Deriving Reliable Model Revisions from Executed Plan Data Analysis

Alice M. Mulvehill, Brian Krisler and Renu Bostwick

BBN Technologies

Presentation Overview

***JAGUAR**: **J**oint **A**ir-to-**G**round **U**nified, **A**daptive **R**eplanning*

- JAGUAR Overview
- Model Adaptor Overview
- Experiment Results
- Summary

What is JAGUAR

JAGUAR: Joint Air-to-Ground Unified, Adaptive Replanning

- JAGUAR is a set of state-of-the-art tools that have been developed to generate plans for and dynamically refine air missions that are published in the Air Tasking Order (ATO).

- JAGUAR has been developed to adapt to changes in the environment
  - Replanning
  - Adaptive modeling

Joint Air Ground Operations: Unified, Adaptive, Replanning (JAGUAR)

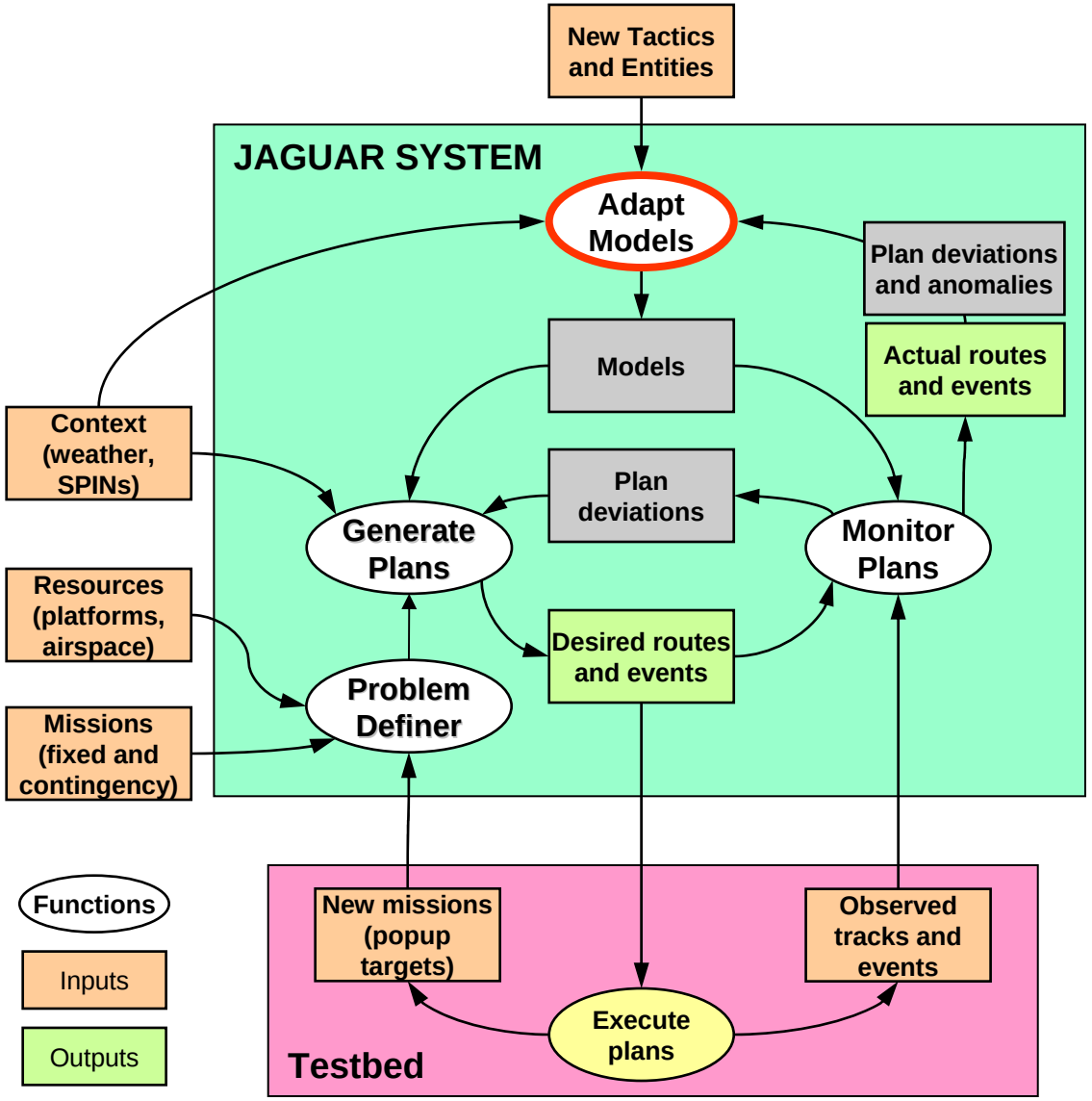
JAGUAR: Joint Air-to-Ground Unified, Adaptive Replanning

JAGUAR COMPONENTS

Model Adaptor: Updates models of platforms and tactics.

Plan Generator: Creates, refines and adjusts air operations plans

Plan Monitor: Correlates reports to plans, anticipates downstream problems.



Model Adaptor Functionality

- Modeling
  - Model Creation
  - Model Publication
  - Model Management
  - Model Repair and Update

- Model Diagnosis
  - Diagnosis of *anomalies* with respect to model errors, or un-modeled world state entities and processes
  - Supports both real time and post mission analysis
  - Provides suggestions for model repair
  - Supports publication of *Model Updates*

- Storage and annotation of historical mission data for *reuse* by Jaguar Components.

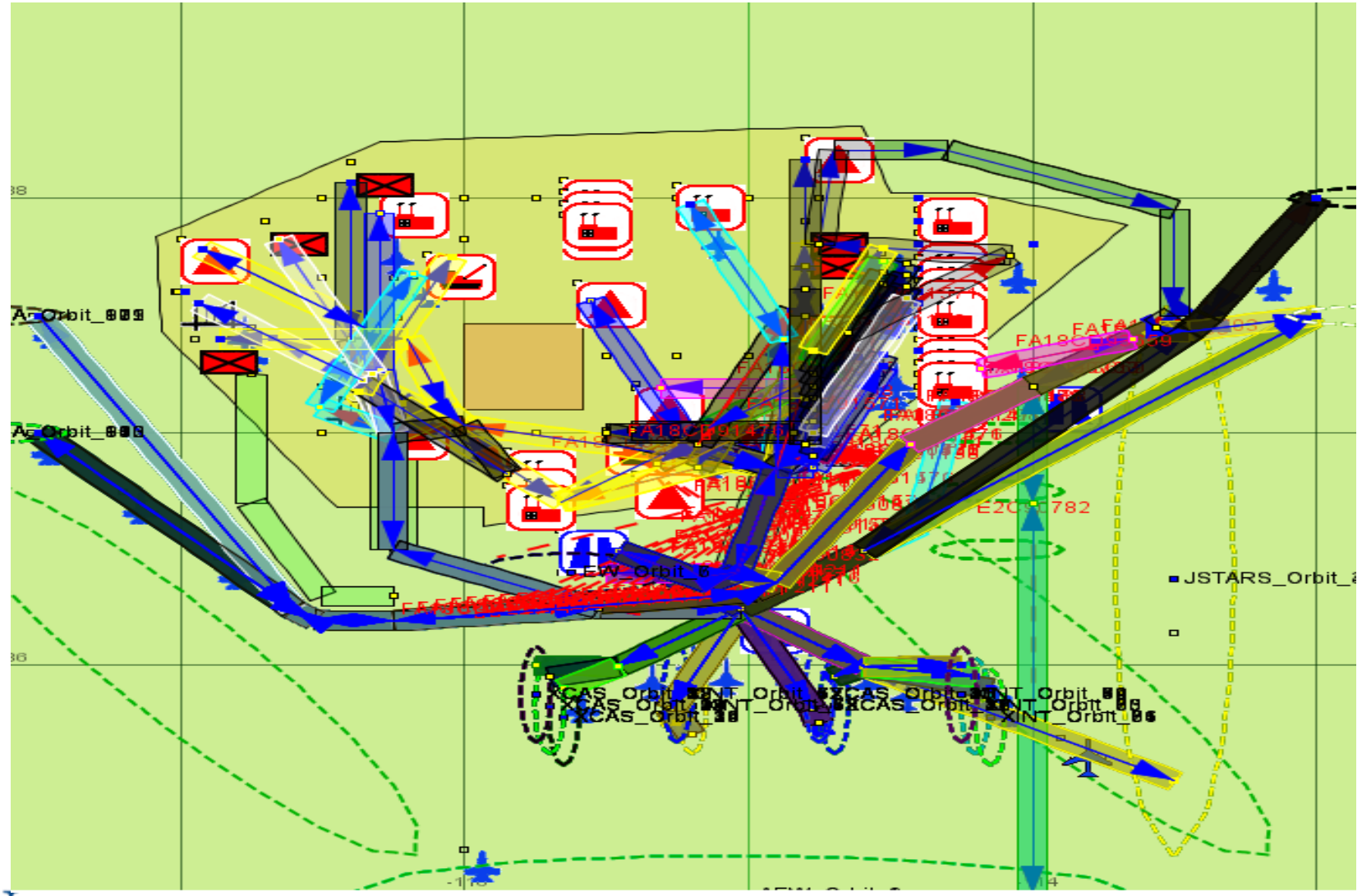
- Analysis and usage of executed mission data for model adaptation, e.g., emergent aircraft capabilities and plan assessment.

Model Adaptation

- Components in JAGUAR are model-driven. Changes to the models do not necessarily require code changes.
- Automated tools support model creation, diagnosis, assessment and repair.
- Model quality and usage can be evaluated through an assessment of the executed plans.

Mission Execution

***JAGUAR**: **J**oint **A**ir-to-**G**round **U**nified, **A**daptive **R**eplanning*



Problem Space

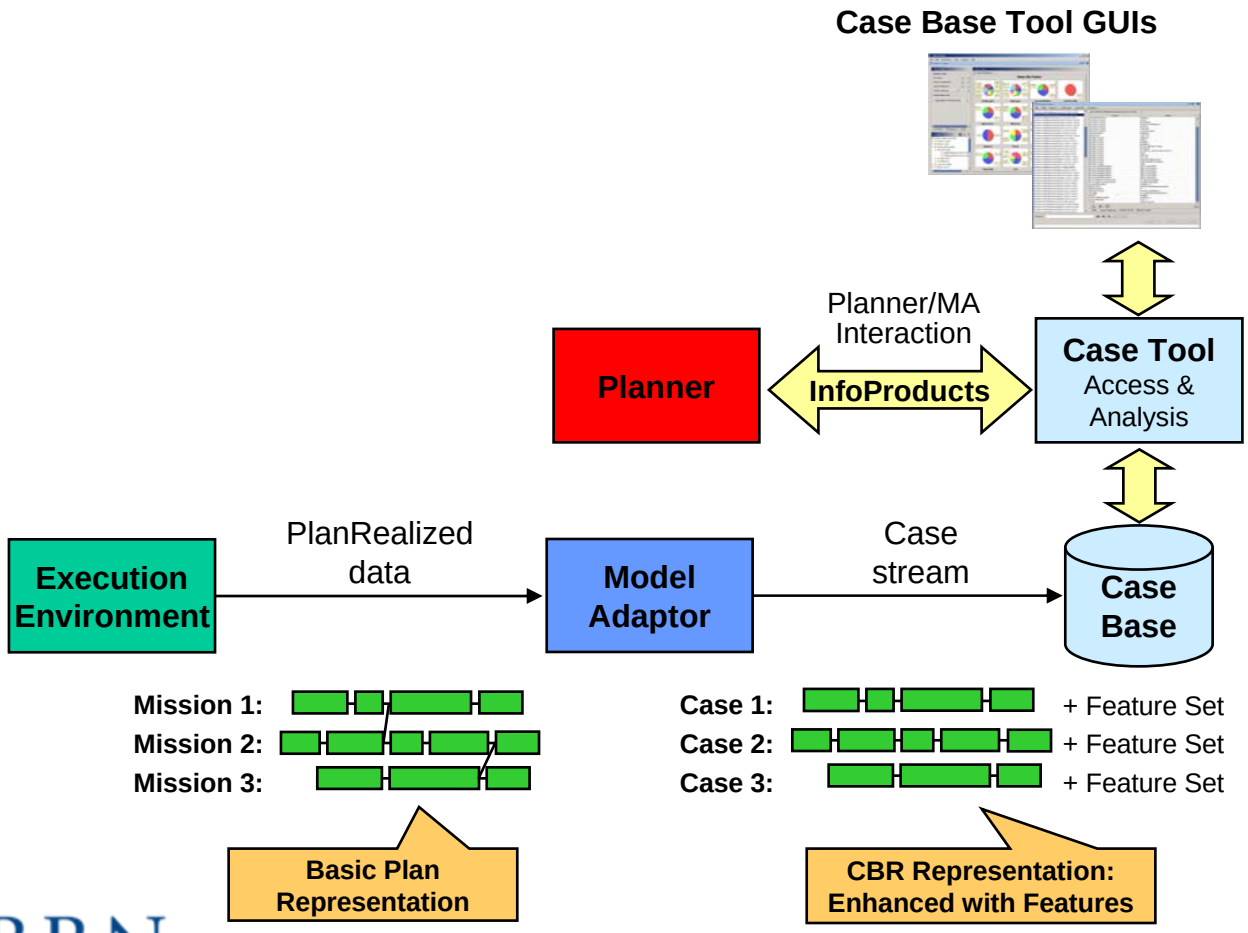
- Situation Space
  - Hundreds of missions executing in parallel.
  - Potentially thousands of anomalies can be detected by the Plan Monitor, some minor and some serious.
- Monitoring Approach:
  - Envelopes included in the plan are used by the Plan Monitor and when unplanned behavior occurs, anomalies are published.
  - Severe anomalies become deviations which require replanning.
- Model Complexity
  - Modeling errors can be the source of some of the anomalies.
  - Detecting the source of a modeling error can be difficult since the models are inter-related and execution problems can be a function of multiple aspects of the models, the component software or of the simulation environment.

Approach to Using Executed Plan Assessment to Support Model Revision

- Premise: Access to the performance of historical executed mission data should be useful for informing a plan generator of good and bad choices.
- Case-Based Reasoning (CBR) technology offers mechanisms to store and reason about historical data.
- The JAGUAR Model Adaptor utilizes CBR technology to store and analyze many aspects of the JAGUAR planning and execution environment. This includes:
  - Models
  - Executed Missions
  - Objectives
  - Anomalies
- Several JAGUAR Model Adaptor tools use the CBR technology to support the following:
  - Model visualization
  - Executed plan assessment
  - Anomaly diagnosis
  - Model repair

Case base Overview

Missions from each realized plan are extracted and included in the Mission Case Base. Case Base Browsing tools allow either a human, or software agents to query, retrieve and analyze case base content.



Case Base Capabilities

- **Case base API** – Supports interaction to obtain case data by software agents or other JAGUAR components.
- **Case base GUI** – For human access to repository.
- **Analysis Tools** – For identifying consistencies and differences across case sets and for generating assessment reports.
- **Multiple Case Repositories –** For storing varying types of cases, e.g., models, executed missions. Queries across case bases are enabled.

The Benefits of the CBR Approach

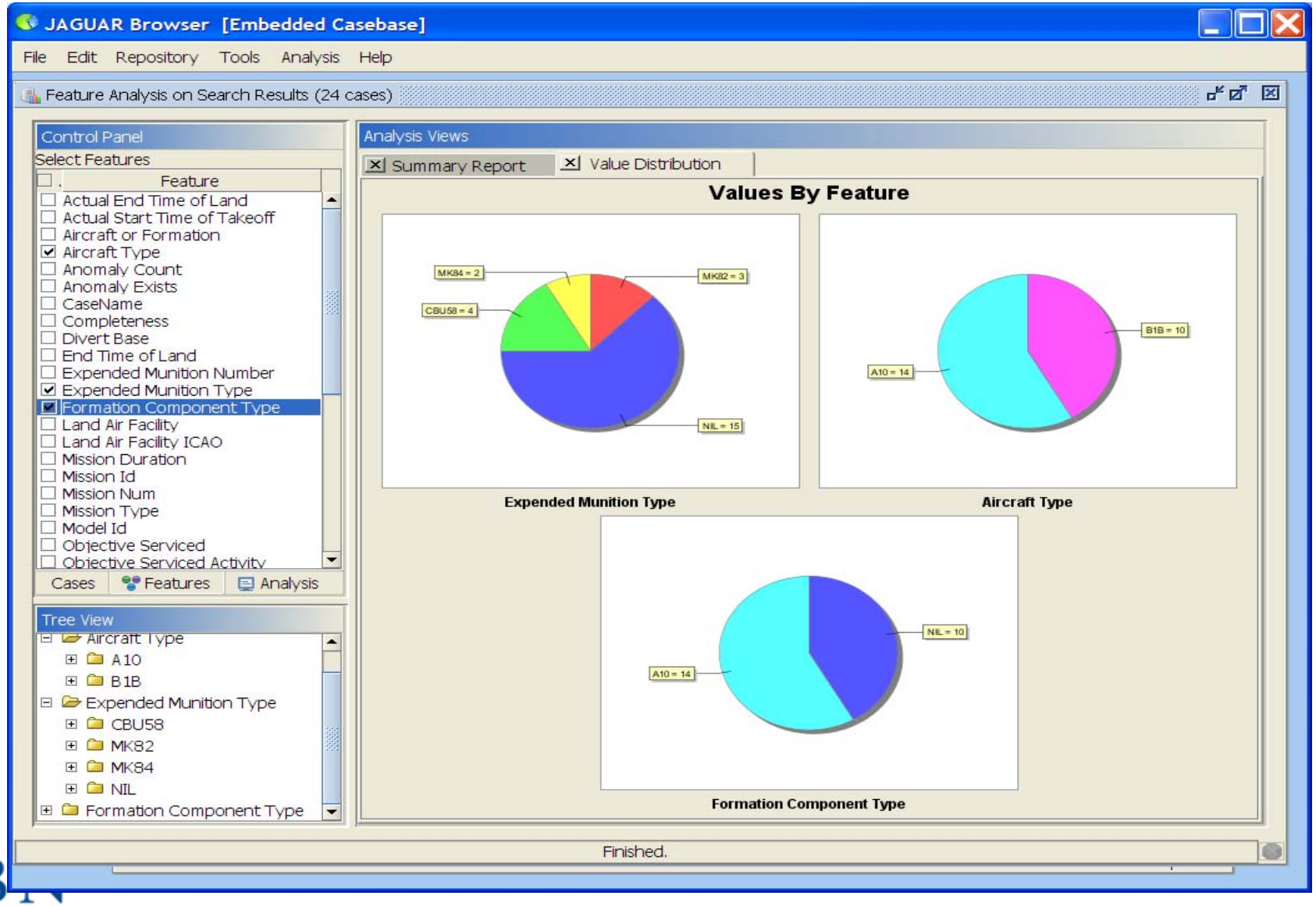
- Cases in a CBR system tend to be described by a feature vector.
  - Case features capture the distinguishing features of an episode, such as a mission plan or execution.
  - Case features support customization of perspectives, e.g., one set of cases can have different features based on the observational perspective.
  - Case features can also be used to provide meta-data about the case, e.g.:
    - performance data
    - subjective evaluation information
    - categorical information
- The feature vector is useful for
  - Search
  - Comparison
  - Report Generation
    - Pie Chart Generation
    - Specialized Feature Table Reports

Deriving Reliable Model Revisions from Executed Plan Data Analysis

- In JAGUAR the plan generator does not directly utilize historical plan data.
- Instead the executed plan data is analyzed with Model Adaptor tools and models may be changed to:
  - Remove or add choice bias
  - Remove errors associated with faulty model parameters, e.g., fuel burn rate, altitude values
  - Add new capability
- Automated tools have been developed to support the operator in the analysis of executed mission data
  - Pie chart tools
  - Feature Report Tools
  - Agent-based Model Repair Suggestion Generation
- Machine learning tools have been investigated to find patterns in the executed plan data.
  - Capturing statistical data
  - Generation of Decision trees that can reveal plan generator choices, especially among available resources within the plan environment.

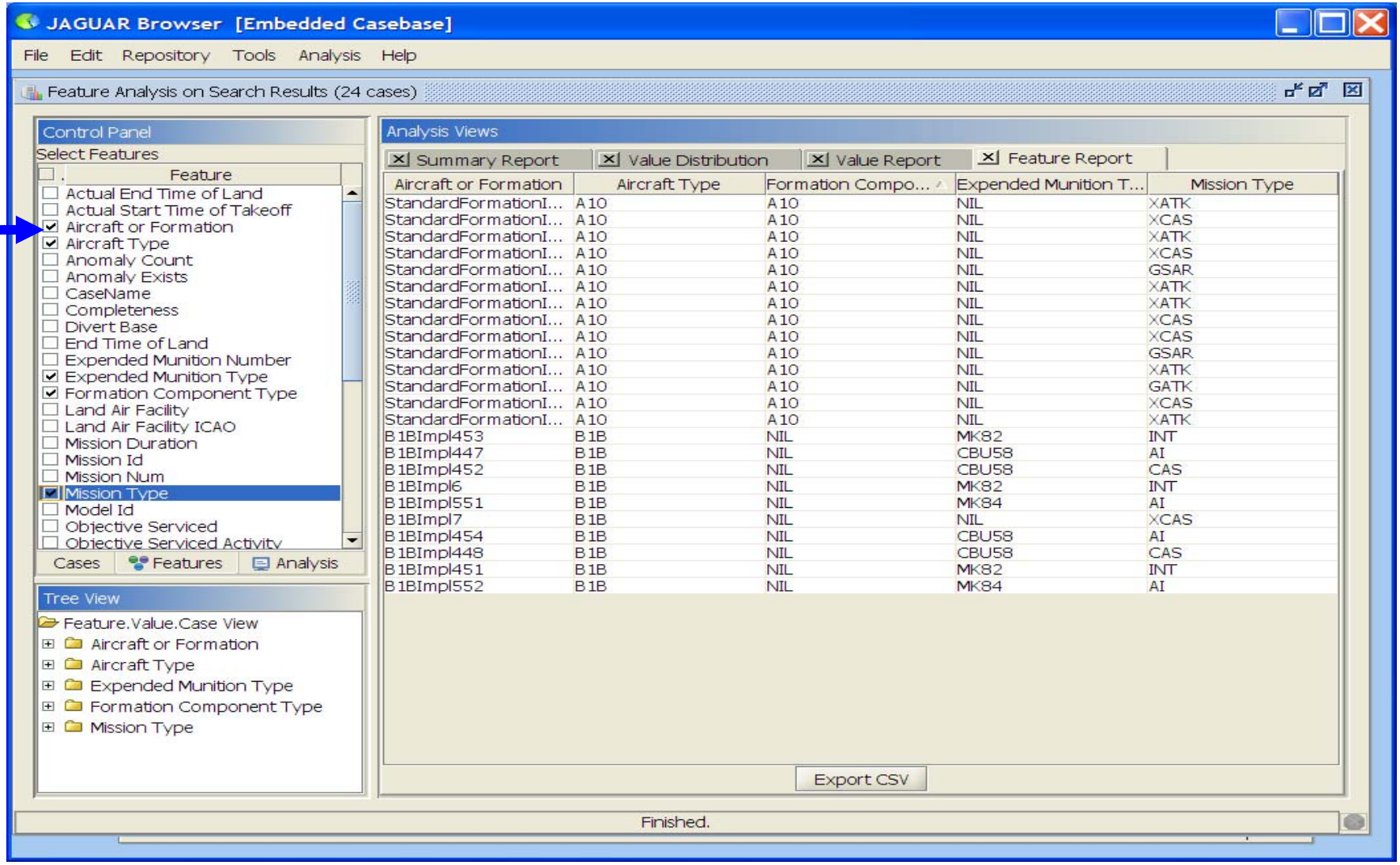
Pie Charts for Analysis

JAGUAR: Joint Air-to-Ground Unified, Adaptive Replanning

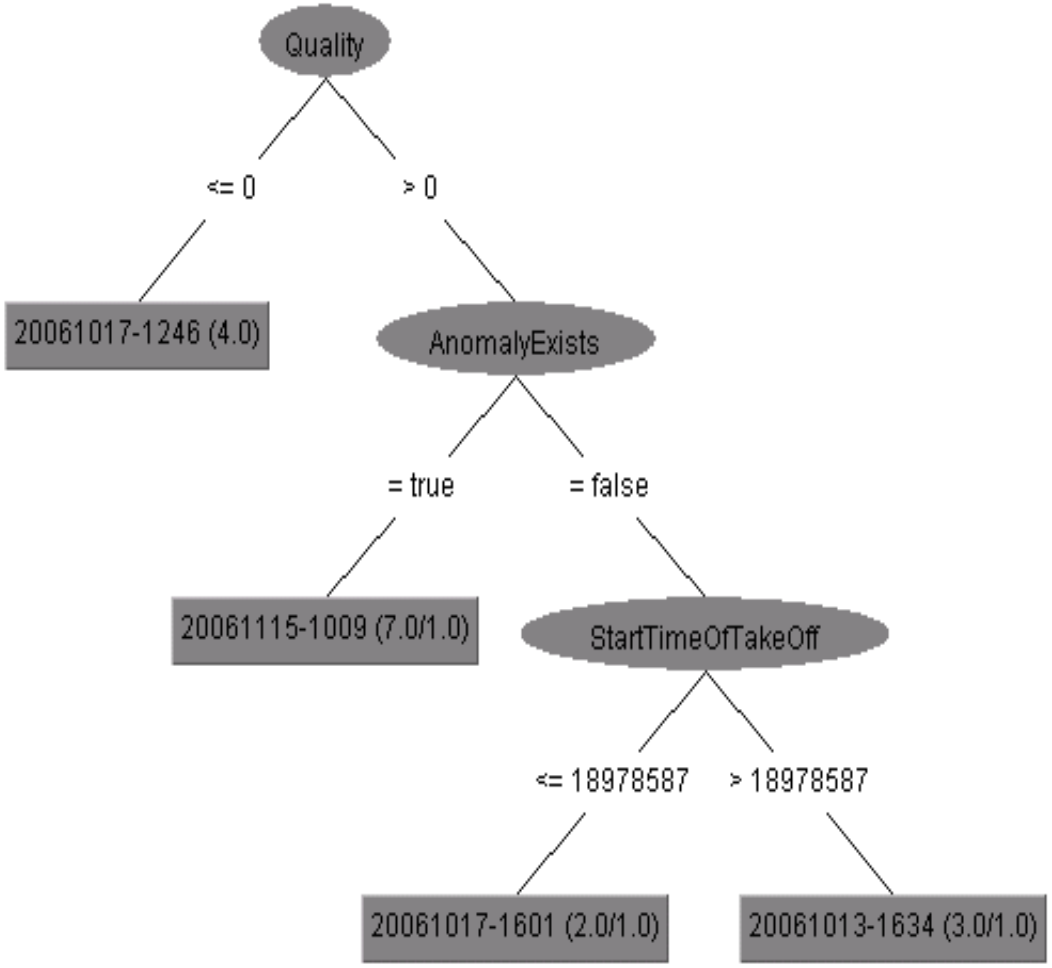


Adding New Features to Explain Data

***JAGUAR**: **J**oint **A**ir-to-**G**round **U**nified, **A**daptive **R**eplanning*



Interpreting Decision Tree Data



- Decision trees like this are useful for identifying anomaly occurrences and can lead to model changes.

- However, this data can easily be misinterpreted. For example, in this tree, "quality" defines the completeness of the executed mission.

- This tree says if the missions did not execute, there were no anomalies. However, there were some trials when the anomalies were not being generated due to software problems. This tree has no information about software problems, so the data can be misleading.

- When the missions did execute, there were anomalies in two of the trials. However, the type of anomaly is not explicitly stated.

Experimentation

- Several experiments were conducted to evaluate the usefulness of JAGUAR tools by operational users.
- The experiments involved operators with a range of expertise. All operators were provided with checklists for guidance in utilizing a particular tool. Some operators participated in multiple experiments.

Experiment #	Experiment Duration	Number of Executed Missions	Subject Participation*
1	1 day	379	A, B, c
2	1 day	440	A, B, d
3	3 days	804	A, B, d, e, f, g, h
4	1 day	186	A, B. d
5	3 days	863	A, B, d, f, i

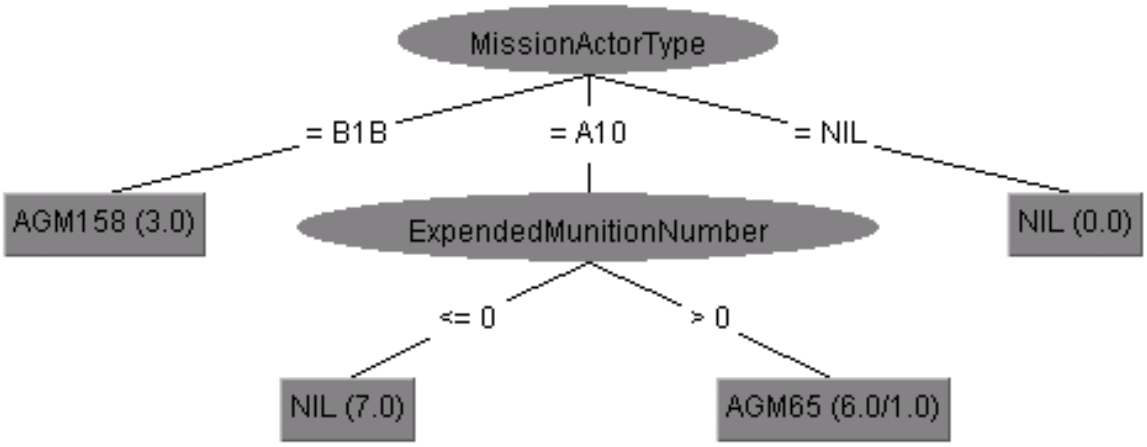
* Each letter represents an actual subject. When A occurs across experiments that indicates that subject A participated in the experiment. Subjects A and B participated in all of the experiments and hence became more skilled in the usage of the JAGUAR tools.

Findings: Creating a New Mission Case Base

- With limited guidance, one of the trained subjects was able to create a new mission case base.
  - Used two days of test data for the new case base.
  - Successfully aggregated appropriate historical information into one case base for easier analysis.
  - Used analysis tools to identify items of interest within the data set.
    - Distribution of mission types
    - Types of aircraft used
    - Number of anomalies

Findings: Decision Trees

- In general, our experience using machine learning tools to generate decision trees about executed plan data was mixed.
- Bad – Knowing what settings to use for creating the decision trees was difficult. Knowing what Nil references.
- Good - In some cases, we were able to quickly determine how weapons were being expended in the selected missions and then refine models to affect how the plan generator would make future resource selections.
  - For example, in this decision tree, the algorithm C4.5 found that the A10 used an AGM65 in 6 out of 7 instances.



Findings: Feature-based Analysis Tools

- Pie charts and feature reports are good ways for determining how aircraft and munitions are used.
  - It is more difficult to determine when an aircraft that should have been used was not used.
- Using existing features to create custom reports was useful.
  - While some custom reports can be saved, the process needs to be easier to implement.
- Features only describe a subset of the executed mission data.
  - When a user wants a report to describe data in the executed mission that has not been described with a feature, one or more new features will need to be added.
  - The addition of new features requires the support of a skilled MA operator.

Summary

- The Model Adaptor tools can effectively assist domain experts, and potentially the operational users:
  - to determine the validity of models
  - to understand how the model set affects overall plan generation
  - to provide credible indications of requirements for model repair
- The use of the case base helps human military operators analyze historical executed mission cases, evaluate and recommend updates to the associated models.
- Operators with intelligence backgrounds were more comfortable using the machine learning tools and data.
- The executed mission case data reports are useful for identifying both successful and failed trends and for aiding operators in plan assessment and reporting.

Conclusion

- Model Adaptor tools:
  - suggest when models should be repaired
  - provide model repair justification
  - provide tools for model creating and repair
  - provide tools for model publication
  - support plan assessment
- The human operator has the final responsibility for implementing and publishing model repairs.