "nominal" propulsion proposition) and detects the spacecraft in its LEO parking orbit (confirming the "non-start" propulsion proposition). Successful detections of each collision object (including the same combined detection of "Coll 0" and "Coll 2" at 02:05) refute each collision proposition, and the navigation status is confirmed nominal by detection of the primary spacecraft in LEO.

5.7.7 Case 3: Propulsion System Explosion

In this test case, a propulsion anomaly occurs resulting in an explosion, scattering debris in the LEO parking orbit. The resulting sensor tasking schedule and hypothesis entropies are



(a) Sensor tasking schedule

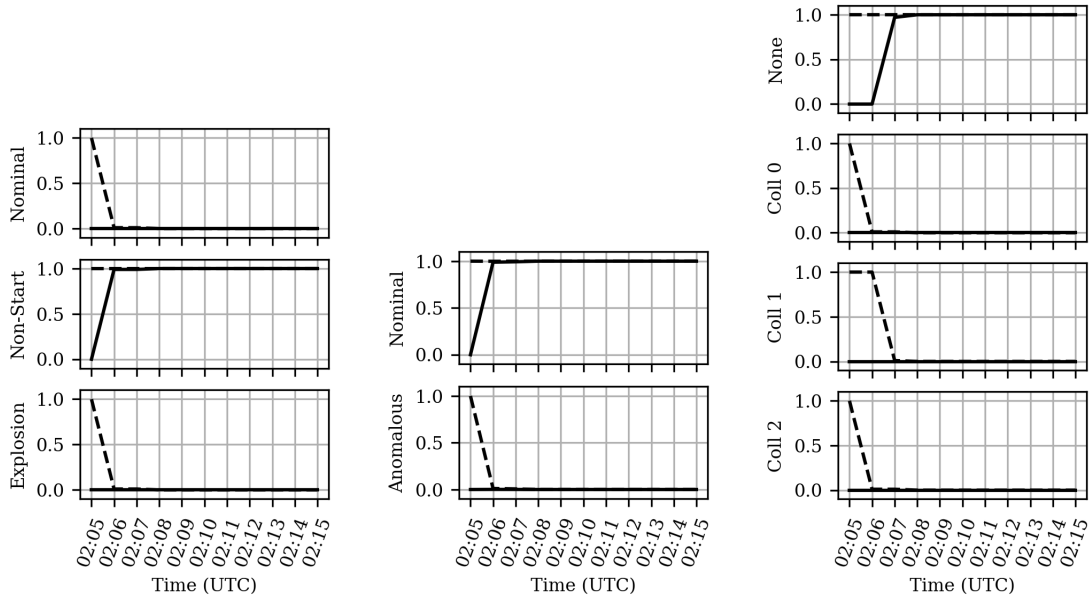
(b) Hypothesis entropy

Figure 5.10: Case 2: propulsion non-start

shown in Figs. 5.12(a), and 5.12(b), respectively. The resolutions of each proposition are plotted in Figs. 5.13(a), 5.13(b), and 5.13(c).

This test case features weaker evidence, resulting in more actions required to reach adequate hypothesis resolution. The initial missed detections in GTO contribute weak evidence toward anomalous propositions for both the propulsion and navigation statuses. The sensor network initiates a search in GTO to confirm that the navigation status is not the cause, searching for the object in an off-nominal GTO state.

In the course of this search, several pieces of debris are detected, contributing evidence toward both the propulsive explosion and collision propositions. This initially inflates the belief in a collision with object “Coll 1” that is later refuted through positive detection of the “Coll 1” object in its nominal orbit. Evidence mounts toward the propulsive explosion proposition as further debris is detected, the target object is not found in GTO or LEO, and each collision object is successfully detected. This test case serves as a prime example of



(a) Propulsion status (b) Navigation status (c) Collision in LEO

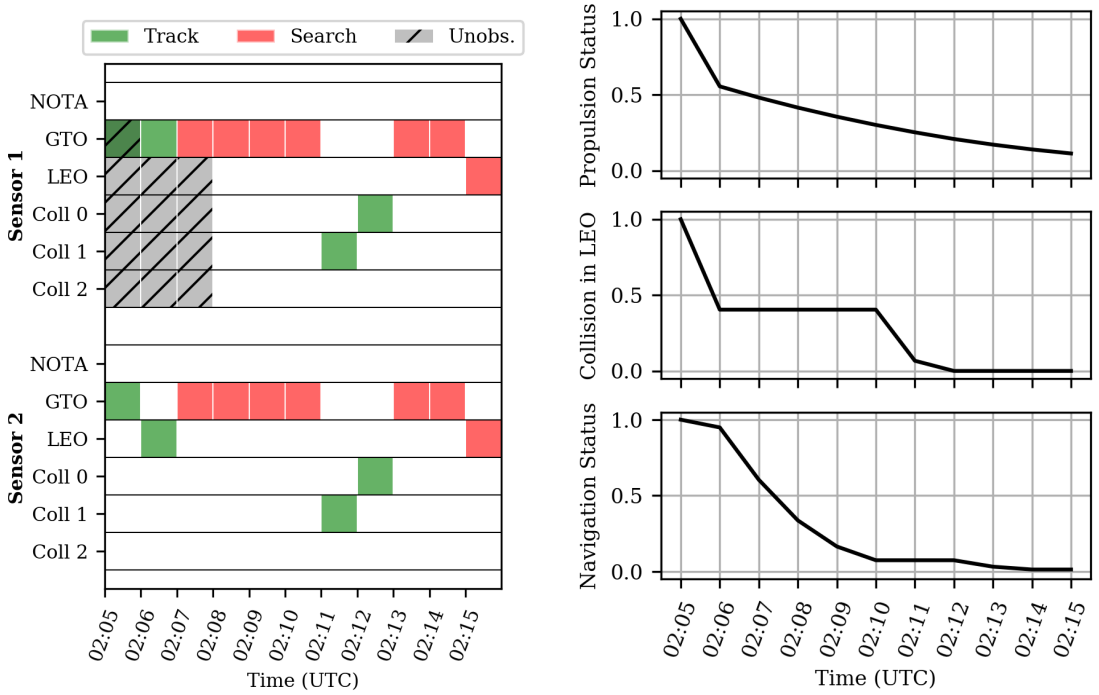
Figure 5.11: Case 2: propulsion non-start, hypothesis resolutions (solid line for belief, dashed line for plausibility)

the unbiased resolution focus of JER, as the prior induced by the initial debris detection is rejected by further inspection.

5.7.8 Case 4: Collision with Object in LEO

In this test case, the true proposition is a collision in LEO with the object labeled "Coll 0." This event generates multiple debris objects in both the LEO parking orbit and the nominal orbit of the collision object. The resulting sensor tasking schedule and hypothesis entropies are shown in Figs. 5.14(a), and 5.14(b), respectively. The resolutions of each proposition are plotted in Figs. 5.15(a), 5.15(b), and 5.15(c).

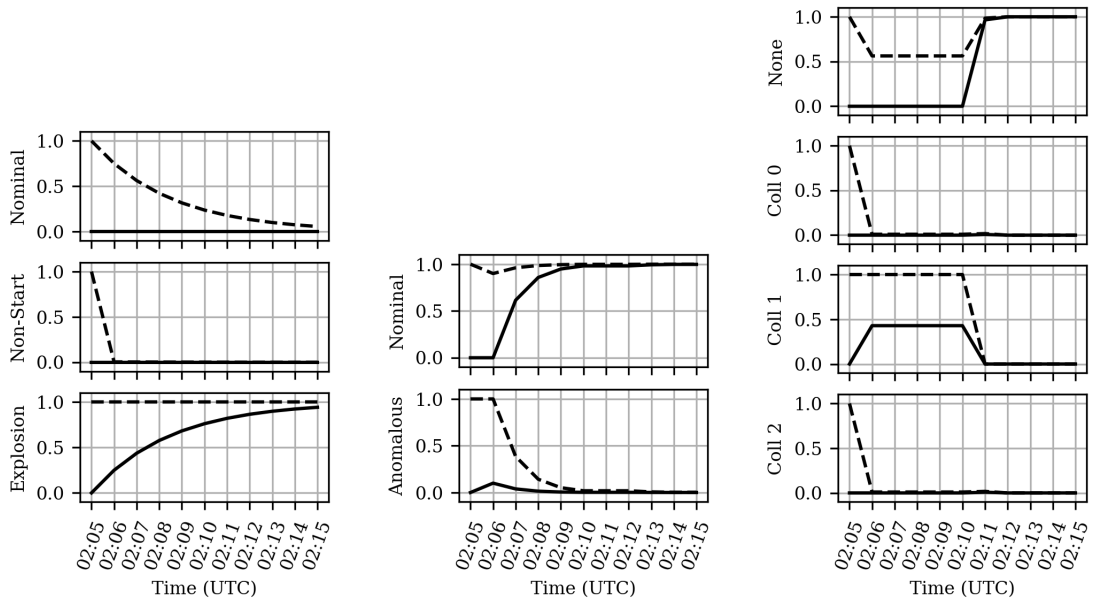
Similar to the propulsion explosion test case, debris detected in the early observations contributes belief in a collision, but successful detections of "Coll 1" and "Coll 2" refute those collision propositions. In this case, debris is detected in both the LEO and "Coll 1" orbits, which differentiates the explosion and collision propositions, contributing additional evidence to refute explosion. A navigation anomaly is also ruled out through search of



(a) Sensor tasking schedule

(b) Hypothesis entropy

Figure 5.12: Case 3: propulsion explosion

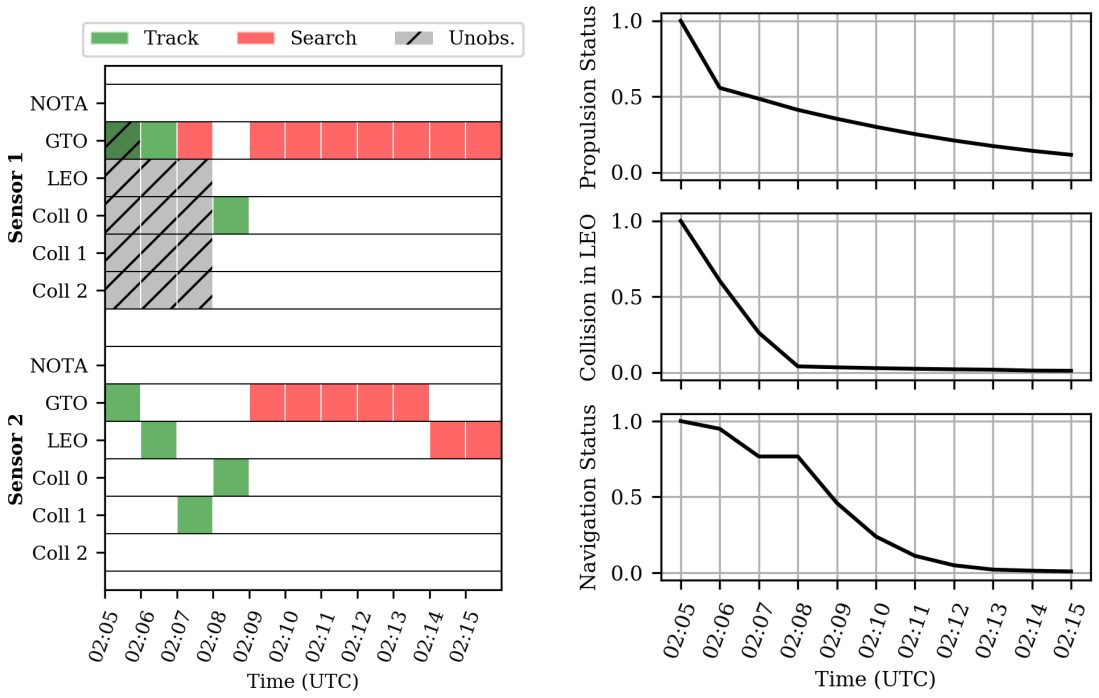


(a) Propulsion status

(b) Navigation status

(c) Collision in LEO

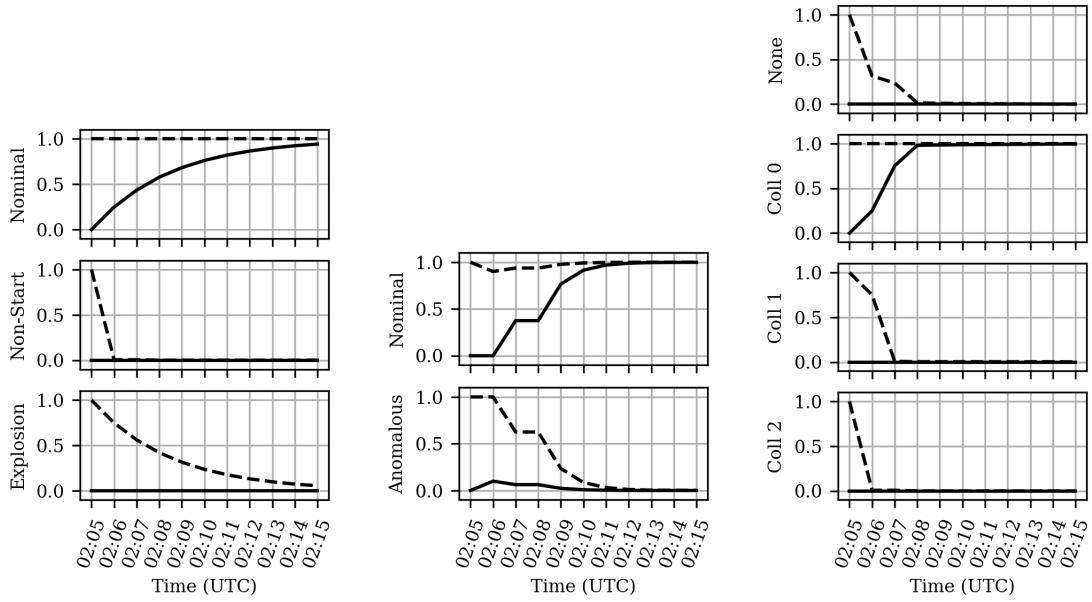
Figure 5.13: Case 3: propulsion explosion, hypothesis resolutions (solid line for belief, dashed line for plausibility)



(a) Sensor tasking schedule

(b) Hypothesis entropy

Figure 5.14: Case 4: collision in LEO



(a) Propulsion status

(b) Navigation status

(c) Collision in LEO

Figure 5.15: Case 4: collision in LEO, hypothesis resolution

the GTO orbit, despite an initial slight belief in navigation anomaly. As the simulation progresses, the evidence builds to strong hypothesis resolutions at the end of the simulation.

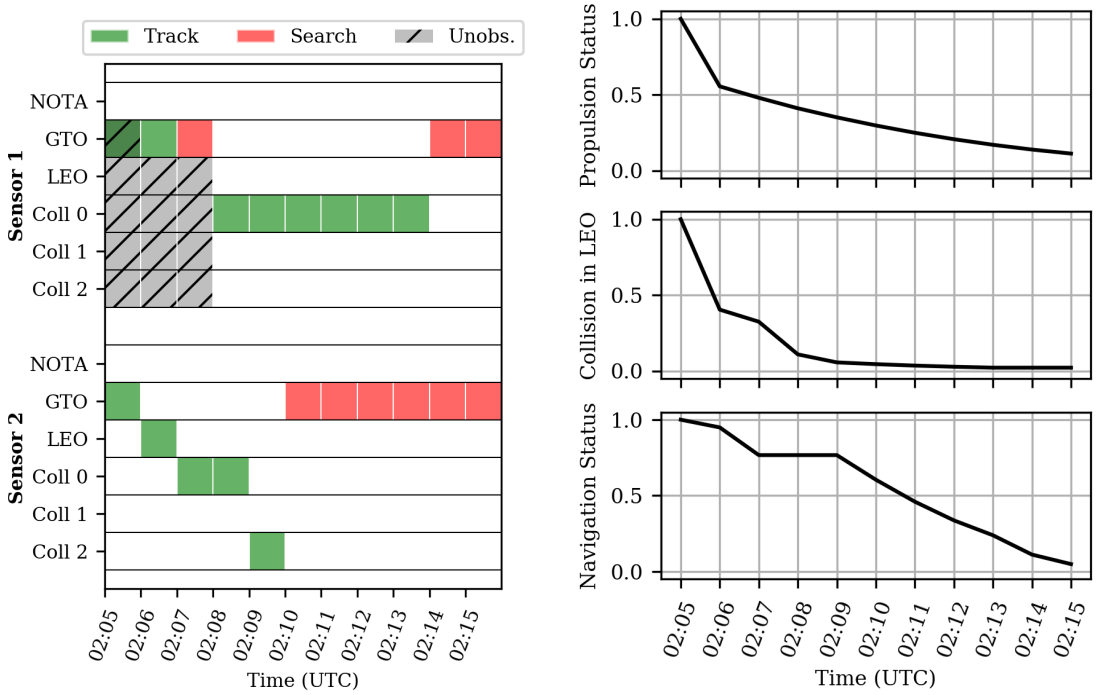
5.7.9 Comparison to Entropy-Greedy Scheduler

As a point of comparison, a brute-force entropy-greedy scheduler was implemented for comparison to the JER approach. The entropy-greedy scheduler evaluates all valid action sequences over the scheduler horizon and selects the action sequence that minimizes the weighted-sum entropy. This represents the brute-force solution to the hypothesis-based evidence-gathering optimization problem in Eqn. (5.8), presented to analyze proposed computational complexity and bias-related improvements of the JER approach.

The biggest immediate difference between the approaches is the number of sequence evaluations required. Even in these low-dimensional scenarios, the brute-force evaluation of all possible tasking sequences (two sensors, five targets, two-step horizon) requires $1,024$ sequence evaluations each iteration for all test cases. In comparison, for the nominal maneuver scenario (Case 1), recall that JER only required a maximum of 271 sequence evaluations (including all agent-pairs and the combination schedule evaluations) in any iteration.

While the entropy-greedy scheduler sometimes performs identically to JER in hypothesis resolution (as in Case 1 and 2), it predictably struggles with confirmation bias. The comparison scenario presented is identical to the propulsion anomaly in Case 3. The resulting sensor tasking schedule and hypothesis entropies are shown in Figs. 5.16(a), and 5.16(b), respectively. The resolutions of each proposition are plotted in Figs. 5.13(a), 5.17(b), and 5.17(c).

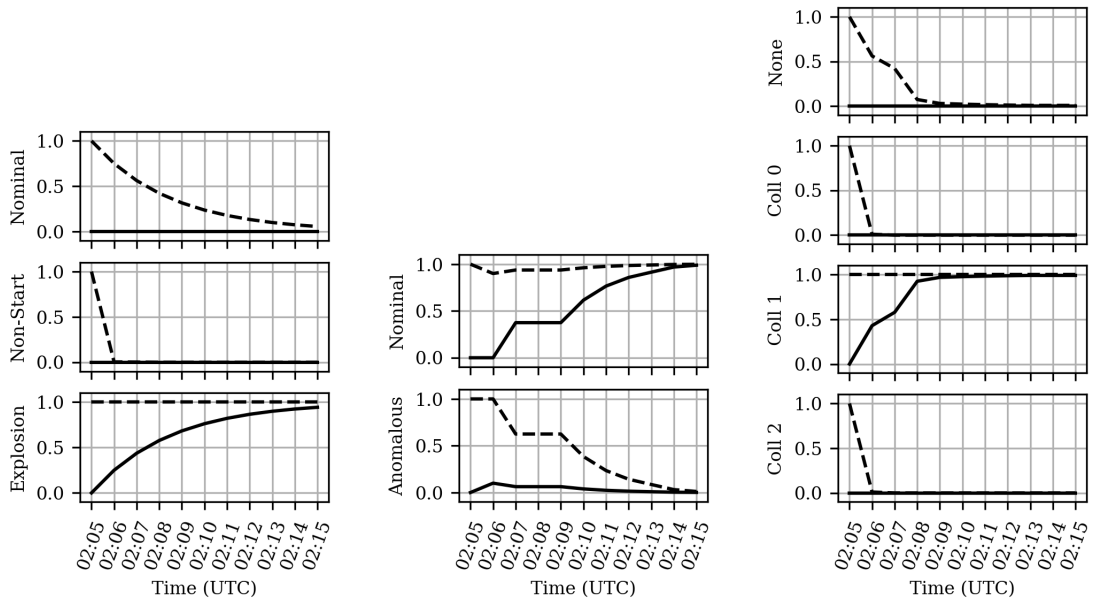
As with Case 3 above, detection of debris early in the simulation contributes evidence toward both the (correct) explosion proposition and the (incorrect) collision propositions. With JER, the alternating-agent scheme overcomes the incorrect collision prior by searching for confirming evidence for each collision proposition. However, in the entropy-greedy



(a) Sensor tasking schedule

(b) Hypothesis entropy

Figure 5.16: Entropy-greedy scheduler: propulsion explosion



(a) Propulsion status

(b) Navigation status

(c) Collision in LEO

Figure 5.17: Entropy-greedy scheduler: propulsion explosion, hypothesis resolutions (solid line for belief, dashed line for plausibility)

approach, once the other propositions are ruled out and only the "Coll 1" proposition remains, the algorithm is satisfied with its hypothesis resolution. This results in an incorrect resolution of the collision status hypothesis.

This is a phenomenon also experienced in previous work using DST for sensor tasking [99], where the mutually exclusive and collectively exhaustive nature of the propositions does not always encourage positive confirmation of the hypothesis resolution. The entropy-greedy results, contrasted with the JER results, further underscore the impact of the alternating-agent scheme in rejecting confirmation bias.

5.7.10 Discussion

These simulated cases show that the JER algorithm performs as designed, seeking strong evidence to resolve hypotheses without fixating on any particular proposition. Weak evidence from missed detections results in the algorithm moving to other hypotheses or propositions that will plausibly produce stronger evidence. Additionally, decomposing the sensor tasking problem into tractable sub-problems through JER agent-pairs increases the feasible time horizon, which is computationally constrained in a brute-force approach, even for this relatively low-dimensional example.

The test cases range from clear scenarios with strong evidence to anomalous scenarios with weak and ambiguous evidence. The nominal maneuver scenario (Case 1) provides a baseline for comparison, quickly resolving the hypotheses with strong evidence. The propulsion non-start scenario (Case 2) shows an ability to ingest both weak evidence (missed detections from GTO) and strong evidence (successful detections in LEO) to explore the hypotheses efficiently. The propulsion explosion and collision scenarios (Cases 3 and 4) highlight the ability to avoid confirmation bias induced by poor prior knowledge or ambiguous evidence by continuing to seek evidence to reject the incorrect propositions. In comparison, the entropy-greedy approach struggled to overcome confirmation bias.

5.8 Conclusion

The proposed Judicial Evidential Reasoning (JER) evidence-gathering framework arranges decision-maker questions as rigorously testable hypotheses to enable predictive evidence-gathering for hypothesis resolution. The use of a hypothesis abstraction supports human decision-making strengths of planning and strategy, off-loading processing work to the algorithm and fusing evidence into intuitive hypothesis resolutions. Recognizing the need to account for ambiguity aversion in decision-making, the use of Dempster-Shafer theory allows for quantification of evidence ambiguity. Finally, applying the principle of equal effort through an alternating-turn adversarial optimization scheme avoids confirmation bias induced by improper prior beliefs or evidence uncertainty and ambiguity, avoiding fixation on incorrect propositions.

This approach values impartiality in addition to time-efficiency in many-hypothesis resolution, while breaking the greater evidence-gathering problem into a number of sub-problems for each hypothesis reduces computational complexity and allows for a receding horizon optimization of the total schedule. Selecting the final optimal schedules as the minimum total weighted entropy ensures that the selected actions minimize conflict and non-specificity according to priorities set by the decision-makers.

The provided example cases illustrate the application of both the JER agent-pairs and the overall JER schedule manager approach to evidence-gathering. The simulated results for a GTO insertion maneuver anomaly scenario show that the algorithm performs as expected: the appropriate hypotheses are confirmed via evidence and in the process the JER algorithm does not fixate on any particular proposition, instead accruing evidence that gradually leads to the correct conclusion. The JER approach also compares well against an entropy-greedy approach that focuses actions on the most-probable propositions only, avoiding improper hypothesis resolution caused by confirmation bias.

The combined emphasis on hypothesis abstraction, quantifying ambiguity, and avoiding

confirmation bias in the JER approach enables predictive evidence-gathering for hypothesis resolution. Future work will investigate the human cognitive effects of a JER-like hypothesis-based evidence gathering approach to further develop decision support systems that effectively support human-in-the-loop decision-making.

CHAPTER 6

DEVELOPMENT OF A PROTOTYPE DSS FOR SSA

The cognitive work analysis in Chapter 2 identified a number insights of and design requirements that drive DSS development for SSA applications. The following chapters developed hypothesis-based applications aimed at addressing SSA goals, culminating in the development of Judicial Evidential Reasoning (JER) as a generalized evidence-gathering hypothesis-resolution framework. In this chapter, a subset of derived design requirements are explicitly addressed in the development and validation of a prototype DSS for SSA. In particular, the design requirements selected for further analysis are ones that align well with the main tenets of JER to further investigate how hypothesis resolution supports decision-making in SSA.

6.1 Design Requirements Addressed

As stated previously, existing methods of sensor tasking in SSA focus on state uncertainty reduction. While this tasking goal correlates well with some SSA goals (e.g. conjunction analysis), not all decision-maker questions can easily be mapped to state covariance estimates. Therefore, state uncertainty minimization methods do not provide a reliable means of tasking to resolve decision-maker hypotheses related to the overarching SSA goals of space asset safety and national security.

In order to make connections between covariance estimates and other SSA hypotheses, current methods require decision-makers to do significant knowledge-based reasoning in the judgment phase (recall the phases of the control task analysis, Fig. 2.4). Specifically, when evaluating different courses of action, the decision-maker must consider if the covariance reduction resulting from a given set of actions also gathers evidence to resolve the high-priority hypotheses. This requires reasoning on several different levels of the ab-

straction hierarchy (recall Fig. 2.2: state covariances on the physical function level, sensor allocation and information fusion on the generalized function level, and hypotheses and priorities on the abstract function level. These are three very different levels of detail of the SSA problem, and requiring the decision-maker to move between them quickly (and especially iteratively, as in the judgment phase) often leads to reduced situation awareness and increased measures of workload (e.g. effort, mental demand, frustration).

Conversely, a DSS that suggests tasking assignments based directly on resolving hypotheses only requires the decision-maker to remain in the abstract function level, considering trades between priorities for the different hypotheses without having to be directly concerned with the sensor allocation or state measurements. It also frees the decision-maker to spend more mental effort on formulating hypotheses to directly support the dynamic list of SSA goals. Humans out-perform automation at abstract tasks such as prioritization, so this application stands to be a better use of human decision-maker effort and cognition. This cognitive engineering study aims to investigate that claim: that SSA decision-maker situation awareness and workload are improved when using a hypothesis-based tasking algorithm as opposed to a more traditional covariance-based scheduler.

In order to investigate this claim, a prototype DSS was developed that incorporated a number of the design requirements from the ConTA in Chapter 2. Specifically, due to the development of the JER framework in the previous chapter, the design requirements selected are ones which evidential reasoning help to address. The following cognitive work requirements (CWRs) are considered in this DSS development:

- **CWR-3:** The DSS shall translate observational data into evidence.
- **CWR-6:** The DSS shall track hypothesis resolution in comparison to prescribed thresholds.
- **CWR-9:** The DSS shall provide capability for operators to adjust hypothesis priorities.

- **CWR-10:** The DSS shall assess expected hypothesis resolution based on current prioritization.
- **CWR-13:** The DSS shall generate specific actions and requests required to reach the target hypothesis resolution.

A primary function of hypothesis-based sensor tasking, applied to SSA, is to analyze candidate tasking schedules to estimate hypothesis resolution based on the current hypothesis prioritization (CWR-10). Evidential reasoning approaches are predicated upon developing mappings from acquired data to evidence, so any evidential reasoning application must satisfactorily address CWR-3. Through the use of hypothesis entropy, quantifying both conflict and ambiguity, evidential reasoning provides a means for tracking hypothesis resolution against prescribed thresholds (CWR-6) Additionally, as demonstrated by JER, applying and adjusting hypothesis weights based on entropy allows decision-makers to prioritize hypotheses appropriately (CWR-9). Finally, the end result of a hypothesis-based tasking approach is the list of actions that gather the required evidence to resolve the hypotheses (CWR-13). Therefore, an evidential reasoning approach to sensor tasking, such as JER, appropriately addresses these five CWRs.

6.2 DSS Design

This section describes the specifics of the prototype DSS design, including the two sensor tasking schedulers and an overview of the functionality of the simulation environment. Recall that this study aims to compare a hypothesis-based tasking algorithm with a more traditional covariance-based scheduler for the purposes of supporting decision-making, situation awareness, and workload in SSA.

6.2.1 Sensor Tasking Schedulers

The two algorithms implemented in the prototype DSS are a covariance minimization approach and an evidential reasoning approach. When using the covariance-based scheduler, the next sensor actions are selected using Eqn. (6.1) to minimize the weighted-sum of covariances:

$$\mathcal{A}^* = \underset{\mathcal{A}\in\mathbb{A}}{\arg\min} \sum_{j=1}^{N} w_j \frac{\mathrm{Tr}\left[\mathbf{P}_j^+\right]}{a_j} \tag{6.1}$$

where $\mathbb{A}$ is the set of valid action sequences at the current time step, N is the number of space objects, $\mathbf{P}_j^+$ is the estimated a posteriori state covariance of the j^th space object, and a_j is the semi-major axis of the j^{th} space object. The semi-major axis scaling factor is included to account for naturally larger uncertainties at higher-altitude orbits and prevent those from dominating the sensor tasking. In this mode, the subject is able to change weighting parameters to attempt to increase or reduce tasking actions taken on a particular object.

When using the hypothesis-based scheduler, the next sensor actions are selected using Eqn. (6.2) to minimize the weighted-sum of hypothesis entropies:

$$\mathcal{A}^* = \underset{\mathcal{A}\in\mathbb{A}}{\arg\min} \sum_{i=1}^{H} w_i \tilde{H}_{JS}^+ \left(m_i\right) \tag{6.2}$$

where $\mathbb{A}$ is again the set of valid action sequences at the current time step, H is the number of hypotheses considered, and $\tilde{H}_{JS}\left(m_i\right)^+$ is the estimated a posteriori normalized Jirousek-Shenoy entropy for the i^{th} hypothesis as defined in Eqn. (5.18). In this mode, the subject is able to change weighting parameters to attempt to increase or reduce tasking actions taken against a particular hypothesis.

In both modes, the subject's weighting parameters may affect the scheduled sensor actions; however, the subject is not able to directly assign any space object or hypothesis

to any sensor. Therefore, if the evidence available is low/weak or the object's covariance is already sufficiently small, the subject may not be able to override the algorithm.

6.2.2 Simulation Environment

The prototype DSS was developed in C# using the Unity engine and modeled after existing SSA support tools such as the Systems Tool Kit (STK) by Analytical Graphics, Inc. (AGI). Figure 6.1 shows a screen-capture of the prototype SSA DSS. The primary display is a 3D view of the Earth and all the sensors (purple boxes) and space objects (green boxes). The operator may use arrow keys or right-click-and-drag motions to rotate the camera around the Earth. The top-middle panel contains information on the current simulation time and the time remaining to until the next actions are taken.

The list on the top-right of the screen contains buttons linked to each sensor and space object. Clicking any of these buttons changes the information panel in the bottom-left to display data relevant to that sensor (e.g. observation conditions, objects above the horizon) or space object (e.g. state estimate and uncertainty). This also changes the focus on the camera to that particular sensor or object, allowing the operator to pivot the camera around that new focus point. The Earth button in the top-right of the screen can be used to re-focus the display on the Earth.

Similarly, the list on the middle-right of the screen contains buttons linked to each hypothesis under consideration. Clicking any of these buttons changes the information panel in the bottom-right to display data relevant to that hypothesis (e.g. current probabilities for each proposition, entropy, number of actions taken, and related space objects).

In the top-left of the screen, another list displays each sensor in the network and its current and next tasks. When the time remaining until the next action hits zero, the next task in each sensor's cue is triggered. At this point, the 3D display updates the sensor fields-of-view accordingly, the measurement updates are applied to each space object, and any gathered evidence is also used to update the hypotheses. Summaries of the actions

taken and any resulting evidence gathered are shown in the message console below the tasking list.

The only control that the user has over the scheduler is set of weighting parameters on either the space objects (for the covariance-based scheduler) or the hypotheses (for the hypothesis-based scheduler). These weights are controlled using the slider that appears in the lower-middle of the screen, divided into 11 segments: 0% to 100% in 10% increments. When using the hypothesis-based scheduler, the slider appears directly to the left of the hypothesis information panel (as shown in Fig. 6.1). For the covariance-based scheduler, the slider appears directly to the right of the space object information panel.

6.3 Human-in-the-Loop Data Collection

A human-in-the-loop test was performed using this prototype DSS (see Appendix B for the approved International Review Board consent form). The specific goal of this test was to investigate the situation awareness, cognitive support, and workload effects of using a hypothesis-based sensor-tasking scheduler as compared to a more traditional covariance-based scheduler. As described in the DSS design section, the test subject's only control over the tasking algorithm is through the set of weighting parameters corresponding to relative priorities on either space objects or hypotheses.

In a one-hour training session, the test participants were given an abbreviated background on the SSA problem and instructed on the use of the DSS for gathering evidence using sensor observations to answer a set of questions. Subjects were informed of the differences between the two scheduling algorithms and their respective controls. At the end of the session, each participant completed a qualification test scenario to ensure he/she could successfully operate the DSS to gather evidence and answer the questionnaires.

Two subsequent data collection sessions were conducted for each participant, one session each for the covariance- and hypothesis-based schedulers. To control for learning effects, each subject was randomly assigned the scheduler type for the first day of testing,

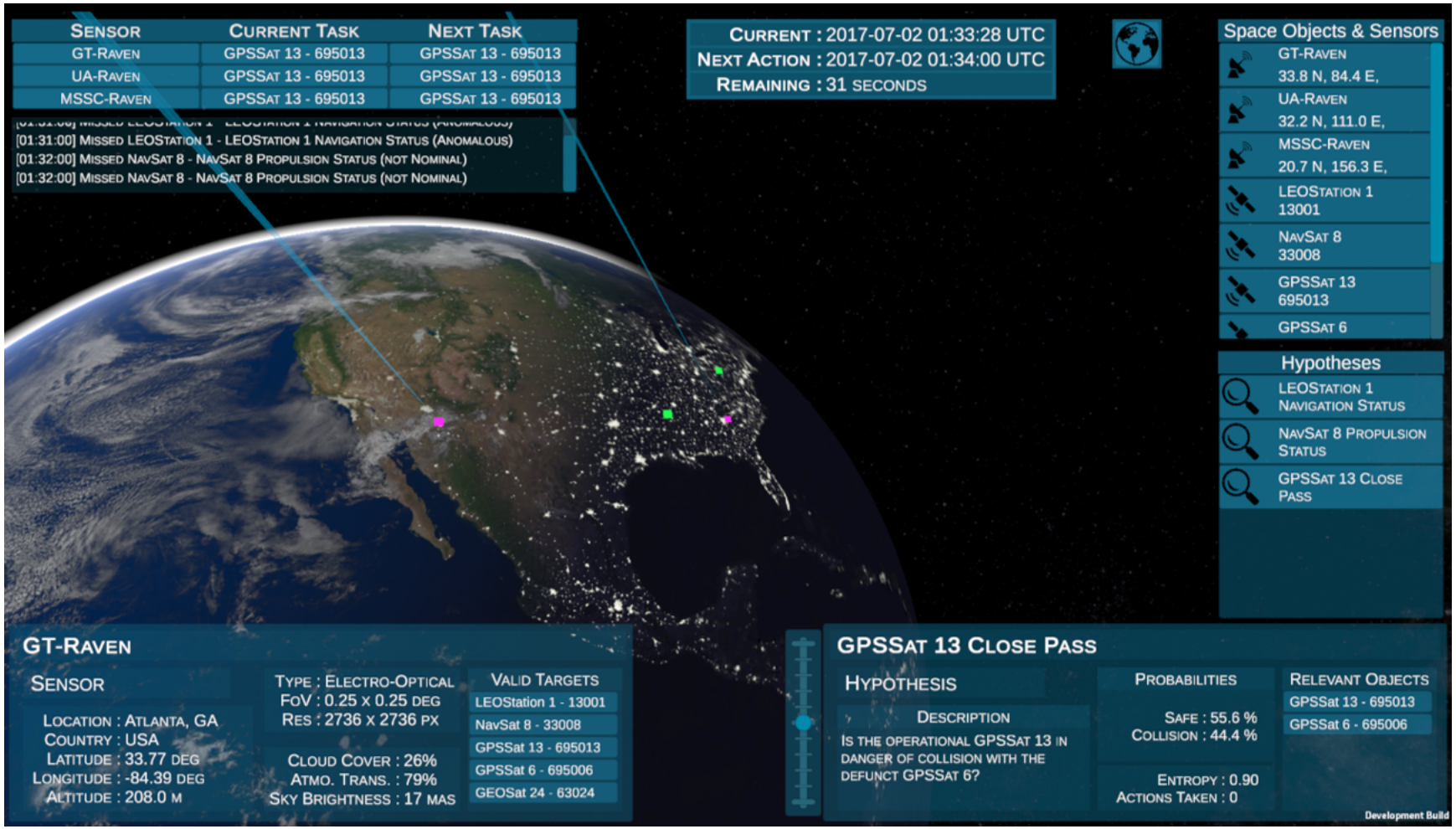


Figure 6.1: Screen-shot of the prototype SSA DSS

so that six subjects used the covariance-based scheduler on the first test day and five used the hypothesis-based scheduler.

The data collection methods are similar to a recent DSS evaluation study published to assess emergency department information displays [125], consisting primarily of questionnaires to assess cognitive support and workload. Before any questionnaire is completed, the simulation display is replaced entirely by the questionnaire so that the subject cannot rely on the DSS to answer the questions [6]. The cognitive support and workload questionnaires were asked at the end of each scenario, and the questions were the same each time. The cognitive support questionnaire assesses the subject's opinion on the ability of each scheduler to support various cognitive objectives. Subject responses to the cognitive support questions are on an 11-point scale, with 1 being "not at all effective" and 11 being "extremely effective." The eight questions asked in this questionnaire are: How effective was this scheduler's support for...

- identifying which questions you still need to answer?
- proposing sensor tasking assignments that answer relevant questions?
- allowing modifications of the scheduled tasking assignments as needed?
- identifying which space objects have high state uncertainty?
- identifying which questions have high uncertainty (entropy)?
- assessing the sensor resources required to answer questions?
- assessing the strength of the evidence received?
- adapting to sensor observation conditions?

The workload questionnaire is the widely-used NASA Task Load Index (NASA-TLX) [126], which assess the subject's perceived workload. Per NASA-TLX convention, subject responses are are on a 21-point scale, where lower-rating responses correspond to a

better perceived workload (e.g. 1 = perfect performance). The six questions asked in this questionnaire are:

- Mental Demand: How mentally demanding was the task?
- Physical Demand: How physically demanding was the task?
- Temporal Demand: How hurried or rushed was the pace of the task?
- Performance: How successful were you in accomplishing what your task?
- Effort: How hard did you have to work to accomplish your level of performance?
- Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?

The responses to the performance question range from "perfect" to "failure" while responses to the remainder of the questions range from "very low" to "very high."

Together, these questionnaires provide subjective measurements of the cognitive and workload support provided by each scheduler.

6.4 Test Scenarios

The test participants executed the evidence-gathering tasks in five different scenarios, each with slight modifications to test responses in different conditions. Running each scenario with both scheduling algorithms allows for direct comparison between the schedulers in support for the given hypothesis-resolution tasks. To control for learning effects, the orders of the tests performed were randomized for each participant. Additionally, the space objects identifiers were changed for each scenario and between the schedulers so that subjects could not rely on recollection of the previous evaluation session to answer questions.

Each scenario consists of two or three sensors, each with different background sky brightness, cloud cover, and atmospheric transmittance that contribute to observation condition quality. Observation conditions of moderate or better allow for successful detections, but conditions of poor or worse would cause missed detections. Additionally, each

Table 6.1: Summary of hypotheses and propositions for simulations

Hypothesis	**Propositions**
Close Pass	Safe, Collision
Propulsion Status	Nominal, Non-Start, Explosion
Navigation Status	Nominal, Anomalous
Custody Status	Maintained, Lost

scenario consists of five space objects: two in low Earth orbit (LEO), two in medium Earth orbit (MEO), and one in geostationary Earth orbit (GEO). Finally, each scenario consists of four hypotheses. The types of hypotheses allowed and their respective propositions are summarized in Table 6.1.

The following sections provide more detail on each scenario, including sensor observation conditions, space object initial one-sigma state uncertainties in the local-vertical local-horizontal (LVLH) frame, and hypothesis priors and strength of evidence available.

Scenario A: In this scenario, three ground-based sensors are available (see Table 6.2) with observation conditions ranging from moderate to very good. The relevant space object parameters are listed in Table 6.3, and the relevant hypotheses in Table 6.4. The first two space objects are in danger of a collision, which must be confirmed by evidence despite vacuous prior knowledge. The third object navigation status and the fourth object propulsion statuses are all nominal, while the GEO object custody status is maintained, all of which must also be confirmed by evidence. This scenario represents a baseline test case with moderate evidence strength and no missed-detections due to observation conditions.

Table 6.2: Sensor parameters, scenario A

ID	**Location**	**Conditions**
MSSC-Raven	$20.7^\circ\ N,\ 156.2^\circ\ W$	Very Good
UA-Raven	$32.2^\circ\ N,\ 111.9^\circ\ W$	Good
GT-Raven	$33.7^\circ\ N,\ 84.3^\circ\ W$	Moderate

Scenario B: In this scenario, three ground-based sensors are available once again (see Table 6.5) though now the GT-Raven observation conditions are too poor for successful detections. The relevant space object parameters are listed in Table 6.6, and the relevant

Table 6.3: Space object parameters, scenario A

ID	Regime	σ_r (km, LVLH)	σ_v (m/s, LVLH)
A1	LEO	$1.0, 2.0, 0.7$	$0.75, 1.25, 0.20$
A2	LEO	$5.4, 7.8, 1.2$	$0.90, 1.05, 0.25$
A3	MEO	$9.3, 10.2, 3.1$	$0.66, 0.94, 0.32$
A4	MEO	$8.2, 9.8, 2.2$	$0.54, 0.79, 0.21$
A5	GEO	$15.5, 42.1, 6.8$	$0.27, 0.95, 0.09$

Table 6.4: Hypothesis parameters, scenario A

ID	Type	Objects	Truth	Evidence	Prior
A1	Close Pass	A1, A2	Collision	Moderate	Vacuous
A2	Navigation Status	A3	Nominal	Moderate	Nominal (slight)
A3	Propulsion Status	A4	Nominal	Moderate	Nominal (slight)
A4	Custody Status	A5	Maintained	Strong	Maintained (slight)

hypotheses in Table 6.7. The first object navigation status is nominal, and its prior has a slight belief for nominal. The second object propulsion status is a non-start, but the prior knowledge is vacuous. The third and fourth object are not in danger of a collision, but this must be confirmed despite vacuous prior knowledge. The fifth object custody status is maintained, with a slight prior belief for the maintained proposition as well. This scenario primarily differs from scenario A by including poor observation conditions, causing all observations from GT-Raven to miss.

Table 6.5: Sensor parameters, scenario B

ID	Location	Conditions
MSSC-Raven	$20.7° N,\ 156.2° W$	Very Good
UA-Raven	$32.2° N,\ 111.9° W$	Good
GT-Raven	$33.7° N,\ 84.3° W$	Poor

Table 6.6: Space object parameters, scenario B

ID	Regime	σ_r (km, LVLH)	σ_v (m/s, LVLH)
B1	LEO	$1.0, 2.0, 0.7$	$0.75, 1.25, 0.20$
B2	LEO	$5.4, 7.8, 1.2$	$0.90, 1.05, 0.25$
B3	MEO	$5.0, 13.0, 2.1$	$0.45, 0.90, 0.15$
B4	MEO	$11.2, 18.4, 3.6$	$0.65, 0.98, 0.35$
B5	GEO	$15.5, 42.1, 6.8$	$0.27, 0.95, 0.09$

Table 6.7: Hypothesis parameters, scenario B

ID	Type	Objects	Truth	Evidence	Prior
B1	Navigation Status	B1	Nominal	Moderate	Nominal (slight)
B2	Propulsion Status	B2	Non-Start	Moderate	Vacuous
B3	Close Pass	B3, B4	Safe	Strong	Safe (slight)
B4	Custody Status	B5	Maintained	Strong	Maintained (slight)

Scenario C: In this scenario, only two ground-based sensors are available (see Table 6.8). The relevant space object parameters are listed in Table 6.9, and the relevant hypotheses in Table 6.10. The navigation status for both objects 1 and 4 are nominal, but the evidence for object 4 is stronger. The propulsion status for object 2 is also nominal, and the custody status for object 5 is maintained. The primary difference in scenario C is the presence of an extra space object (object 3) that does not relate to any hypotheses, testing the operators awareness to focus tasking actions on other objects.

Table 6.8: Sensor parameters, scenario C

ID	Location	Conditions
MSSC-Raven	$20.7° \, N, \, 156.2° \, W$	Very Good
UA-Raven	$32.2° \, N, \, 111.9° \, W$	Good

Table 6.9: Space object parameters, scenario C

ID	Regime	σ_r (km, LVLH)	σ_v (m/s, LVLH)
C1	LEO	$1.0, 2.0, 0.7$	$0.75, 1.25, 0.20$
C2	LEO	$5.4, 7.8, 1.2$	$0.90, 1.05, 0.25$
C3	MEO	$5.0, 13.0, 2.1$	$0.45, 0.90, 0.15$
C4	MEO	$11.2, 18.4, 3.6$	$0.65, 0.98, 0.35$
C5	GEO	$15.5, 42.1, 6.8$	$0.27, 0.95, 0.09$

Table 6.10: Hypothesis parameters, scenario C

ID	Type	Objects	Truth	Evidence	Prior
C1	Navigation Status	C1	Nominal	Moderate	Nominal (slight)
C2	Propulsion Status	C2	Nominal	Moderate	Vacuous
C3	Navigation Status	C4	Nominal	Strong	Nominal (slight)
C4	Custody Status	C5	Maintained	Moderate	Maintained (moderate)

Scenario D: In this scenario, three ground-based sensors are available (see Table 6.11),

with poor observation conditions at GT-Raven causing missed detections. The relevant space object parameters are listed in Table 6.12, and the relevant hypotheses in Table 6.13. Objects 1 and 2 experience a safe close pass, but the prior knowledge of this event is vacuous. Object 2 also has an anomalous navigation status, though the prior knowledge incorrectly indicates nominal status. The propulsion status for object 3 is nominal, while the propulsion status for object 5 is a non-start. The prior knowledge for object 5's propulsion status is also incorrect, with a moderate preference toward nominal. This scenario once again contains an extra space object (object 4) that does not relate to any hypotheses. Additionally, since object 2 is associated with two separate hypotheses, this tests whether the algorithm and operator recognize that this object can be leveraged for strong evidence. Finally, two of the hypotheses begin with incorrect prior assumptions, testing whether the operators will still prioritize gathering evidence to resolve these hypotheses.

Table 6.11: Sensor parameters, scenario D

ID	**Location**	**Conditions**
UA-Raven	$32.2° \, N, \, 111.9° \, W$	Good
GT-Raven	$33.7° \, N, \, 84.3° \, W$	Poor
MSSC-Raven	$20.7° \, N, \, 156.2° \, W$	Moderate

Table 6.12: Space object parameters, scenario D

ID	**Regime**	σ_r **(km, LVLH)**	σ_v **(m/s, LVLH)**
D1	LEO	$1.0, 2.0, 0.7$	$0.75, 1.25, 0.20$
D2	LEO	$5.4, 7.8, 1.2$	$0.90, 1.05, 0.25$
D3	MEO	$8.2, 9.8, 2.2$	$0.54, 0.79, 0.21$
D4	MEO	$9.3, 10.2, 3.1$	$0.66, 0.94, 0.32$
D5	GEO	$15.5, 42.1, 6.8$	$0.27, 0.95, 0.09$

Table 6.13: Hypothesis parameters, scenario D

ID	**Type**	**Objects**	**Truth**	**Evidence**	**Prior**
D1	Close Pass	D1, D2	Safe	Moderate	Vacuous
D2	Navigation Status	D2	Anomalous	Moderate	Nominal (slight)
D3	Propulsion Status	D3	Nominal	Moderate	Nominal (slight)
D4	Propulsion Status	D5	Non-Start	Moderate	Nominal (moderate)

Scenario E: In this scenario, only two ground-based sensors are available (see Table 6.14) with moderate to very good observation conditions. The relevant space object parameters are listed in Table 6.15, and the relevant hypotheses in Table 6.16. The first and second objects experience a safe close pass. The second object's navigation status must also be correctly identified as anomalous despite a prior belief in the nominal status. The fourth object has experienced a propulsive anomaly resulting in explosion, so the detection of debris is required to correctly resolve this hypothesis. The fifth object's custody status must be verified as maintained. Similar to scenario D, this scenario features a limited number of sensors and two incorrect prior beliefs, as well as an extra object that is not related to any hypotheses. This scenario stretches the sensor network's evidence gathering capability through limited observability to see how that affects hypothesis resolution and operator workload.

Table 6.14: Sensor parameters, scenario E

ID	**Location**	**Conditions**
UA-Raven	$32.2° \, N, \, 111.9° \, W$	Very Good
GT-Raven	$33.7° \, N, \, 84.3° \, W$	Moderate

Table 6.15: Space object parameters, scenario E

ID	**Regime**	σ_r **(km, LVLH)**	σ_v **(m/s, LVLH)**
E1	LEO	$1.0, 2.0, 0.7$	$0.75, 1.25, 0.20$
E2	LEO	$5.4, 7.8, 1.2$	$0.90, 1.05, 0.25$
E3	MEO	$9.3, 10.2, 3.1$	$0.66, 0.94, 0.32$
E4	MEO	$8.2, 9.8, 2.2$	$0.54, 0.79, 0.21$
E5	GEO	$15.5, 42.1, 6.8$	$0.27, 0.95, 0.09$

Table 6.16: Hypothesis parameters, scenario E

ID	**Type**	**Objects**	**Truth**	**Evidence**	**Prior**
E1	Close Pass	E1, E2	Safe	Moderate	Safe (slight)
E2	Navigation Status	E2	Anomalous	Moderate	Nominal (slight)
E3	Propulsion Status	E4	Explosion	Moderate	Nominal (slight)
E4	Custody Status	E5	Maintained	Moderate	Maintained (moderate)

6.5 Results and Discussion

The data obtained for all 11 test subjects was compiled and analyzed to develop statistical measures for subjects' responses. The box-and-whisker plots, as illustrated in Fig. 6.2, display the first quartile, second quartile (median), and third quartile, as well as the $\pm 1\sigma$ points of the data set. This concisely presents the distribution of responses on a single axis to allow comparison between scheduling methods and scenarios. Small boxes with short whiskers indicate clustered responses, while large boxes or long whiskers indicate a wider distribution and less agreement. Therefore, the values of the response statistics can be compared (i.e. the all three quartiles are better for scheduler 1, indicating better scores) as well as the distribution of the values (i.e. the response distribution is much wider for scheduler 1, indicating less agreement in responses). Though the response values are subjective, each subject answers the same questions in each scenario using each scheduler, so comparing between the distributions normalizes for response scaling differences between test subjects

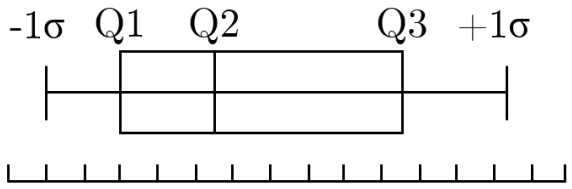


Figure 6.2: Box-and-whisker plot illustration

6.5.1 NASA-TLX

Figures 6.3 and 6.4, and 6.7 summarize the responses to the NASA-TLX questionnaires. Strong trends are identified in the mental demand, temporal demand, and effort responses, indicated by better scores for the hypothesis-based scheduler in Figs. 6.3(a), 6.3(c), and 6.4(b). The hypothesis-based scheduler allowed operators to accomplish their tasks with lower workload requirements. As expected, physical demand (see Fig. 6.3(b)) for both schedulers is very low. The responses for performance and frustration are less conclusive. The subjects' subjective opinion of their performance at the task varies between scenario, as

shown in Fig. 6.4(a). Interestingly, the performance response distribution for scenario C is much wider for the hypothesis-based scheduler, indicating that subjects could not identify their performance as well in this scenario. This is more indicative of a poor scenario design, and further results from the cognitive support questionnaire (below) corroborate this view. Finally, the frustration responses in Fig. 6.4(c) are low, scoring well with both schedulers in most scenarios. Both schedulers score nearly identically in scenario A, and the hypothesis-based scheduler scores slightly better in scenarios B, C, and E while the covariance-based scheduler scores better in scenario D.

While the frustration and performance responses are inconclusive, responses to mental demand, temporal demand, and effort all indicate improved workload while using the hypothesis-based scheduler. In particular, the performance trends can be further illuminated by the following analysis of the cognitive support responses.

6.5.2 Cognitive Support

Figures 6.5, 6.6, and 6.7 summarize the responses to the cognitive support questionnaires. For this initial set of tests, the cognitive support results provide limited intuition for comparisons between the schedulers but do provide useful insight into the scenario design. For example, responses to prompts 1 (identifying unresolved questions, Fig. 6.5(a)) and 2 (proposing tasking to answer relevant questions, Fig. 6.5(b)) do not clearly identify any trends between the schedulers. In scenarios A and B, for instance, the hypothesis scheduler scores better on both prompts, but scenarios C, D, and E show similar responses for both schedulers.

Prompt 3 (allowing modifications to schedules, Fig. 6.5(c)) does not appear conclusive in comparing between schedulers either, but provides useful insight on the scenario design. The subjects scored the hypothesis scheduler lower in this prompt for both scenarios A and C, which are scenarios where all three available sensors exhaust the search space of evidence before the scenario terminates. In this case, the hypothesis scheduler becomes

unresponsive to changes in certain hypothesis weights because no evidence is available to gather for those hypotheses, refusing to task on them even if the weight is maximized. It is interesting to note that this condition was seemingly not noticed in the covariance scheduler, leading subject to take actions that did not gather useful evidence. This indicates that these scenarios could be improved to avoid exhausting all evidence for certain hypotheses, and that the DSS should be improved to more clearly indicate incoming evidence to alert the users in the covariance-scheduler to vacuous evidence.

Prompts 4 (identifying high state uncertainty, Fig. 6.6(a)) and 5 (identifying high hypothesis entropy, Fig. 6.6(b)) exhibit expected trends. The covariance scheduler mostly scores better in identifying state uncertainty as the operator must be aware of covariance values in changing covariance weights. Similarly, the hypothesis scheduler scores better across the board in identifying hypothesis entropy.

Prompt 6 (assessing resources required to answer questions, Fig. 6.6(c)) indicates slightly better scores for the hypothesis scheduler, likely because the hypothesis scheduler directly allocates resources to answer questions. Prompt 7 (assessing strength of evidence received, Fig. 6.7(a)) also indicates better scores for the hypothesis scheduler, as expected since the suggested tasking would select strong evidence first. Finally, prompt 8 (adapting to observation conditions, Fig. 6.7(b)) is mostly inconclusive as comparisons between scenarios with poor observation conditions (scenarios B and D) and those with good observation conditions (A, C, and E) do not identify a trend.

Overall, the cognitive support responses confirmed expectations for the state uncertainty, hypothesis uncertainty, and evidence strength questions. However, the lack of conclusiveness in the remainder of the responses lends some insight on potential improvements required in the test scenarios and the overall test apparatus. For instance, clearer indication of how each proposed action addresses the hypotheses would allow for stronger comparisons in how the proposed actions resolve questions. Additionally, scenarios A and C should be improved to include more evidence to avoid exhausting evidence for any partic-

ular hypothesis.

6.6 Conclusions

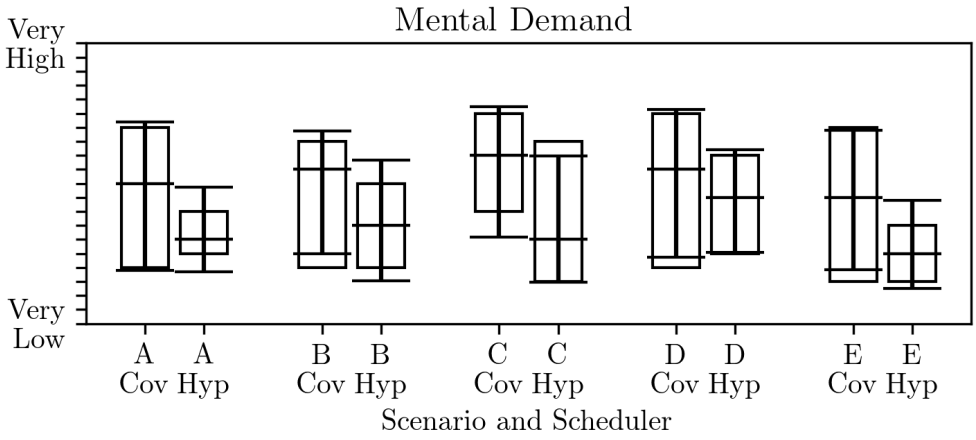
This chapter presents the development and evaluation of a prototype SSA DSS. The DSS design requirements from the cognitive work analysis in Chapter 2 are revisited to highlight five cognitive work requirements addressed by a hypothesis-based sensor tasking algorithm using evidential reasoning, such as JER. The components of the prototype DSS are displayed along with the human-in-the-loop test specifics.

Analysis of the data from these preliminary tests provide indications that the hypothesis-based approach does improve decision-maker mental demand, temporal demand, and effort. The cognitive support questionnaire results are less conclusive but instead illuminated important deficiencies in the preliminary test scenarios. Particularly, certain scenarios do not contain enough evidence to limit the number of actions taken, which meant the hypothesis-based scheduler could collect all the available evidence and then become unresponsive to changes in tasking.

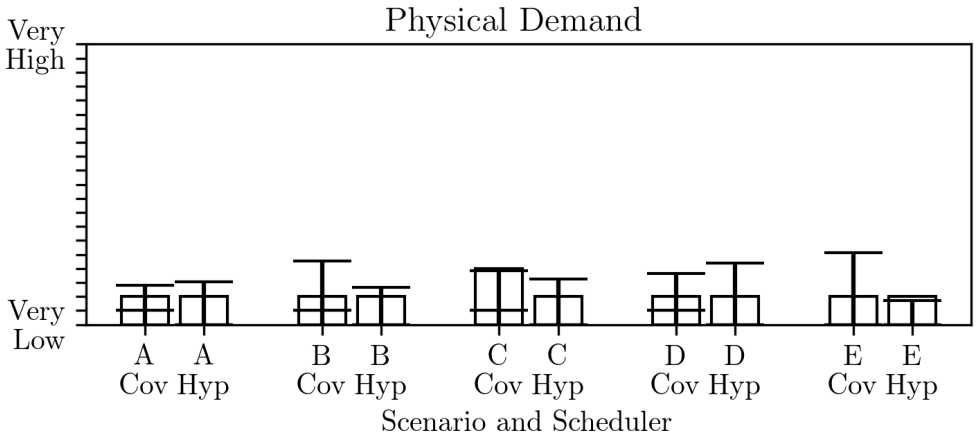
A second round of testing is planned to build on these findings. The test scenarios will be revised to contain more varied evidence and ensure that evidence exhaustion does not occur. Additionally, the 3D orbit display will be deemphasized as it does not directly relate to the hypothesis-resolution testing goals. Though the design was modeled from existing SSA software, this particular task does not rely heavily on the 3D visualization, so it only serves as a distraction from the tasks. Another planned change is to distance the measured hypotheses and the situation awareness hypotheses, asking decision-makers to report their awareness of broader events (e.g. which space object most probably experienced a navigation anomaly?). The preliminary SAGAT questionnaires featured too many level 1 questions, related to perception of elements of the environment, whereas level 2 questions, related to comprehension of the significance of the situation, are desired. This encourages more knowledge-based reasoning in the judgment portion of the task instead of

encouraging memorization of the display elements. This is similar to the approach taken by McGeorge et al. [125]. An objective measure of the operator's performance will also be computed by comparing SAGAT responses to the simulation truth data. This measure is omitted from this preliminary analysis due to oversimplification of the SAGAT questions resulting in very little variation in objective performance between all scenarios and schedulers. Inclusion of this objective performance measurement also allows for interesting comparisons to the operator's own subjective assessment of task performance as measured through the NASA-TLX questionnaire.

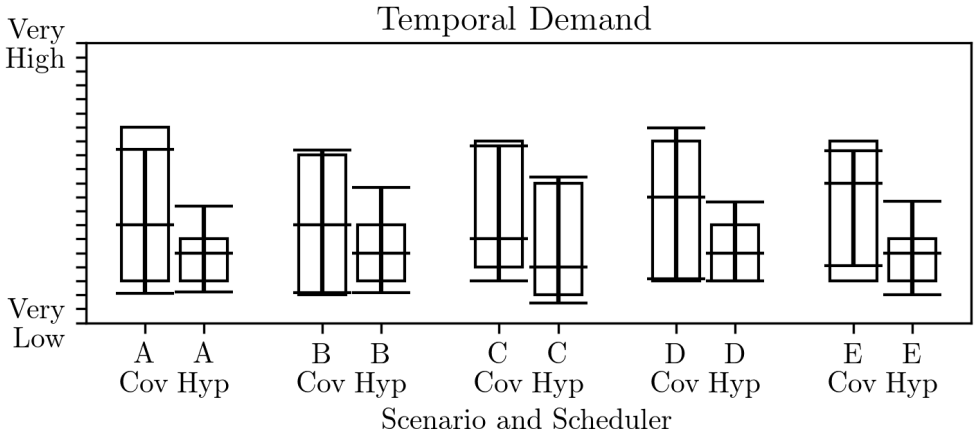
The revised tests are aimed at verifying situation awareness impacts of the hypothesis-based tasking approach and clarifying the cognitive support responses, while the preliminary findings indicate improvements in workload. Importantly, both the preliminary and revised test results should be interpreted with appropriate caveats in light of the limited complexity of laboratory simulation to avoid misleading or invalid conclusions. The results should be used primarily to identify trends for the development of more in-depth human-in-the-loop tests closer to operational scenarios.



(a) Prompt 1

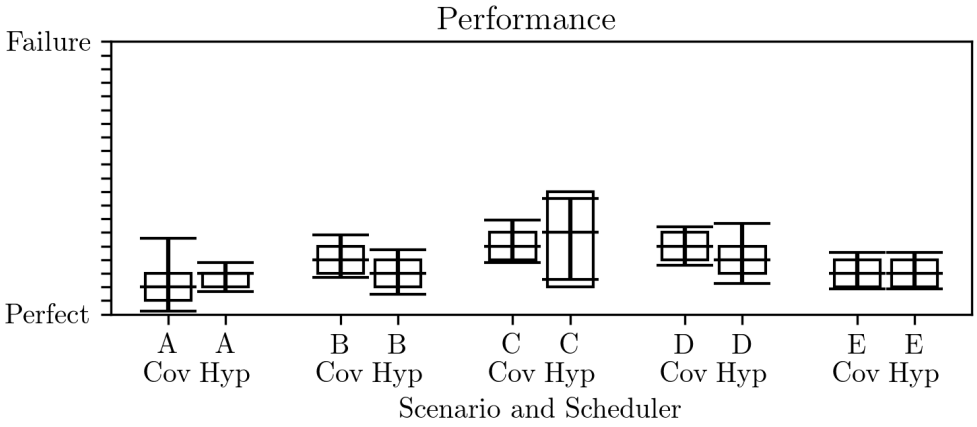


(b) Prompt 2

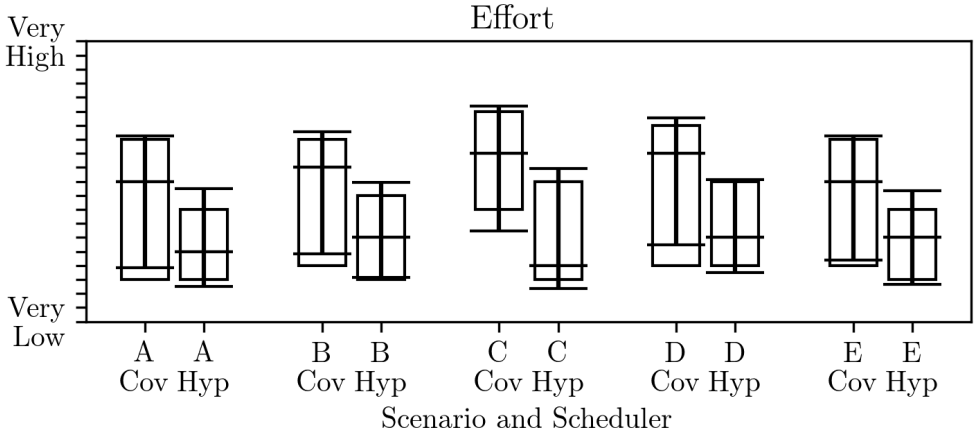


(c) Prompt 3

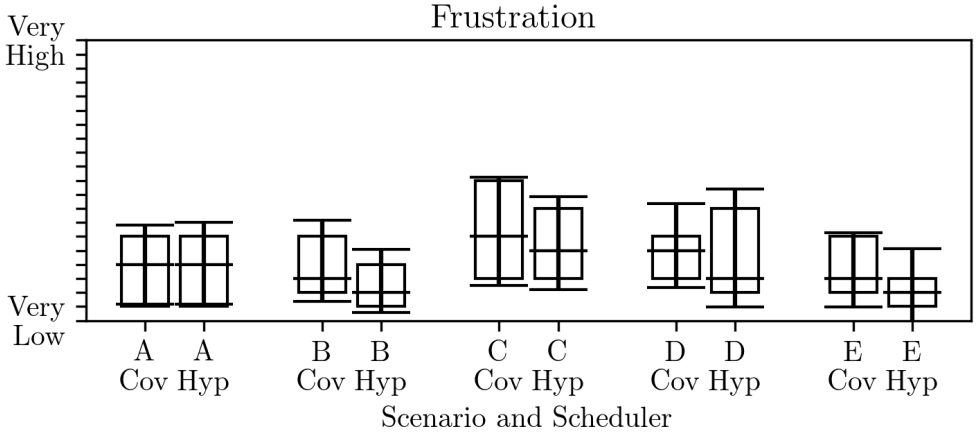
Figure 6.3: NASA-TLX questionnaire response statistics, prompts 1-3



(a) Prompt 4

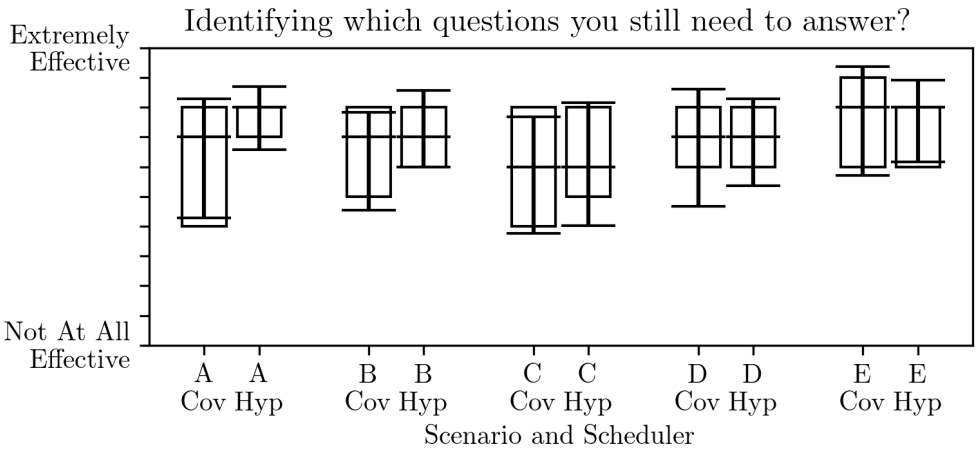


(b) Prompt 5

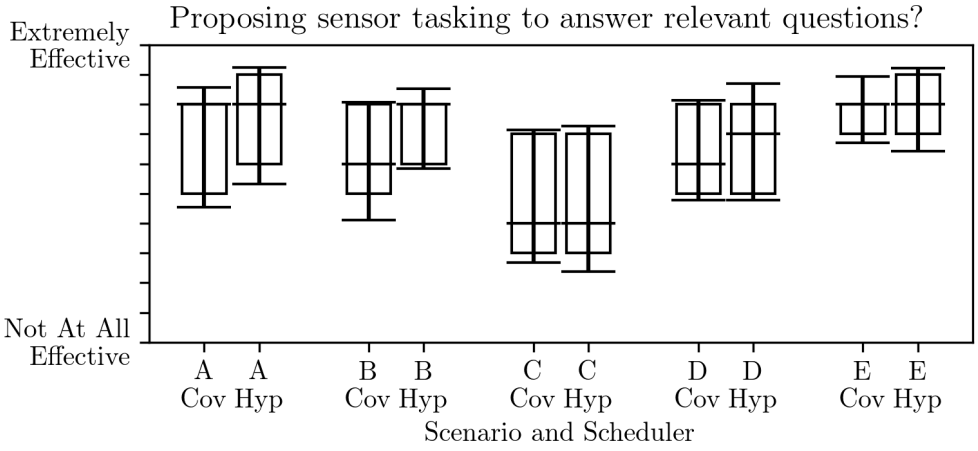


(c) Prompt 6

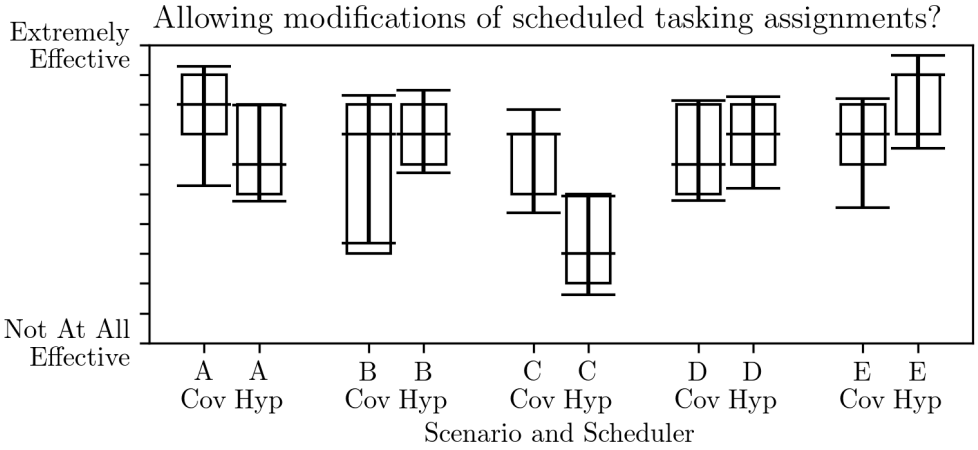
Figure 6.4: NASA-TLX questionnaire response statistics, prompts 4-6



(a) Prompt 1

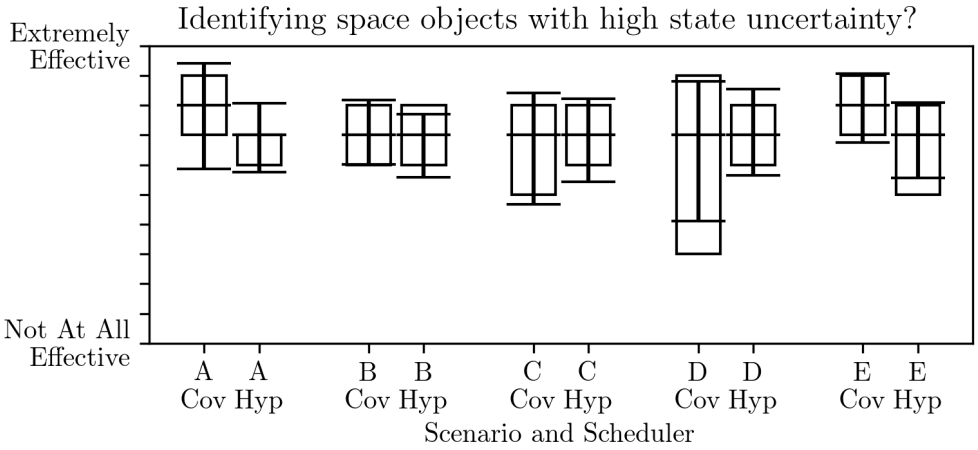


(b) Prompt 2

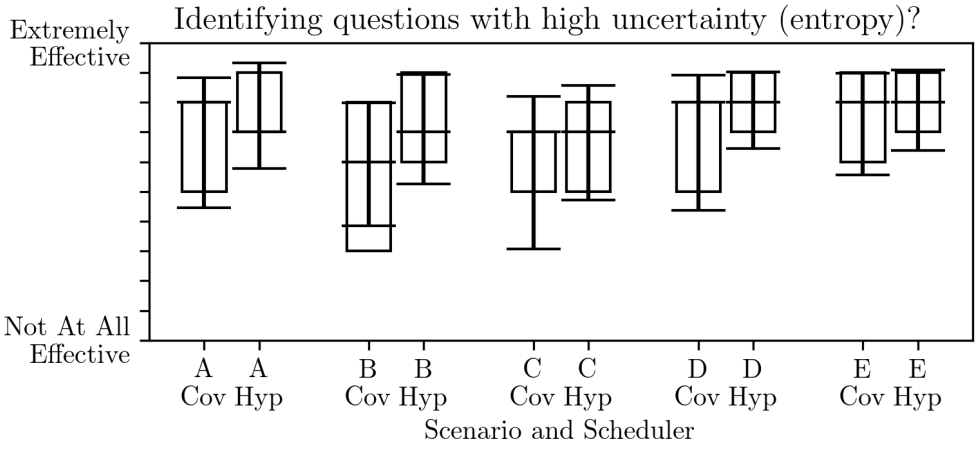


(c) Prompt 3

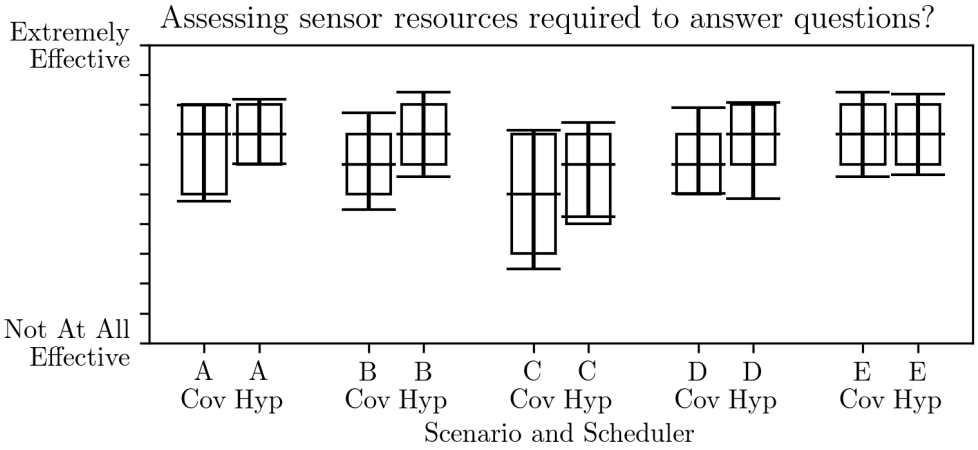
Figure 6.5: Cognitive support questionnaire response statistics, prompts 1-3



(a) Prompt 4

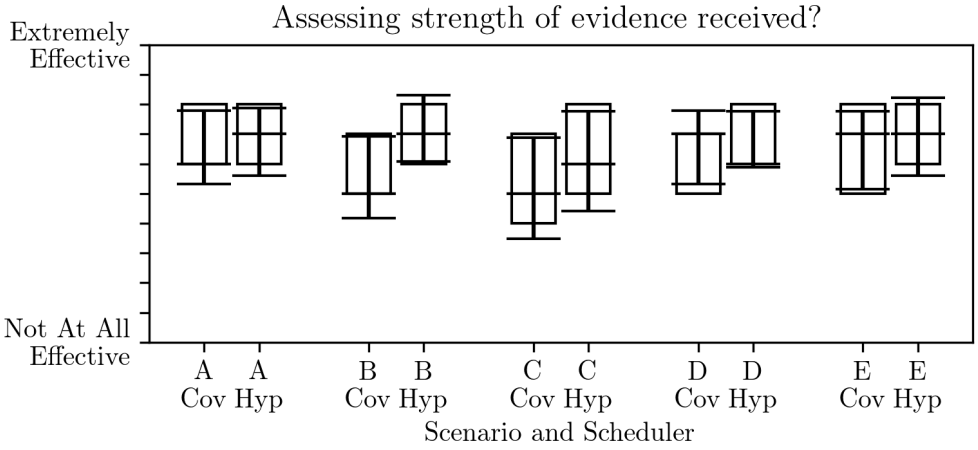


(b) Prompt 5

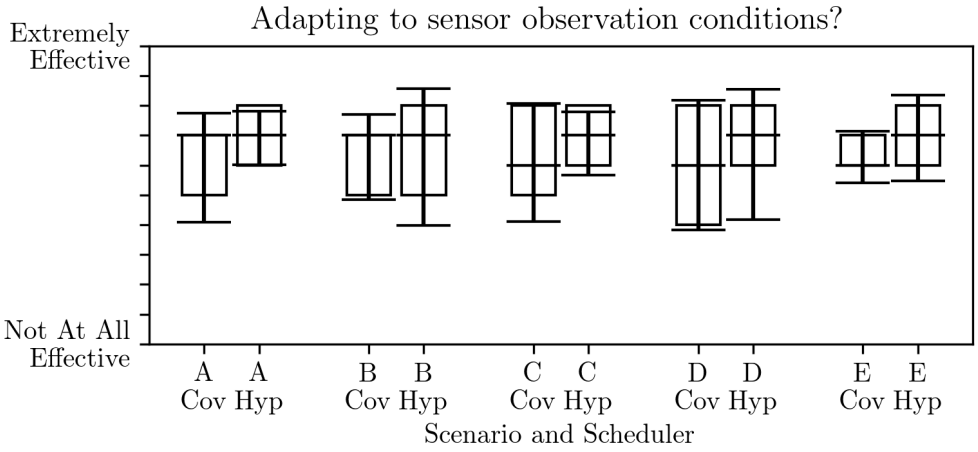


(c) Prompt 6

Figure 6.6: Cognitive support questionnaire response statistics, prompts 4-6



(a) Prompt 7



(b) Prompt 8

Figure 6.7: Cognitive support questionnaire response statistics, prompts 7-8

CHAPTER 7
FUTURE WORK

The work presented in this thesis focuses on improved methods for hypothesis resolution and decision support in space situational awareness to provide decision-quality information, enable predictive tasking, and improve decision-maker situation awareness. The following sections outline promising areas of future research related to this work.

7.1 Non-SSA Applications of JER

The JER approach is intentionally developed for generalized evidence-gathering to resolve hypotheses. The derivation includes non-SSA examples, but future work should apply the full methodology to non-SSA hypothesis resolution tasks. The non-SSA examples included are intentionally simplified to be illustrative, demonstrating the function of both the JER agent-pair optimization and the sub-problem schedule combination. Therefore, future work should include more nuanced analyses of these or other non-SSA scenarios to further validate application outside sensor network tasking.

The medical diagnosis application is a compelling candidate for further development and validation of the JER approach. Medical diagnosis exhibits many similar constraints to SSA hypothesis resolution. Time optimality is a concern as early diagnosis is often a key factor for successful recovery; however, an incorrect or missed diagnosis can be catastrophic, so positive confirmation of all hypothesis resolutions is a priority as well. Additionally, medical test and procedure costs are often very high, limiting the number of actions that may be performed to reach the conclusion. The JER approach was formulated to address these kinds of constrained evidence gathering tasks, where the number of actions available or feasible cannot exhaust all the evidence in the search space. A more detailed and nuanced evaluation of JER used to select diagnostic tests would facilitate translation

from the relatively small field of SSA to a much larger community.

7.2 Improved Confirmation Bias Metrics in JER

As discussed in the development of JER, confirmation bias measurement is an active field of study. Current confirmation bias measurements amount to differences in the number or strength of chosen evidence sources [92, 93], leading to the use of the principle of equal effort for balancing between advocate and critic agent actions. This approach ensures that both agents have equal numbers of actions to gather and present the strongest available evidence. While the approach is ad hoc, it is also necessary to ensure avoid local minima that might be found in an entropy reduction approach, induced by attempting to confirm an incorrect prior belief. The alternating agents also allow application of adversarial optimization approaches, which helps reduce computational complexity.

With the principle of equal effort, the current approach of measuring turns for and against a proposition might be replaced by the amount of time spent on the evidence search, or the number of pieces of evidence returned from the previous actions. Ideally, an information-based metric can be developed to compute and display the confirmation bias inherent in the knowledge state based on previous actions taken and the strength of evidence. If such a metric is developed, the agent turn-ordering can be adjusted to incorporate it: instead of alternating turns, the advocate and critic agent pairs can alternate the currently-active agent based on the current measured confirmation bias. However, this task may prove difficult as confirmation bias is typically measured entirely separate from the actual hypothesis resolution. An actor can be entirely biased toward one proposition and still reach a high-quality resolution if that proposition happens to be correct: in this case, the actor is still experiencing confirmation bias. This is why the ad hoc measurements of confirmation bias use numbers of evidence or subjectively-assigned evidence strengths in laboratory settings, but improved information-based confirmation bias metrics would allow for more nuanced agent-pair scheduling without sacrificing confirmation bias.

7.3 Principle of Equal Effort and Asynchronous Evidence Gathering

Related to the improvement of confirmation bias metrics, the principle of equal effort also suffers when dealing with highly asynchronous action sequences. If a particular evidence-gathering actor has a fast action cadence (e.g. a bot that scans news websites and articles for references of keywords) and another actor has a slow action cadence (e.g. a telescope slewing and making observations), the current formulation of JER would require recomputing and re-estimating action sequences for both the fast and slow actor. This adds many valid sequences for evaluation and adversely affects computational tractability of the approach.

One possible approach is to use sub-intervals for different actors by "fixing" longer time-scale actions and still allow re-evaluation of the short time-scale actions in between. The set of valid actions would only include actions for the non-fixed actors, reducing the search space significantly once again. But here again, the confirmation bias measurement becomes an issue as the principle of equal effort can be affected, especially if the equal effort objective includes time instead of just number of actions. Additionally, since the fast agent takes far more actions in the same timespan, even weak evidence gathered by the fast agent may bias hypothesis resolution. Addressing this asynchronous tasking issue would enable efficient application of JER in even more real-world situations.

7.4 Further Application of Decision Support Requirements

The decision-support prototype developed using cognitive systems engineering addresses a number of the derived decision-support design requirements. Future work can extend upon this prototype to incorporate more decision support design requirements and conduct further tests on situation awareness and workload.

Currently, a follow-on series of tests incorporating lessons-learned from the first prototype is planned. This revised test will modify the situation awareness questions to investigate comprehension (i.e. level 2 situation awareness) as opposed to perception (i.e.

level 1) by abstracting the questions from the available display elements, relying less on memory-based questions.

Additionally, the prototype UI is being altered to deemphasize the 3-dimensional visualization of the space environment. Though industry-standard software for SSA includes these types of displays, the added visualization is only a distraction from the relevant situation awareness questions in this test. By incorporating more complex hypotheses and level 2 situation awareness questions, while also deemphasizing the 3-dimensional visualization, the follow-on tests aim to provide stronger indications of decision-maker cognitive support effectiveness.

7.5 Hypothesis Generation and Evidence Mapping

Though the JER approach considers methods for removing resolved hypotheses from consideration, the topic of generating new hypotheses based on available data was not addressed in this work. Currently, generating the hypotheses and outlining the mappings that transform sensor evidence into BPAs is a very human-intensive process that involves outlining the possible resolutions of each decision-making question and rigorously mapping the incoming evidence to different propositions. Tools can be designed to aid in this process, but the abstract reasoning required to derive the relationships between data or evidence and hypothesis resolution is a knowledge-based reasoning task preferably assigned to the human decision-makers. Future work should investigate methods for supporting human cognition in forming these abstract relationships and generating new hypotheses with associated evidence-to-BPA mappings.

CHAPTER 8
CONCLUSIONS

As research and development efforts in SSA increasingly focus on supporting decision-making, the methods for information fusion and sensor allocation must adapt to support decision-making as well. The number of trackable space objects is increasing and, along with improved sensor technology, contributes to increasing data processing requirements. Current sensor tasking approaches often rely upon human-intensive techniques developed over half a century ago, which are infeasible with increasing SSA data requirements. Proposed improvements often focus on minimizing space object state uncertainty, but the relationships between state uncertainty and decision-maker needs are not always readily apparent, motivating this analysis of sensor tasking and decision support in SSA.

Chapter 2 provides background on cognitive systems engineering practices, particularly cognitive work analysis, and applies these techniques to the SSA domain. Careful consideration of the capabilities and constraints in the SSA work domain and environment leads to the development of abstraction hierarchies to model linkages between purposes, priorities, functions, and resources in SSA. The work domain analysis provides insight that existing sensor allocation and information fusion approaches do not provide a robust or clear mapping to decision-maker goals of space asset safety and national security. The control task analysis focuses on the tasks of information fusion and sensor allocation, leading to the derivation of 14 cognitive work requirements and information relationship requirements. Importantly, this study motivates the development of hypothesis-based methods for information fusion and sensor allocation, which would allow decision-makers to remain at the abstract levels of prioritizing between difference decision-making goals when selecting sensor tasking.

Chapter 3 provides a concrete application of hypothesis-based methods to an existing

problem of spacecraft association and anomaly detection. Using both control distance and Mahalanobis distance metrics, the existing approach for anomaly detection is extended for application to non-Gaussian boundary conditions using Gaussian mixture models. The inclusion of a binary hypothesis testing approach further improves the existing method by more realistically modeling the control cost distributions for a non-maneuver and allowing prescription of allowable rates of false alarm. Control distance is shown to have advantages in anomaly detection consistency over longer observation gaps and computational complexity, while Mahalnobis distance exhibits higher sensitivity in anomaly detection. Overall, this study provides an introduction to applications of hypothesis-based methods to specific SSA tasks.

Chapter 4 continues with spacecraft anomaly detection while also considering the custody problem. Here, evidential reasoning is motivated through its ability to quantify evidence ambiguity, allowing for more realistic modeling of SSA sensor evidence. This motivates the initial evidential reasoning sensor tasking approach based on reducing hypothesis ambiguity, or ignorance. Several SSA sensors (i.e. All-Sky cameras, electro-optical sensors) and auxiliary data sources (i.e. weather forecast reports) are cast as Dempster-Shafer evidence experts and applied in simulated spacecraft custody and anomaly detection cases. The ignorance-reduction criterion does allow for gathering strong evidence and resolves the hypotheses as well as a covariance-based method while using far fewer actions. When limiting the number of actions taken by both algorithms to be equal, the ignorance-reduction method significantly out-performs covariance-minimization in hypothesis resolution.

Chapter 5 generalizes the evidential reasoning application of Chapter 4 to general hypothesis resolution tasks and addresses additional concerns in evidence-gathering for hypothesis resolution. This leads to the development of the Judicial Evidential Reasoning (JER) framework, which has three primary considerations: hypothesis abstraction, ambiguity aversion, and confirmation bias. The hypothesis abstraction is addressed as in Chapter 4 through the application of evidential reasoning. Ambiguity aversion is a well-documented

cognitive phenomenon in decision-making wherein decision-makers overwhelmingly prefer to take bets on known probability distributions than unknown or ambiguous distributions. Bayesian probability theory struggles to capture this difference, casting both distributions as non-informative priors with equal probability. In JER, entropy is used as an optimization criterion as it accounts for both internal conflict and ambiguity in the hypothesis knowledge; reducing both leads to decision-quality hypothesis resolutions. Computational complexity concerns are addressed by decomposing the multi-hypothesis problem into sub-problems, greatly reducing the number of sequence evaluations required as compared to brute force evaluation. The sub-problem decomposition also addresses confirmation bias through the principle of equal effort, wherein the sub-problem schedules are solved on an alternating-turn basis to equally distribute evidence-gathering actions between the propositions. This ensures that prior evidence does not overwhelm the entropy-minimization process. The sub-problem schedules are combined, resolving incongruities by minimizing the priority-weighted total entropy. Examples developed for non-SSA applications illustrate the function of each component of JER, and the simulated results apply JER to a multi-hypothesis anomaly resolution task. Balancing the competing objectives of time-optimality and impartiality enables predictive tasking through JER to resolve hypotheses despite ambiguous evidence, uncertainty, and incorrect priors.

Finally, Chapter 6 returns to the decision support system requirements to develop a prototype for SSA. A primary insight from the cognitive work analysis indicates that hypothesis-based methods allow the decision-maker to reason about sensor tasking options at the abstract function level, without having to resort to physical function level data such as space object covariances. By applying a subset of these requirements relevant to evidential reasoning and the hypothesis-based sensor tasking insights gained through SSA applications in Chapters 3 through 5, the prototype is designed to improve situation awareness and workload. A human-in-the-loop test is conducted to evaluate differences between hypothesis- and covariance-based sensor tasking. The preliminary test results indicate that

the hypothesis-based approach reduces mental demand, temporal demand, and effort in SSA tasks, and follow-on testing is planned to further investigate the cognitive support and situation awareness effects of the hypothesis-based approach.

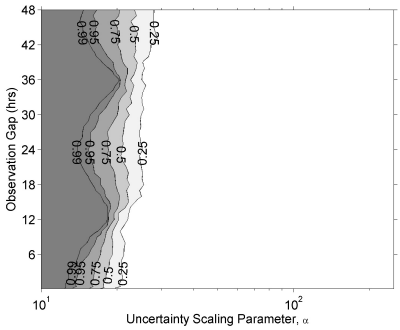
Sensor tasking for SSA is a high-dimensional, multi-objective, non-linear, mixed-integer optimization problem. The work in this thesis formulates the SSA sensor tasking problem to interrogate specific hypothesis that support decision-maker needs. The application of evidential reasoning provides a framework for modeling these hypotheses and evidence while accounting for ambiguity. The specific applications show promise in applying hypothesis-based approaches to a wide array of SSA and activities, with the overal goal of providing decision-quality information, enabling predictive tasking, and improving decision-maker situation awareness and workload in SSA.

Appendices

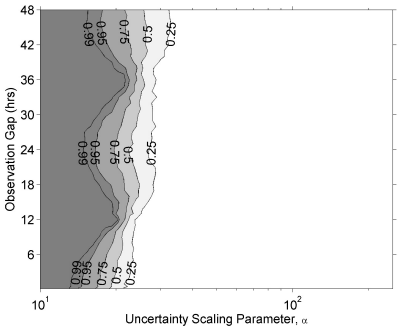
APPENDIX A

ANOMALY DETECTION SENSITIVITY STUDY RESULTS

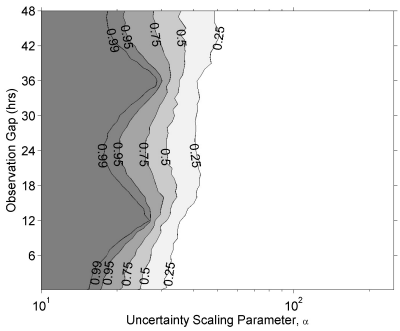
This appendix contains a larger selection of data from the anomaly detection hypothesis testing sensitivity studies. Trends using this data are highlighted in the Simulation Results and Empirical Results sections of Chapter 2.



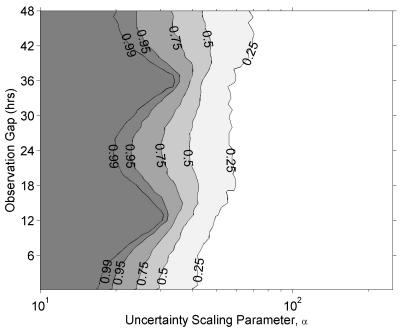
(a) $P_{FA} = 0.005$



(b) $P_{FA} = 0.01$

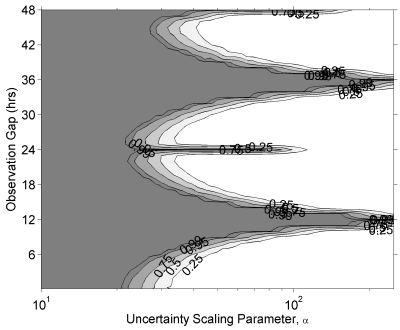


(c) $P_{FA} = 0.05$

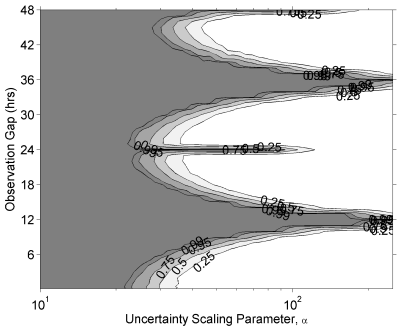


(d) $P_{FA} = 0.1$

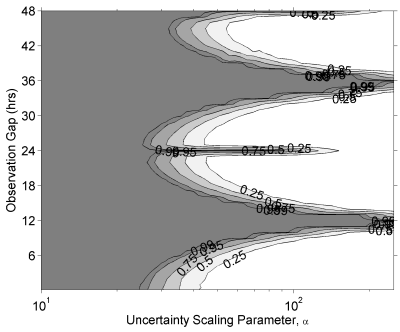
Figure A.1: Control distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated inclination change maneuver.



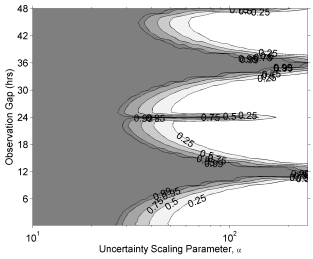
(a) $P_{FA} = 0.005$



(b) $P_{FA} = 0.01$

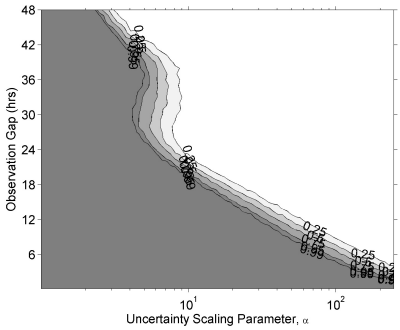


(c) $P_{FA} = 0.05$

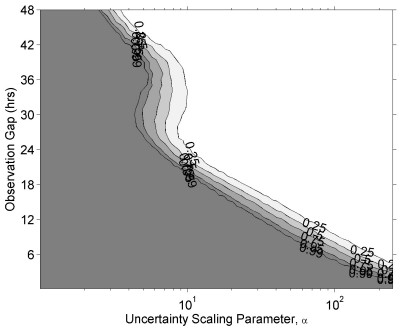


(d) $P_{FA} = 0.1$

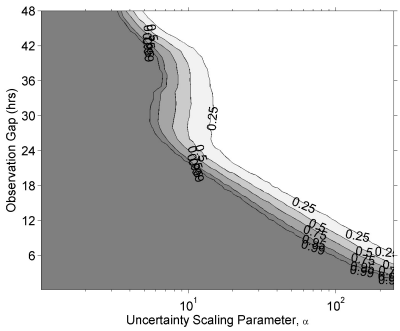
Figure A.2: Mahalanobis distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated inclination change maneuver.



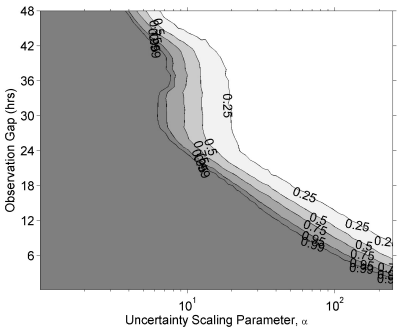
(a) $P_{FA} = 0.005$



(b) $P_{FA} = 0.01$

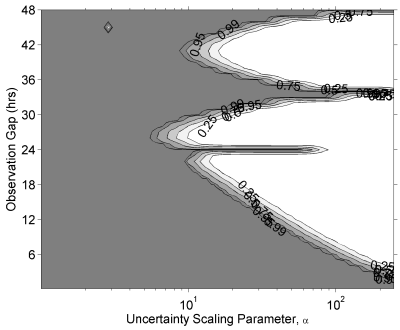


(c) $P_{FA} = 0.05$

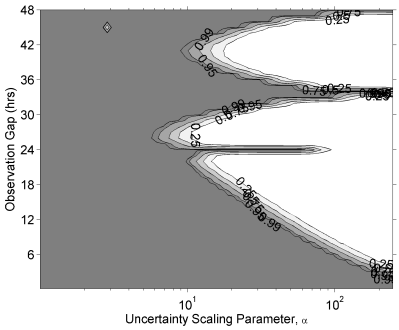


(d) $P_{FA} = 0.1$

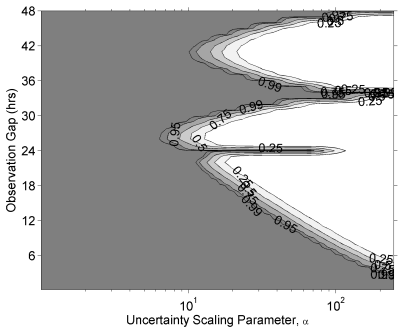
Figure A.3: Control distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated phasing maneuver.



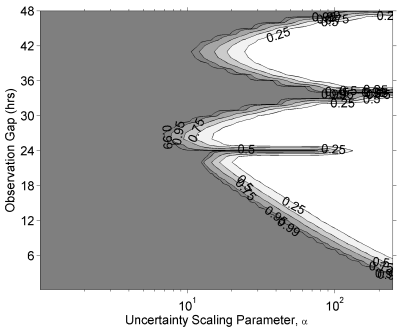
(a) $P_{FA} = 0.005$



(b) $P_{FA} = 0.01$

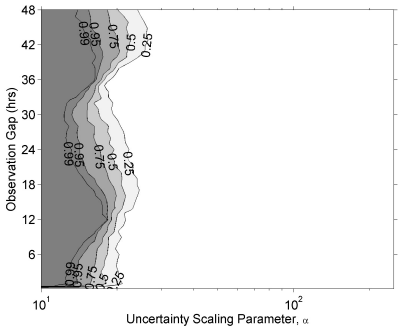


(c) $P_{FA} = 0.05$

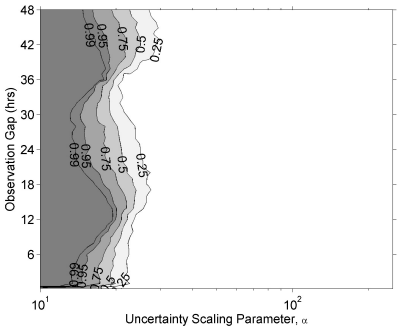


(d) $P_{FA} = 0.1$

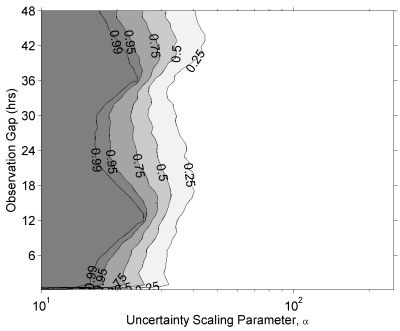
Figure A.4: Mahalanobis distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, simulated phasing maneuver.



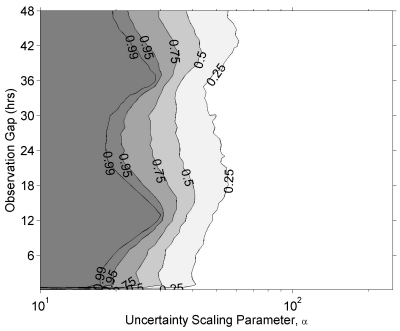
(a) $P_{FA} = 0.005$



(b) $P_{FA} = 0.01$

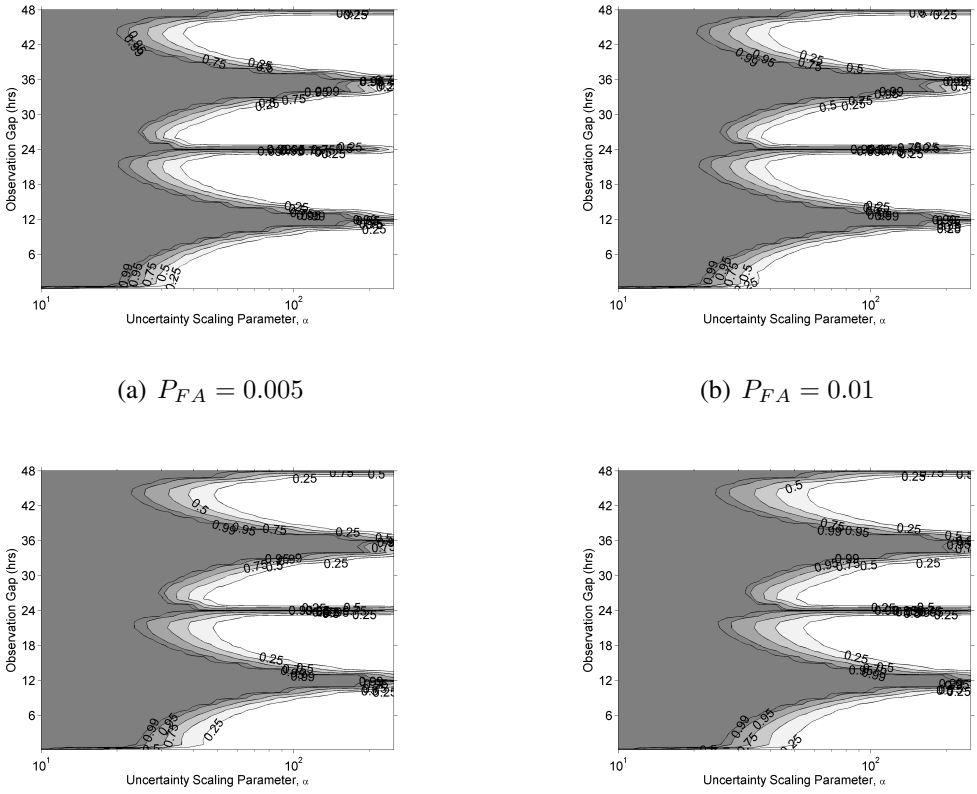


(c) $P_{FA} = 0.05$



(d) $P_{FA} = 0.1$

Figure A.5: Control distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, real-data (WAAS) inclination change maneuver.



(a) $P_{FA} = 0.005$

(b) $P_{FA} = 0.01$

(c) $P_{FA} = 0.05$

(d) $P_{FA} = 0.1$

Figure A.6: Mahalanobis distance probability of anomaly contours vs uncertainty scaling parameter and observation gap, real-data (WAAS) inclination change maneuver.

APPENDIX B

SSA DSS HUMAN-IN-THE-LOOP EXPERIMENT MATERIALS

This appendix contains forms and materials from the SSA DSS human-in-the-loop study.

B.1 Informed Consent

School of Aerospace Engineering
Space Systems Design Lab and Cognitive Engineering Center
Georgia Institute of Technology
Human Subject Consent

1. **Project Title:** Decision Support System Development for Sensor Tasking in Space Situational Awareness

2. **Principal Investigators:** Dr. Karen Feigh, (404) 385-7686, karen.feigh@gatech.edu

   Graduate Students: (Lead) Andris Jaunzemis, (443) 845-2306, adjaunzemis@gatech.edu

3. **Protocol and Consent Title:** Decision Support System Development for Sensor Tasking in Space Situational Awareness

4. **Introduction:** You are being asked to volunteer in a research study. We are seeking your participation in a human in the loop study to evaluate computer support systems for space object tracking. You will be asked to complete a series of test scenarios to help the researchers better understand your experiences as a participant in the simulation to inform the design of future computer support systems.

5. **Purpose:** The purpose of this study is to study your decision making in a simulated space situational awareness task.

6. **Exclusion/Inclusion Criteria:** Participants in this study must be enrolled as graduate students at Georgia Tech and must have completed, or must currently be enrolled in, a graduate-level course in orbital mechanics. Participants must be 18 years or older and must be proficient in English.

7. **Procedures:**
   This study has three phases:
   Training and qualification (1 hour)
   Data collection session 1 (1 hour)
   Data collection session 2 (1 hour)

   If you decide to be in this study, your participation will involve at most three visits.

   The initial visit, which will last no more than 1 hour, will consist of a training session to familiarize you with the operating procedures and tools you will be required to use during the evaluation trials. The session will end with a qualification simulation. If you pass the qualification, you will be invited to return for the data collection sessions.

   The second session, which will last no more than 1 hour, will consist of a series of simulations similar to those presented to you during the initial training session. Additionally, you will be given set of questionnaires both during and after each testing condition.

   The third session, which will last no more than 1 hour, will consist of a series of simulations, similar to those performed in session two, using the prototype computer support system. Similarly to Session 2, you will be asked to complete questionnaires during and after each testing condition.

   The total amount of time you will be in the laboratory is no more than 3 hours, divided into three sessions. Each session will be scheduled at a mutually agreed upon time and your monetary

Figure B.1: Human subject study consent form approved by Georgia Tech Institute Review Board - page 1

compensation will be prorated on how many sessions you complete. Remember, you may stop at any time.

Note: The simulation environment will be recording audio and video of your simulation execution, as well as your interaction with the computer system and simulation environment.

8. **Foreseeable Risks or Discomforts:** The probability and magnitude of harm or discomfort anticipated in the proposed research are not greater than those ordinarily encountered in everyday life or during performance of routine computer based office tasks or coursework.

9. **Benefits:** There are no direct benefits to you for participating in this research study. The results of this study will inform the design and development of future decision support systems for space situational awareness and other fields.

10. **Compensation/Costs:** Participants may be compensated up to $60 upon completion of all three sessions. Participants will be compensated $15 for Session 1, $20 for Session 2 and $25 for Session 3.

11. **Costs to You:** There are no costs to you, other than your time, for being in this study.

12. **Storing and Sharing Your Information:** Your participation in this study is gratefully acknowledged. It is possible that your information/data will be enormously valuable for other research purposes. By signing below, you consent for your de-identified information/data to be stored by the researchers and to be shared with other researchers in future studies. If you agree to allow such future sharing and use, your identity will be completely separated from your information/data. Future researchers will not have a way to identify you.

13. **Use of Audio or Video Recordings:** Audio and video recordings of your simulation execution will be taken during the data-collection scenarios. These files will be stored on the secure (access-controlled and password-protected) laboratory fileserver. Prior to completion of the study, these recordings will be destroyed by secure digital erasure along with any other identifiable information.

14. **Confidentiality:** Your privacy will be protected to the extent required by the law. As a participant, you will be assigned a number and your data will be recorded only in terms of that number. A key linking the number to each participant will be recorded in a password protected file on a password-protected computer during data collection. All electronic data (notes, reports, audio, and video) will be kept in password-protected files on password-protected computers. De-identified information resulting from the testing sessions may be disclosed and published in technical reports, conference papers, journal papers, and academic dissertations. Prior to the end of this study, all data that are identifiable, including the audio and video recordings and the key linking the participant identity to de-identified data, will be destroyed by secure erasure of digital copies or shredding of paper-copies. To make sure that this research is being carried out in the proper way, the GT Office of Research Integrity Assurance may review study records. The Office of Human Research Protections may also look over study records during required interviews. The Georgia Institute of Technology IRB and/or the Office of Human Research Protections may look over study records during required reviews.

15. **Contact Person:** If you have any questions about the research, call or write Andris Jaunzemis at (443) 845-2306 or adjaunzemis@gatech.edu, or the Principal Investigator, Karen Feigh, Ph.D., at (404) 385-7686 or karen.feigh@ae.gatech.edu.

Consent Form Approved by Georgia Tech IRB: December 07, 2017 - December 06, 2018

Figure B.2: Human subject study consent form approved by Georgia Tech Institute Review Board - page 2

16. **Voluntary Participation/Withdrawal:** You have the right to withdraw from the study at any time without penalty. Your recordings, questionnaires, and results will be destroyed upon withdrawal.

17. **In Case of Injury or Harm:** If you are injured as a result of being in this study, please contact the Principal Investigator, Karen Feigh, Ph.D., at (404) 385-7686 or karen.feigh@ae.gatech.edu. Neither the Principal Investigator nor Georgia Institute of Technology has made provision for payment of costs associated with any injury resulting from participation in this study.

18. **Questions about Your Rights as a Research Participant:**
    1. Your participation in this study is voluntary. You do not have to be in this study if you don't want to be.
    2. You have the right to change your mind and leave the study at any time without giving any reason and without penalty.
    3. Any new information that may make you change your mind about being in this study will be given to you.
    4. You will be given a copy of this consent form to keep.
    5. You do not waive any of your legal rights by signing this consent form.

If you have any questions about your rights as a research participant, you may contact:

Ms. Melanie Clark, Georgia Institute of Technology
Office of Research Integrity Assurance, at (404) 894-6942.

Ms. Kelly Winn, Georgia Institute of Technology
Office of Research Integrity Assurance, at (404) 385- 2175.

If you sign below, it means that you have read (or have had read to you) the information given in this consent form, and you would like to be a volunteer in this study.

_____
Participant Name (printed)

_____ _____
Participant Signature Date

_____ _____
Signature of Person Obtaining Consent Date

Consent to Store and Share Your Information:

I agree that my de-identified information/data may be stored and shared for future, unspecified research.

SIGNATURE _____

Figure B.3: Human subject study consent form approved by Georgia Tech Institute Review Board - page 3

REFERENCES

[1] J.-C. Liou, "Modeling the large and small orbital debris populations for environment remediation," NASA Orbital Debris Program Office, Johnson Space Center, Houston, TX, Tech. Rep., Jun. 2014.

[2] "Orbital debris quarterly news," NASA Orbital Debris Program Office, Tech. Rep. 1, Feb. 2018.

[3] C. Cox, E. J. Degraaf, R. J. Wood, and T. H. Crocker, "Intelligent data fusion for improved space situational awareness," in *Space 2005*, 2005.

[4] M. Hart, M. Jah, D. Gaylor, B. C. T. Eyck, E. Butcher, E. L Corral, R. Furfaro, E. H. Lyons, N. Merchant, M. Surdeanu, R. L. Walls, and B. Weiner, "A new approach to space domain awareness at the university of arizona," in *NATO Symposium on Considerations for Space and Space-Enabled Capabilities in NATO Coalition Operations*, 2016.

[5] M. R. Endsley, "Situation awareness global assessment technique (sagat)," in *IEEE 1988 National Aerospace and Electronics Conference*, May 1988.

[6] M. R. Endsley, "Situation awareness analysis and measurement," in, M. R. Endsley and D. J. Garland, Eds. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 2000, ch. Theoretical Underpinnings of Situation Awareness: A Critical Review, pp. 3–29.

[7] J. D. Ianni, D. L. Aleva, and S. A. Ellis, "Overview of human-centric space situational awareness science and technology," Air Force Research Laboratory, Human Effectiveness Directorate, Tech. Rep., Sep. 2012.

[8] J. D. Rendleman and S. M. Mountin, "Responsible ssa coopertion to mitigate on-orbit space debris risks," in *7th International Conference on Recent Advances in Space Technologies (RAST)*, 2015.

[9] T. Kelso, "Analysis of the iridium 33 cosmos 2251 collision," in *Advanced Maui Optical and Space Surveillance Technologies*, Maui, HI, Sep. 2009.

[10] P. A. Brown, "Promoting the safe and responsible use of space: Toward a 21st century transparency framework," *High Frontier*, 5, no., 2008.

[11] F. R. Hoots and P. W. Schumacher, "History of analytical orbit modeling in the u.s. space surveillance system," *Journal of Guidance, Control, and Dynamics*, 27, no., pp. 174–185, 2004.

[12] T. Blake, M. Sanchez, J. Krassner, M. Georgen, and S. Sundbeck, "Space domain awareness," DARPA, Tech. Rep., Sep. 2012.

[13] G. R. Madey, "Enhanced situational awareness: Application of dddas concepts to emergency and disaster management," in *Computational Science - ICCS 2007*, 2007, pp. 1090–1097.

[14] G. R. Madey, M. B. Blake, C. Poellabauer, H. Lu, R. R. McCune, and Y. Wei, "Applying dddas principles to command, control and mission planning for uav swarms," *Procedia Computer Science*, 9, no., pp. 1177–1186, 2012.

[15] R. D. Coder and M. J. Holzinger, "Multi-objective design of optical systems for space situational awareness," *Acta Astronautica*, no., 2016.

[16] T. A. Hobson, "Sensor management for enhanced catalogue maintenance of resident space objects," PhD thesis, University of Queensland Australia, 2014.

[17] C. Daase and O. Kessler, "Knowns and unknowns in the 'war on terror': Uncertainty and the political construction of danger," *Security Dialogue*, 38, no., pp. 411–434,

[18] D. H. Rumsfeld, *Dod news briefing - secretary rumsfeld and gen. myers*, United States Department of Defense News Transcript, Feb. 2002.

[19] J. Luft and H. Ingham, "The johari window, a graphic model of interpersonal awareness," in *Proceedings of the Western Training Laboratory in Group Development*, 1955.

[20] E. D'Appolonia, "Statement of evidence of e. d'appolonia, d'appolonia consulting engineers, pittsburgh, pennsylvania," in *Proceedings of the British Columbia Royal Commission of Inquiry into Uranium Mining, Phase V: Waste Disposal*, 1980.

[21] G. Shafer, "The combination of evidence," *International Journal on Intelligent Systems*, 1, no., pp. 155–179, 1986.

[22] T. Kelecy, D. Hall, K. Hamada, and D. Stocker, "Satellite maneuver detection using two-line element (tle) data," in *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*, 2007.

[23] T. Kelecy and M. Jah, "Detection and orbit determination of a satellite executing low thrust maneuvers," *Acta Astronautica*, 66, no., pp. 798–809, 2010.

[24] M. J. Holzinger, D. J. Scheeres, and K. T. Alfriend, "Object correlation, maneuver detection, and characterization using control distance metric," *Journal of Guidance Navigation and Control*, 35, no., pp. 1312–1325, Jul. 2012.

[25] J. Huang, W.-D. Hu, Q. Xin, and X.-Y. Du, "An object correlation and maneuver detection approach for space surveillance," *Research in Astonomy and Astrophysics*, 12, no., pp. 1402–1416, 2012.

[26] K. J. DeMars, I. I. Hussein, C. Frueh, M. K. Jah, and R. Scott Erwin, "Multiple-object space surveillance tracking using finite-set statistics," *Journal of Guidance, Control, and Dynamics*, no., pp. 1–16, Mar. 2015.

[27] G. Tommei, A. Milani, and A. Rossi, "Orbit determination of space debris: Admissible regions," *Celestial Mechanics and Dynamical Astronomy*, 97, no., pp. 289–304, 2007.

[28] J. M. Maruskin, D. J. Scheeres, and K. T. Alfriend, "Correlation of optical observations of objects in earth orbit," *Journal of Guidance, Control, and Dynamics*, 32, no., pp. 194–209, Jan. 2009.

[29] K. Fujimoto and D. Scheeres, "Correlation of optical observations of earth-orbiting objects and initial orbit determination," *Journal of Guidance, Control, and Dynamics*, 35, no., pp. 208–221, 2012.

[30] K. Hill, K. T. Alfriend, and C. Sabol, "Covariance-based uncorrelated track association," in *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, 2008.

[31] G. A. McIntyre and K. J. Hintz, "An information theoretic approach to sensor scheduling," in *Proceedings of SPIE Signal Processing, Sensor Fusion, and Target Recognition V*, vol. 2755, 1996, pp. 304–312.

[32] A. D. Jaunzemis, M. V. Mathew, and M. J. Holzinger, "Control cost and mahalanobis distance binary hypothesis testing for spacecraft maneuver detection," *Journal of Guidance, Control, and Dynamics*, 39, no., pp. 2058–2072, Sep. 2016.

[33] R. D. Coder, "Multi-objective design of small telescopes and their application to space object characterization," PhD thesis, Georgia Institute of Technology, Aug. 2016.

[34] I. I. Hussein, K. J. DeMars, C. Fruh, M. K. Jah, and R. S. Erwin, "An aegis fisst algorithm for multiple object tracking in space situational awareness," in *AIAA/AAS Astrodynamics Specialist Conference, Guidance, Navigation and Control and Co-located Conferences*, 2012.

[35] C. Fruh and M. Jah, "Initial taxonomy and classification scheme for artificial space objects based on ancestral relation and clustering," in *Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference*, Maui, Hawaii, Sep. 2013.

[36] J. C. Zingarelli, E. Pearce, R. Lambour, T. Blake, C. J. R. Petersen, and S. Cain, "Improving the space surveillance telescope's performance using multi-hypothesis testing," *The Astronomical Journal*, 147, no., pp. 111–124, May 2014.

[37] T. Hardy, S. Cain, J. Jeon, and T. Blake, "Improving space domain awareness through unequal-cost multiple hypothesis testing in the space surveillance telescope," *Applied Optics*, 54, no., pp. 5481–5494, Jun. 2015.

[38] T. S. Murphy, M. J. Holzinger, and B. Flewelling, "Space object detection in images using matched filter bank and bayesian update," *Journal of Guidance, Control, and Dynamics*, 40, no., pp. 497–509, Mar. 2017.

[39] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[40] G. Shafer and R. Logan, "Implementing Dempster's rule for hierarchical evidence," *Artificial Intelligence*, 33, no., pp. 271–298, 1987.

[41] A. P. Dempster, "The Dempster-Shafer calculus for statisticians," *International Journal of Approximate Reasoning*, 48, no., pp. 365–377, 2008.

[42] T. Denoeux, "A k-nearest neighbor classification rule based on dempster-shafer theory," *IEEE Transactions on System, Man, and Cybernetics*, 25, no., pp. 804–813, May 1995.

[43] S. L. Hegarat-Mascle, I. Blouch, and D. Vidal-Madjar, "Application of dempster-shafer evidence theory to unsupervised classification in multisource remote sensing," *IEEE Transactions on Geoscience and Remote Sensing*, 35, no., pp. 1018–1031, Jul. 1997.

[44] C. R. Parikh, M. J. Pont, and N. B. Jones, "Application of dempster-shafer theory in condition monitoring applications: A case study," *Pattern Recogniztion Letters*, no., 2001.

[45] B.-S. Yang and K. J. Kim, "Application of dempster-shafer theory in fault diagnosis of induction motors using vibration and current signals," *Mechanical Systems and Signal Processing*, 20, no., pp. 403–420, Feb. 2006.

[46] W. F. Caselton and W. Luo, "Decision making with imprecise probabilities: Dempster-shafer theory and application.," *Water Reources Research*, 28, no., pp. 3071–3083, Dec. 1992.

[47] P. Orponen, "Dempster's rule is sharp-p-complete," *Artificial Intelligence*, 44, no., pp. 245–253, 1990.

[48] J. Gordon and E. H. Shortliffe, "A method for managing evidential reasoning in a hierarchical hypothesis space," *Artificial Intelligence*, 26, no., pp. 323–357, 1985.

[49] V. Kreinovich, A. Bernat, W. Borrett, Y. Mariscal, and E. Villa, "Monte-carlo methods make dempster-shafer formalism feasible," NASA, Tech. Rep. NASA-CR-192945, 1994.

[50] D. Dubois and H. Prade, "Consonant approximations of belief functions," *International Journal of Approximate Reasoning*, 4, no., pp. 419–449, 1990.

[51] M. Bauer, "Approximation algorithms and decision making in the dempster-shafer theory of evidence - an empirical study," *International Journal of Approximate Reasoning*, 17, no., pp. 217–237, 1997.

[52] J. R. Boyd, "A discourse on winning and losing," Maxwell Air Force Base, Briefing slides Air University Library Document No. M-U 43947, 1987.

[53] W. S. Angerman, "Coming full circle with boyd's ooda loop ideas: An analysis of innovation diffusion and evolution," Master's thesis, Air Force Institute of Technology, Mar. 2004.

[54] G. L. Nemhauser and L. A. Wolsey, *Integer and Combinatorial Optimization*. Wiley, 1999.

[55] B. R. Cobb and P. P. Shenoy, "On the plausibility transformation method for translating belief function models to probability models," *International Journal of Approximate Reasoning*, 41, no., pp. 314–330, Apr. 2006.

[56] B. Tessem, "Approximations for efficient computation in the theory of evidence," *Artificial Intelligence*, 61, no., 1993.

[57] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, 66, no., pp. 191–234, 1994.

[58] L. G. Militello, C. O. Dominguez, G. Lintern, and G. Klein, "The role of cognitive systems engineering in the systems engineering design process," *Journal of Systems Engineering*, 13, no., pp. 261–273, 2010.

[59] K. J. Vicente, *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. CRC Press, 1999.

[60] D. Oltrogge, S. Alfana, and R. Gist, "Satellite mission operations improvements through covariance based methods," in *AIAA 2002-1814, SatMax 2002: Satellite Performance Workshop*, Apr. 2002.

[61] A. Ray, "Symbolic dynamic analysis of complex systems for anomaly detection," *Mechanical Systems and Signal Processing*, 21, no., 2007.

[62] J. T. Horwood, N. D. Aragon, and A. B. Poore, "A gaussian sum filter framework for space surveillance," in *Signal and Data Processing of Small Targets*, vol. 8137, San Diego, CA, 2011.

[63] G. Gallego, C. Cuevas, R. Mohedano, and N. Garcia, "On the mahalanobis distance classification criterion for multidimensional normal distributions," *IEEE Transactions on Signal Processing*, 61, no., pp. 4387–4396, Jun. 2013.

[64] M. J. Holzinger, "Delta-v metrics for object correlation, maneuver detection, and maneuver characterization," in *AIAA Guidance, Navigation and Control Conference*, Portland, OR, Aug. 2011.

[65] Q. Yan, "Distributed signal detection under the neyman-pearson criterion," *IEEE Transactions on Information Theory*, 47, no., pp. 1368–1377, 2001.

[66] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society A: Mathetmatical, Physical and Engineering Sciences*, 231, no., pp. 239–337, 1933.

[67] J. E. Prussing and B. A. Conway, *Orbital Mechanics*. Oxford University Press, 2012.

[68] A. D. Jaunzemis, M. Mathew, and M. J. Holzinger, "Control metric maneuver detection with gaussian mixtures and real data," in *25th AAS/AIAA Space Flight Mechanics Meeting*, Williamsburg, VA, Jan. 2015.

[69] M. J. Holzinger, D. J. Scheeres, and K. T. Alfriend, "Delva-v distance object correlation and maneuver detection with dynamis parameter uncertainty and generalized constraints," in *Proceedings of the 219th American Astronomical Society Meeting*, Austin, TX, Jan. 2012.

[70] X. Q. Li and I. King, "Gaussian mixture distance for information retrieval," in *IEEE International Joint Conference on Neural Netrowks*, vol. 4, 1999.

[71] M. J. Holzinger, C. J. Wetterer, K. K. Luu, C. Sabol, and K. Hamada, "Photometric attitude estimation for agile space objects with shape uncertainty," *Journal of Guidance, Control, and Dynamics*, 37, no., pp. 921–932, May 2014.

[72] R. D. Coder, M. J. Holzinger, and R. Linares, "3dof estimation of agile space objects using marginalized particle filters," *Journal of Guidance, Control, and Dynamics*, no., Submitted 2016.

[73] T. S. Murphy, M. J. Holzinger, K. K. Luu, and C. Sabol, "Generalized minimum-time follow-up approaches applied to electro-optical sensor tasking," *IEEE Transactions on Aerospace and Electronic Systems*, no., 2017 (submitted).

[74] R. Jirousek and P. P. Shenoy, "A new definition of entropy of belief functions in the Dempster-Shafer theory," *International Journal of Approximate Reasoning*, 92, no., pp. 49–65, 2018.

[75] R. R. Yager and L. Liu, *Classic Works of the Dempster-Shafer Theory of Belief Functions*. Springer, 2008.

[76] N. Kami, "Integrated formulation of the theory of belief functions from a linear algebraic perspective," in *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, Kongresshaus Baden-Baden, Germany, Sep. 2016.

[77] K. Sentz and S. Ferson, "Combination of evidence in dempster-shafer theory," Systems Science, Industrial Engineering Department, Thomas J. Watson School of Engineering, and Applied Science, Tech. Rep., 2002.

[78] L. A. Zadeh, "A simple view of the dempster-shafer theory of evidence and its implication for the rule of combination," *AI Magazine*, 7, no., pp. 85–90, 1986.

[79] R. R. Yager, "Arithmetic and other operations on dempster-shafer structures," *International Journal of Man-Machine Studies*, 25, no., pp. 357–366, 1986.

[80] ——, "On the dempster-shafer framework and new combination rules," *Information Sciences*, 41, no., pp. 93–137, 1987.

[81] N. Wilson, "A monte-carlo algorithm for dempster-shafer belief," in *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*, 1991, pp. 414–417.

[82] A. D. Jaunzemis, D. Minotra, M. J. Holzinger, K. M. Feigh, M. W. Chan, and P. P. Shenoy, "Judicial evidential reasoning for decision support applied to orbit insertion failure," in *1st IAA Conference on Space Situational Awarness*, Orlando, FL, Nov. 2017.

[83] Y.-T. Hsia, "Characterizing belief with minimum commitment," in *IJCAI'91 Proceedings of the 12th International Joint Conference on Artificial Intelligence*, vol. 2, 1991, pp. 1184–1189.

[84] M. J. Holzinger and D. J. Scheeres, "Reachability results for nonlinear systems with ellipsoidal initial sets," *IEEE Transactions on Aerospace and Electronic Systems*, 48, no., pp. 1583–1600, Apr. 2012.

[85] J. Brew and M. J. Holzinger, "Reachability subspace exploration using continuation methods," in *27th AAS/AIAA Space Flight Mechanics Meeting*, San Antonio, TX, Feb. 2017.

[86] B. K. Horn and B. G. Schunck, "Determining optical flow," in *Techniques and Applications of Image Understanding*, J. J. Pearson, Ed., vol. 281, Washington, D.C., Apr. 1981.

[87] R. H. Garstang, "Night sky brightness at observatories and sites," *Publications of the Astronomical Society of the Pacific*, 101, no., pp. 306–329, 1989.

[88] K. Krisciunas and B. E. Schaefer, "A model of the brightness of moonlight," *Publications of the Astronomical Society of the Pacific*, 103, no., pp. 1033–1039, Sep. 1991.

[89] R. D. Coder, A. D. Jaunzemis, M. V. Mathew, J. L. Worthy, and M. J. Holzinger, "The georgia tech space object research telescope (gt-sort)," *Journal of Spacecraft and Rockets*, no., Submitted 2017.

[90] A. D. Jaunzemis and M. J. Holzinger, "Evidence-based sensor tasking for space domain awareness," in *Advanced Maui Optical and Space Surveillance Technologies*, Kihei, HI, Sep. 2016.

[91] D. Frey and M. Rosch, "Information seeking after decisions," *Personality and Social Psychology Bulletin*, 10, no., pp. 91–98, Mar. 1984.

[92] E. Jonas, S. Schulz-Hardt, D. Frey, and N. Thelen, "Confirmation bias in sequential information search after preliminary decisions: An expansion of dissonance theoretical research on selective exposure to information," *Journal of Personality and Social Psychology*, 80, no., pp. 557–571, Apr. 2001.

[93] M. B. Cook and H. S. Smallman, "Human factors of the confirmation bias in intelligence analysis: Decision support from graphical evidence landscapes," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50, no., pp. 745–754, Oct. 2008.

[94] D. Ellsberg, "Risk, ambiguity, and the Savage axioms," *The Quarterly Journal of Economics*, 75, no., pp. 643–669, Nov. 1961.

[95] P. P. Shenoy and G. Shafer, "Axioms for probability and belief-function propagation," in *Uncertainty in Artificial Intelligence*, R. D. Shachter, T. Levitt, J. F. Lemmer, and L. N. Kanal, Eds., vol. 4, Amsterdam: Elsevier, 1990, pp. 169–198.

[96] D. E. Knuth and R. W. Moore, "An analysis of alpha-beta pruning," *Artificial Intelligence*, 6, no., pp. 293–326, 1975.

[97] R. D. Coder, C. J. Wetterer, K. M. Hamada, M. J. Holzinger, and M. K. Jah, "Inferring active control mode of the Hubble space telescope using unresolved imagery," *Journal of Guidance, Control, and Dynamics (accepted)*, no., 2017.

[98] G. Shafer, "Belief functions and parametric models," *Journal of the Royal Statistical Society. Series B*, 44, no., pp. 322–352, 1982.

[99] A. D. Jaunzemis, M. J. Holzinger, and K. K. Luu, "Sensor tasking for spacecraft custody maintenance and anomaly detection using evidential reasoning," *Journal of Aerospace Information Systems*, no., 2018.

[100] A. Kirlik and R. Strauss, "Situation awareness as judgment i: Statistical modeling and quantitative measurement," *International Journal of Industrial Ergonomics*, 36, no., pp. 463–474, May 2006.

[101] R. Strauss and A. Kirlik, "Situation awareness as judgment II: Experimental demonstration," *International Journal of Industrial Ergonomics*, 36, no., pp. 475–484, May 2006.

[102] R. M. Taylor and S. J. Selcon, "Cognitive quality and situational awareness with advanced aircraft attitude displays," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 34, 1990, pp. 26–30.

[103] R. W. Pew, "Situation awareness analysis and measurement," in, M. R. Endsley and D. J. Garland, Eds. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers, 2000, ch. The State of Situation Awareness Measurement: Heading Toward the Next Century, pp. 29–45.

[104] M. J. Miller, K. M. McGuire, and K. M. Feigh, "Decision support system requirements definition for human extravehicular activity based on cognitive work analysis," *Journal of Cognitive Engineering and Decision Making*, 11, no., pp. 136–165, 2017.

[105] P. Keen, "Decision support systems: The next decade," *Decision Support Systems*, 3, no., pp. 253–265, 1987.

[106] J. P. Shim, M. Warkentin, J. F. Courney, D. J. Power, R. Sharda, and C. Carlsson, "Past, present, and future of decision support technology," *Decision Support Systems*, 33, no., pp. 111–126, 2002.

[107] P. J. Smith, K. B. Bennett, and R. B. Stone, "Representation aiding to support performance on problem-solving tasks," *Reviews of Human Factors and Ergonomics*, 2, no., pp. 74–108, 2006.

[108] P. J. Smith, D. Johnson, S. Pruchnicki, J. Schimmel, A. Spencer, and S. Young, "Cognitive engineering considerations in the development of an information retrieval system: Avoid fixation on technological substitutes," *Journal of Cognitive Engineering and Decision Making*, 10, no., pp. 126–130, Jun. 2016.

[109] E. Hollnagel and D. D. Woods, "Cognitive systems engineering: New wine in new bottles," *International Journal of Man-Machine Studies*, 18, no., pp. 583–600, 1983.

[110] J. Rasmussen and A. M. Pejtersen, *Cognitive Systems Engineering*. New York: Wiley, 1994.

[111] H. Beyer and K. Holtzblatt, *Contextual Design: Defining Customer-Centered Systems*. San Francisco, CA: Morgan Kauffman Kluwer Academic Publishers, 1998.

[112] J. Annett, "Handbook of cognitive task design," in, E. Hollnagel, Ed. Boca Raton, FL: CRC Press, 2003, ch. Hierarchical Task Analysis, pp. 17–36.

[113] G. Klein, "Naturalistic decision making," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50, no., pp. 456–460, 2008.

[114] M. J. Miller, "Decision support system development for human extravehicular activity," PhD thesis, Georgia Institute of Technology, Dec. 2017.

[115] J. Rasmussen, "The role of hierarchical knowledge representation in decision and system management," *IEEE Transactions on Systems Man and Cybernetics*, 15, no., pp. 234–243, 1985.

[116] A. Bisantz and E. M. Roth, "Analysis of cognitive framework," *Reviews of Human Factors and Ergonomics*, 3, no., pp. 1–43, Nov. 2008.

[117] R. C. McIlroy and N. A. Stanton, "Ecological interface design two decades on: Whatever happended to the srk taxonomy?" *IEEE Transactions on Human-Machine Systems*, 45, no., pp. 145–163, 2015.

[118] N. Naikar, *Work Domain Analysis: Concepts, Guidelines, and Cases*. CRC Press, 2013.

[119] C. M. Burns, D. J. Bryant, and B. A. Chalmers, "Boundary, purpose, and values in work-domain models: Models of naval command and control," *Systems, Man, and Cybernetics, Part A: Systems and Humans*, 35, no., pp. 603–616, 2005.

[120] D. Aleva, J. Ianni, and V. Schmidt, "Space situation awareness human effectiveness research trends," in *World Congress in Computer Science, Computer Engineering and Applied Computing*, 2012.

[121] D. Aleva and J. McCracken, "Jspoc cognitive task analysis," in *Advanced Maui Optical and Space Surveillance Technologies Conference*, Maui, HI, Sep. 2009.

[122] J. Rasmussen, "Skills, rules, knowledge; signals, signs, and symbols, and other distinctions in human performance models," *IEEE Transactions on Systems, Man, and Cybernetics*, 13, no., pp. 257–266, May 1983.

[123] S. S. Potter, J. W. Gualtieri, and W. C. Elm, "Handbook for cognitive task design," in. London, UK: Lawrence Erlbaum Associates, 2003, ch. Case Studies: Applied Cognitive Work Analysis in the Design of Innovative Decision Support, pp. 1–32.

[124] W. C. Elm, S. S. Potter, J. W. Gualtieri, E. M. Roth, and J. R. Easter, "Handbook for cognitive task design," in. London, UK: Lawrence Erlbaum Associates, 2003, ch. Applied Cognitive Work Analysis: A Pragmatic Methodology for Designing Revolutionary Cognitive Affordances, pp. 357–382.

[125] N. McGeorge, S. Hedge, R. L. Berg, T. K. Guarrera-Schick, D. T. LaVergne, S. N. Casucci, A. Z. Hettinger, L. N. Clark, L. Lin, R. J. Fairbanks, N. C. Benda, L. Sun, R. L. Wears, S. Perry, and A. Bisantz, "Assessment of innovative emergency department information displays in a clinical simulation center," *Journal of Cognitive Engineering and Decision Making*, 9, no., pp. 329–346, Dec. 2015.

[126] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Advances in Psychology*, 52, no., pp. 139–183, 1988.