Micro Analysis & Design

# Review of Human Performance Models for the Prediction of Human Error

Workshop on Human Error Modeling

NASA Ames Research Center

Ron Laughery

Ken Leiden

Micro Analysis & Design

October 18, 2001

Micro Analysis & Design

# Outline of the presentation

- Introduction
- Approach
- Taxonomies of human error
- Using HPM to predict human error
- Review of HPM architectures for error prediction
- Architecture Recommendations

Micro Analysis & Design

# Introduction

- MA&D received two HEM contracts
  - IMPRINT/ACT-R hybrid approach
  - Investigate integration issues for the various HPM architectures
- NASA requested that we change the focus of the integration study
  - Instead, review the state-of-the-art in HPM as it applies to HEM
  - Open-ended, only specific request was that we discuss error chains at some level

Micro Analysis & Design

# Overview

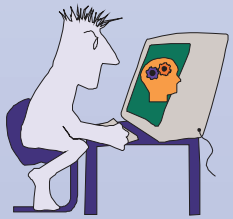
- Introduction
- **Approach**
- Taxonomies of human error
- Using HPM to predict human error
- Review of HPM architectures for error prediction
- Architecture Recommendations

Micro Analysis & Design

# Approach

- What do we mean by human error?
  - 1st step - literature search on error taxonomies
    - Focused on aviation literature

    - Provided insight into causal factors of human error
    - Bound the error categories that we would evaluate
- What models are there?
  - Literature search for *any* computational approach used for human performance modeling or error prediction.
  - Only a few efforts aimed specifically at the prediction of human error.
    - Confirmed our beliefs that HEM is in its infancy



Micro Analysis & Design

# Approach (cont)

- Computational approaches explored were limited to existing HPM architectures
  - no stand-alone theories or algorithms
- For each architecture we gathered information to allow us to prepare a report including:
  - Brief overview about high level features
    - Pragmatic aspects (e.g., what platforms support it)
  - Underlying theory of operation
  - Demonstrated or purported error prediction capabilities

Micro Analysis & Design

# Overview

- Introduction
- Approach
- **Taxonomies of human error**
- Using HPM to predict human error
- Review of HPM architectures for error prediction
- Architecture Recommendations

Micro Analysis & Design

# Taxonomies of Human Error

- Understand underlying error mechanisms that should be accounted for by HEM
  - e.g., certain types of errors occur in high workload situations, therefore HEMs used to study those errors should address operator workloads
- Statistics of error types within each taxonomy provide an estimate of the frequency and criticality of occurrence
  - Allows us to focus on modeling the error types that are more likely to occur and have consequences.

# Taxonomies (cont)

- No single taxonomy is generally accepted for addressing all causal factors
- Focus on understanding the cognitive process involved in the production of human error, rather than the observable characteristic of the error
  - Situation Awareness (Endsley) – 14 error categories
  - Internal Human Malfunction (based on Rasmussen and extended by O'Hare) – 6 error categories
  - Unsafe Acts (Reason) – 8 error categories
  - Information Processing (Wickens) – 5 error categories

Micro Analysis & Design

# Error Taxonomy: Situation Awareness

| Error Type | Error Description | Freq (%) |
|---|---|---|
| **Level 1: Failure to correctly perceive information** | | |
| Data not available | Data is not available due to failure of the system design to present it or failure in the communication process. | 13 |
| Data hard to detect | Examples are poor runway markings or inadequate lighting, noise in the cockpit, or obstructions blocking view. | 11.1 |
| Failure to monitor or observe data | Data is available, but is not scanned due to simple omission, attentional narrowing, distractions due to multi-tasking, or high workload. | 35.1 |
| Misperception of data | Data is misperceived due to influence of prior expectations or misunderstood due to task distraction. | 8.7 |
| Memory loss | Forgetting information is due to disruptions in normal routine or high workload. | 8.4 |
| **Level 2: Failure to correctly integrate or comprehend information** | | |
| Poor mental model | Poor mental model does not enable the combining of information needed to meet goals, often associated with automated systems. | 6.9 |
| Use of incorrect mental model | Interpretation of cues through an expected, but wrong, mental model of a system's behavior leads to the incorrect assessment. | 6.5 |
| Over-reliance on default values | Routine expectations of the system is assumed even though conflicting information is available, but not accessed. | 4.6 |
| Other | Information is not properly integrated or comprehended due to working memory lapses or other cognitive reasons. | 2.3 |

Micro Analysis & Design

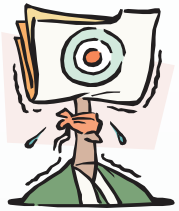
# Error Taxonomy: Situation Awareness (cont)

| Error Type | Error Description | Freq (%) |
|---|---|---|
| **Level 3: Failure to project future actions or state of the system** | | |
| Poor mental model | Information of current state is correctly understood, but projection of that state into the future fails because of poor understanding of how to do so. | 0.4 |
| Over-projection of current trends | The current state is projected into the future correctly. However, it is projected further into the future than for which the data is realistically valid. | 1.1 |
| Other | Projection of current state into the future fails because it is a demanding task that in a multi-tasking environment is not always performed. | 1.9 |
| **General** | | |
| Failure to maintain multiple goals | Failure to maintain multiple goals in memory degrades SA across all three levels. | 0 |
| Executing habitual schema | Performing task automatically can result in important system cues being overlooked. | 0 |

Jones and Endsley utilized voluntary reports from NASA's Aviation Safety Reporting System for 111 incidents involving pilots and 32 incidents involving controllers.

Micro Analysis & Design

# Error Taxonomy: Internal Human Malfunction



| Error Type | Error Description | Freq. (%) |
|---|---|---|
| Information error | The pilot did not detect cues arising from the change in system states. | 6.1 |
| Diagnostic error | The pilot did not accurately diagnose the state of the system based on the information available. | 21.7 |
| Goal Setting error | The pilot did not choose a goal that was reasonable given the circumstances. | 11.5 |
| Strategy Selection error | The pilot did not choose a strategy that would achieve the intended goal. | 12.9 |
| Procedure error | The pilot did not execute procedures consistent with the strategy selected. | 39.5 |
| Action error | The pilot did not execute procedures as intended. | 8.2 |

- Based on flight related mishaps for the US Navy and Marine Corps between 1977 and 1992. (Wiegmann and Shappell)
- 91.3% (264) of the mishaps attributed to pilot causal factors fit this model

Micro Analysis & Design

# Error Taxonomy: Unsafe Acts

| Error Type | | Freq (%) |
|---|---|---|
| Skill-based error | Attentional slip | 14.3 |
| | Memory lapse | 11.2 |
| Rule-based mistake | Misapplication of good rules | 57.1 |
| | Application of bad rules | |
| Knowledge-based mistake | Solving a unique problem that involves conscious reasoning | |
| Violations (deviations from normal operating procedures) | Skill-based | 17.4 |
| | Rule-based | |
| | Knowledge-based | |

- Based on flight related mishaps for the US Navy and Marine Corps between 1977 and 1992. (Wiegmann and Shappell)
- 91.3% (264) of the mishaps attributed to pilot causal factors fit this model

Micro Analysis & Design

# Error Taxonomy: Information Processing

| Error Type | Error Description | Freq. (%) |
|---|---|---|
| Sensory error | Error converting physical phenomena (e.g., light, sound) into neural manifestations. The information lasts briefly and does not require attention resources. | 2.8 |
| Pattern recognition error | Error mapping the physical codes of the sensory stores into meaningful elements (markings into letters, letters into words) | 14.9 |
| Decision/response selection error | Information used incorrectly -- can be stored in working memory to be used in the near future; combined with other information; or initiate a decision process to immediately select a response | 29.5 |
| Response execution error | Occurs when the high level response is decomposed incorrectly into the required auditory, motor and cognitive steps. | 45.5 |
| Attention resources error | Can be viewed as a limiting factor for the previous three stages. | 7.3 |

- Based on flight related mishaps for the US Navy and Marine Corps between 1977 and 1992. (Wiegmann and Shappell)
- 86.9% (251) of the mishaps attributed to pilot causal factors fit this model

Micro Analysis & Design

# Overview

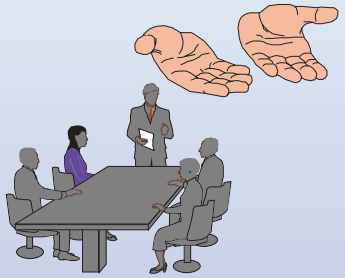
- Introduction
- Approach
- Taxonomies of human error
- **Using HPM to predict human error**
- Review of HPM architectures for error prediction
- Architecture Recommendations

Micro Analysis & Design

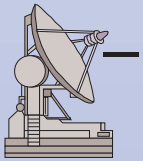
# Modeling Human Error: What Matters?

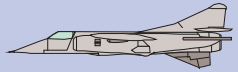
- Taxonomies provided a sense of what aspects of human behavior are linked to human error

- Also, what aspects of hardware, software, and scenario design contribute to human error



- These drive what *may* need to be represented in a model
  - What aspects of the human/machine system to represent in the model and what to leave out?
  - *Focus, focus, focus…*

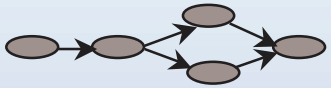




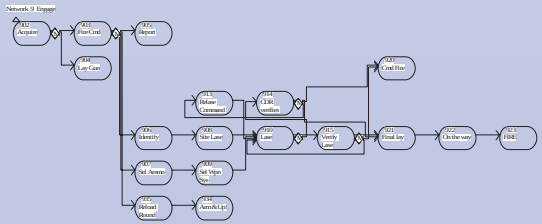
Micro Analysis & Design

# Error Chains

- Incidents/accidents attributed to human error almost always result from a chain of errors rather than a single error
  - What level in the error chain should the impact be considered unsafe or otherwise significant?
    - Option 1 – Human reliability assessment



      - Error rates for tasks determine probability of getting to a level in the error chain that is statistically significant
    - Option 2 – Focus on likely chains of errors leading up to the point where:
      - the next level in the chain is unsafe
      - or the next level requires automation and/or equipment to break the chain
    - Option 2 is needed for human error prediction



Micro Analysis & Design

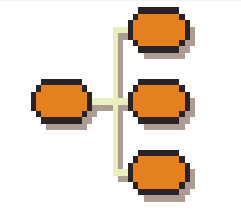
# Overview

- Introduction
- Approach
- Taxonomies of human error
- Using HPM to predict human error
- **Review of HPM architectures for error prediction**
- Architecture Recommendations

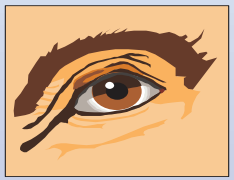
Micro Analysis & Design

# HPM Architectures

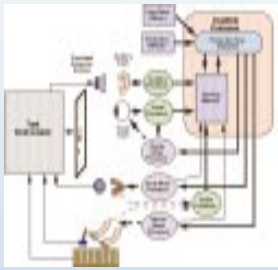
- Task Network Models
  - Micro Saint
  - IMPRINT
  - WinCrew
  - IPME



- Vision Models
  - ORACLE
  - OptiMetrics
  - Georgia Tech



- Cognitive
  - GOMS-based
    - GLEAN
    - EPIC
    - APEX

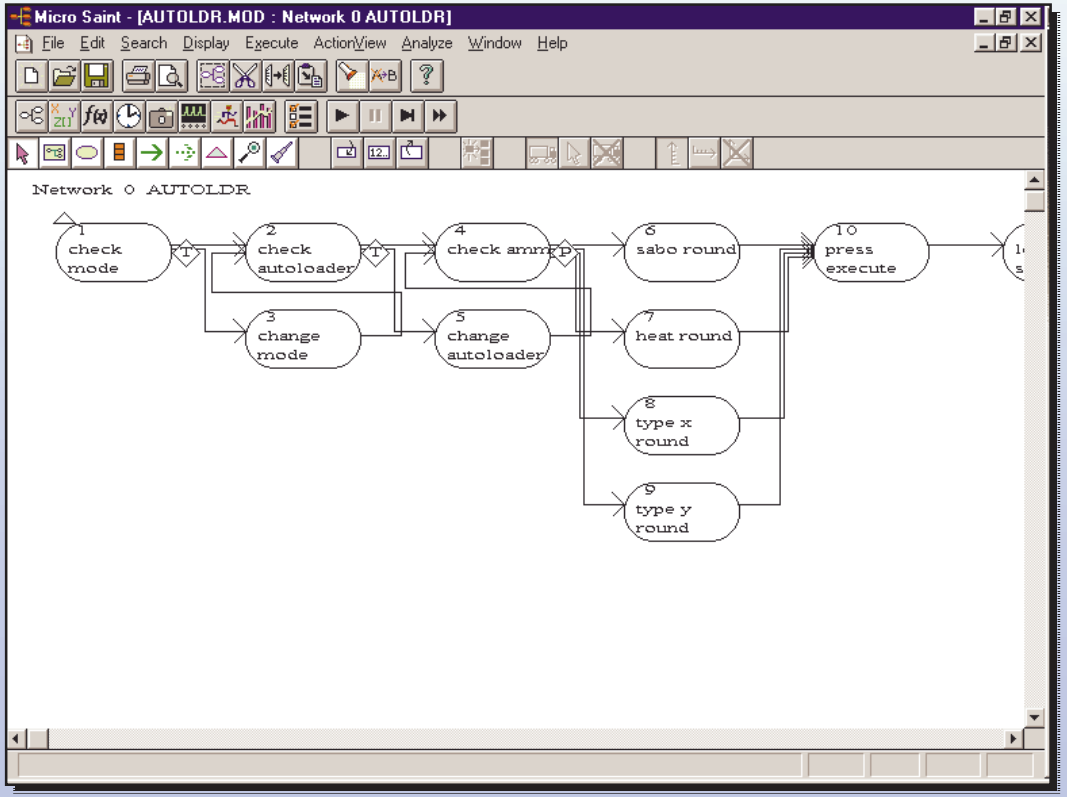


  - ACT-R
  - MIDAS (Air & Core)
  - D-OMAR
  - SAMPLE
  - COGNET
  - Soar

Micro Analysis & Design

# Task Network Modeling

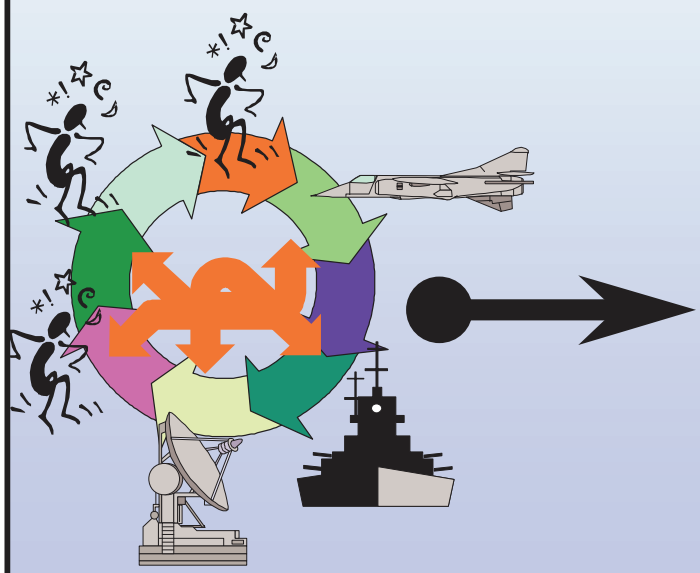
- Largely involves the extension of a task analysis into a network defining sequencing

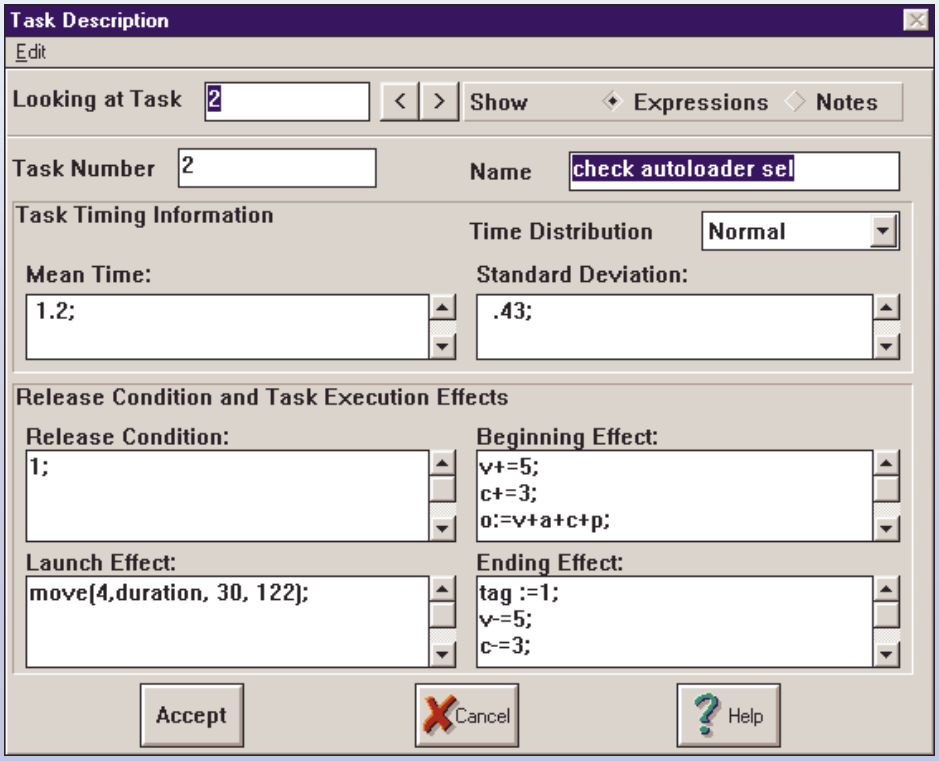


Micro Analysis & Design

# Going from a task network to a running computer model

- Add timing information and task/system interdependencies

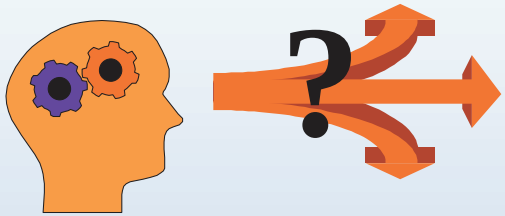




Micro Analysis & Design

# Add human decision making strategies

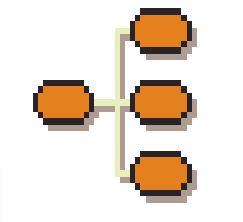
- Any defined branch point represents a need for a decision



- Logic and rule sets of any level of complexity can be included

Micro Analysis & Design

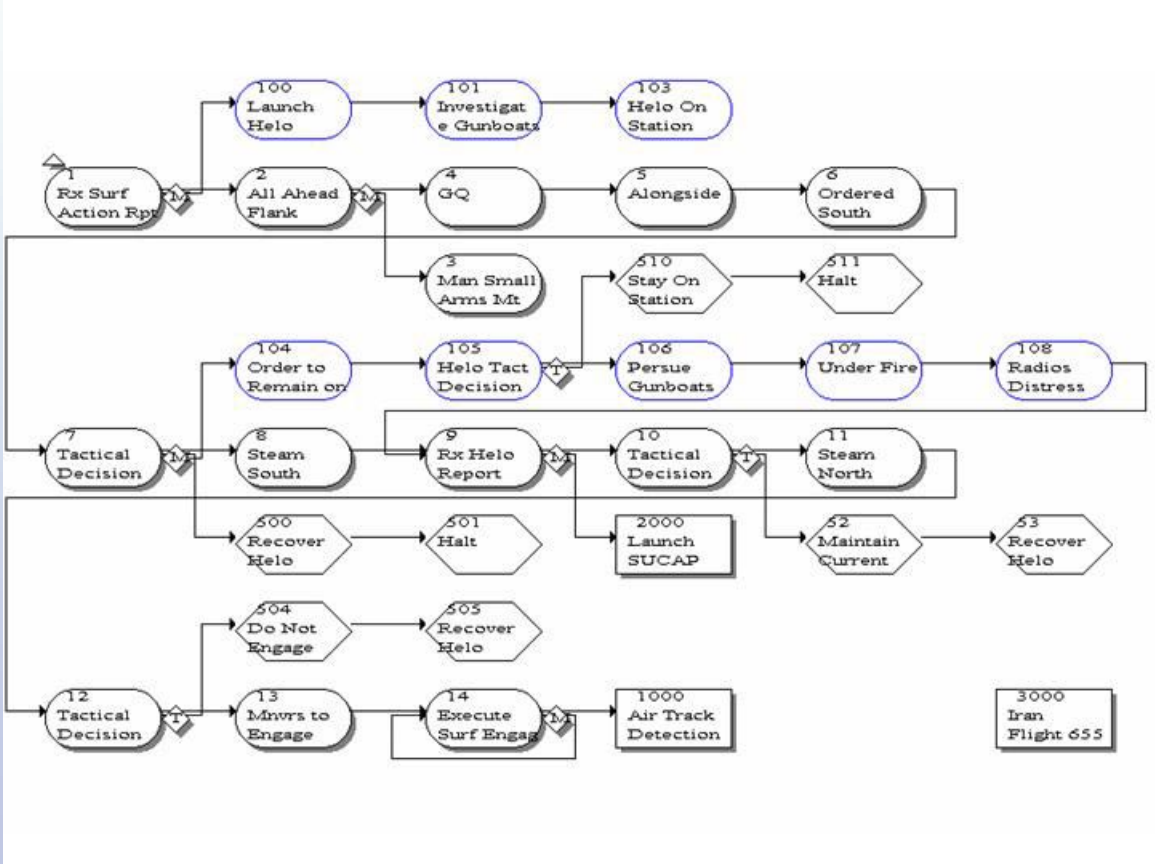
# Task Network Models



- Use human/system task sequence as organizing structure
  - Micro models of human performance determine task time (e.g., how long to read a message depends on # of words and reading rate)
- How finely we decompose the tasks and the amount of the system which is simulated depends on the particular problem
- Multi-tasking is easily modeled
- Embedded models of cognitive, perceptual, and motor behavior can be embedded

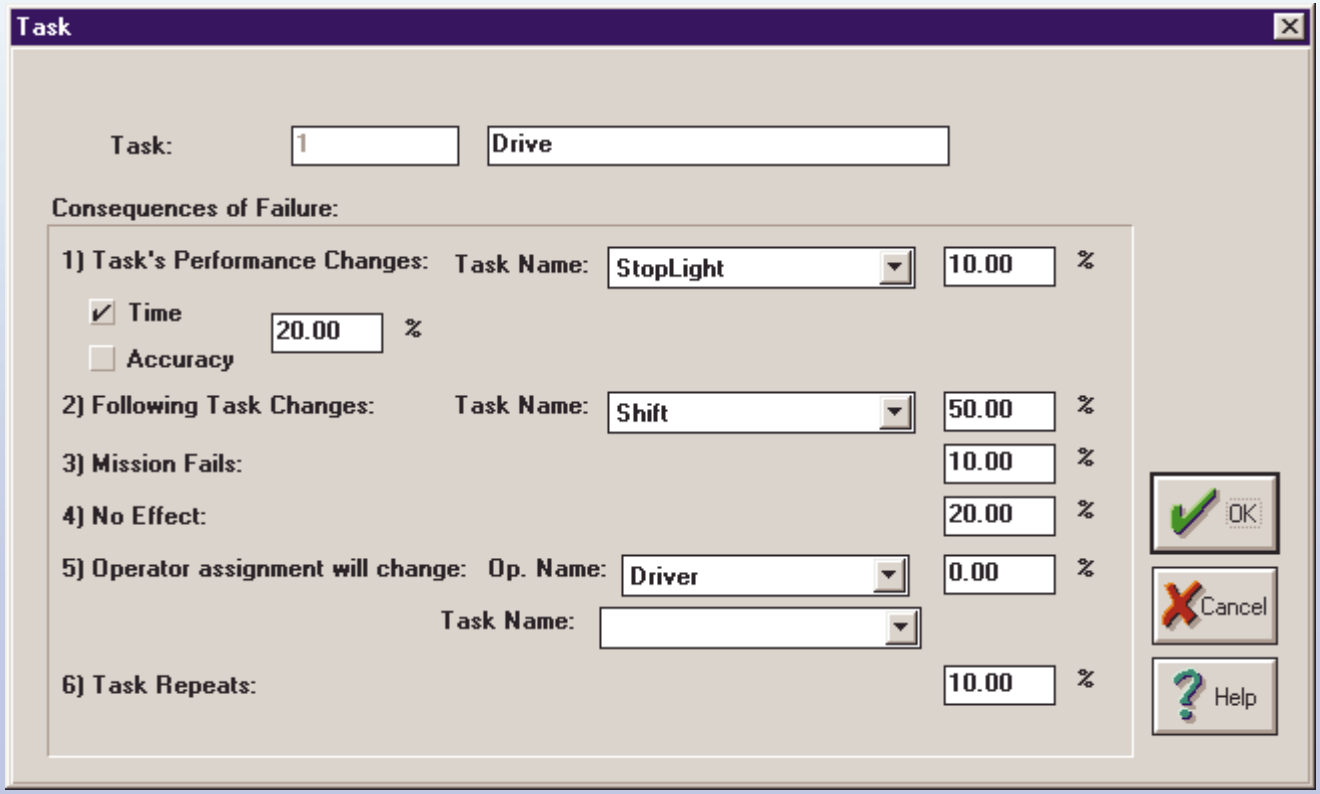
Micro Analysis & Design

# Example: Task Network Model for USS Vincennes Accident Recreation



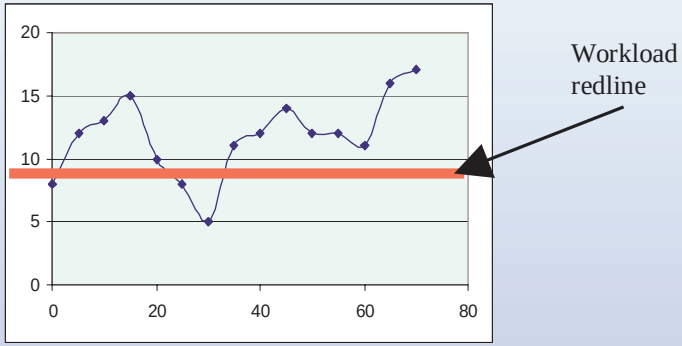
# Use of Task Network Modeling in Error Prediction

- Previous HEM studies haven't predicted error, but shown impact of error and error recovery strategies on system performance



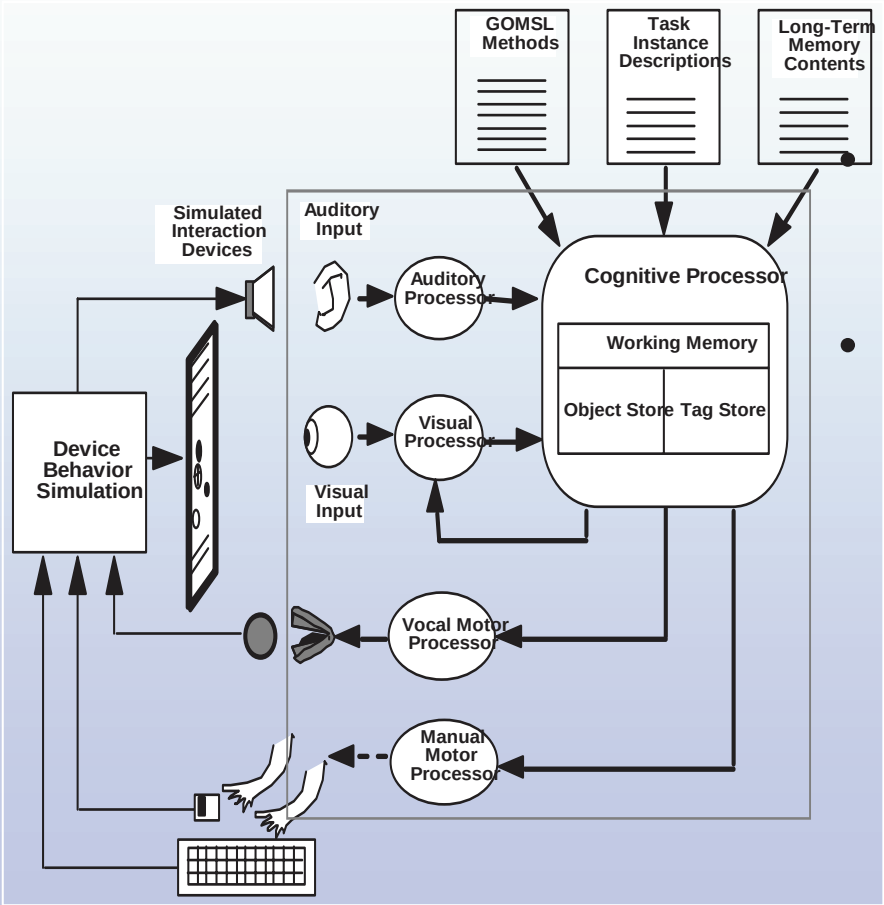
# Use of Task Network Modeling in Error Prediction

- Good potential for error prediction in the context of multi-tasking and high operator workload.
  - Embedded workload models predict:
    - Error-prone operator conditions as well as the specific human behaviors that result
      - e.g., task shedding due to high workload




- IMPRINT/ACT-R hybrid approach for current runway incursion study shows potential for embedding cognitive models

# GLEAN Overview




- *GOMS Language Evaluation and Analysis*
- Developed at U of Michigan with DARPA and Navy support
- Simulates user interaction with a simulated device - assists with UID
- Automates GOMS - Goals, Operators, Methods, and Selection Rules
  - Goals - what the user is trying to accomplish, often in a hierarchy
  - Operators - set of atomic-level operations to compose solution to a goal
  - Methods - sequences of ‘operators’ to accomplish a single goal.
  - Selection Rules are used to decide which method to use to achieve a goal when several are applicable

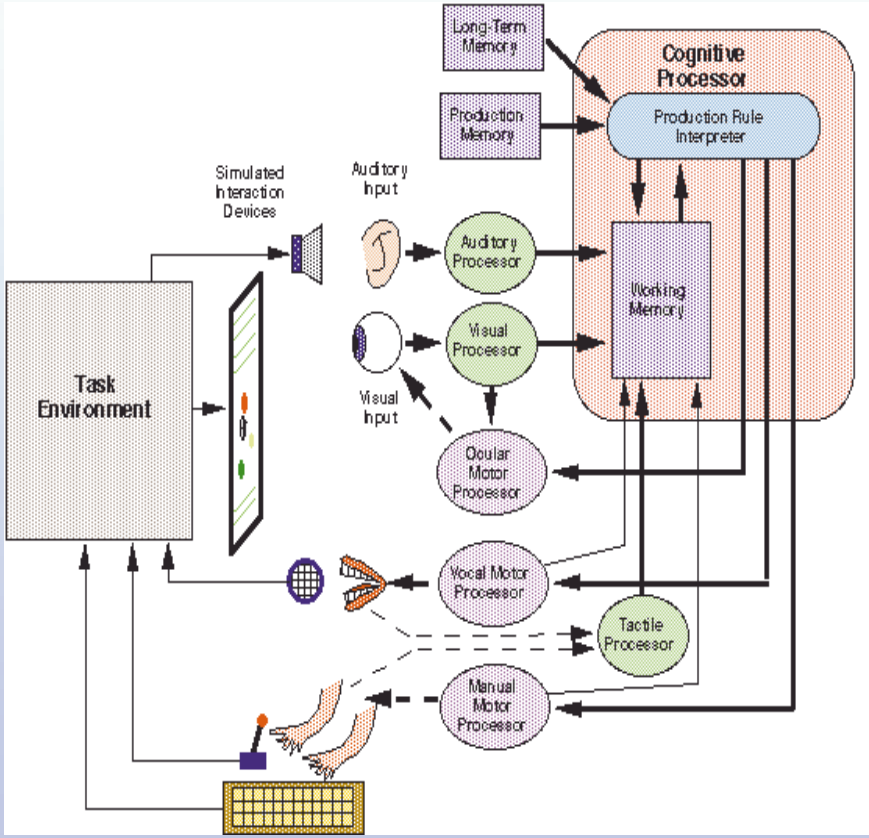
Micro Analysis & Design

# Use of GLEAN in Error Prediction

- Three studies
  - Investigated effects that task structure had on error production due to limited working memory
    - e.g., small differences in the way spreadsheet interface required the user to enter cell formulas had a dramatic impact on the number of errors that were produced
    - One study also considered the decay of goal items in working memory
- Scott Wood's Ph.D. dissertation discussed extensions to GOMS-like architectures to support modeling of human error

Micro Analysis & Design

# EPIC Overview




- *Executive Process-Interactive Control Language*
- Developed at U of Michigan
- Goal to determine the extent that humans can multi-task
  - Separate auditory, visual, and tactile processors for perceptual inputs
  - Separate vocal, ocular, and manual motor processors for producing actions
  - All controlled by the cognitive processor
- Production rule system

Micro Analysis & Design

# Use of EPIC in Error Prediction

- No published error modeling research, but some work has been conducted
  - Minor changes to working memory component, allowing goal and task elements to decay and be confused with similar items allows modeling
    - Errors of omission and commission, reversals, transpositions, and capture errors
    - Other error types, such as those that are caused by missing data and data corruptions, can already be seen in EPIC models of high performance tasks
    - Errors while managing multiple task strategies are well represented

Micro Analysis & Design

# APEX Overview

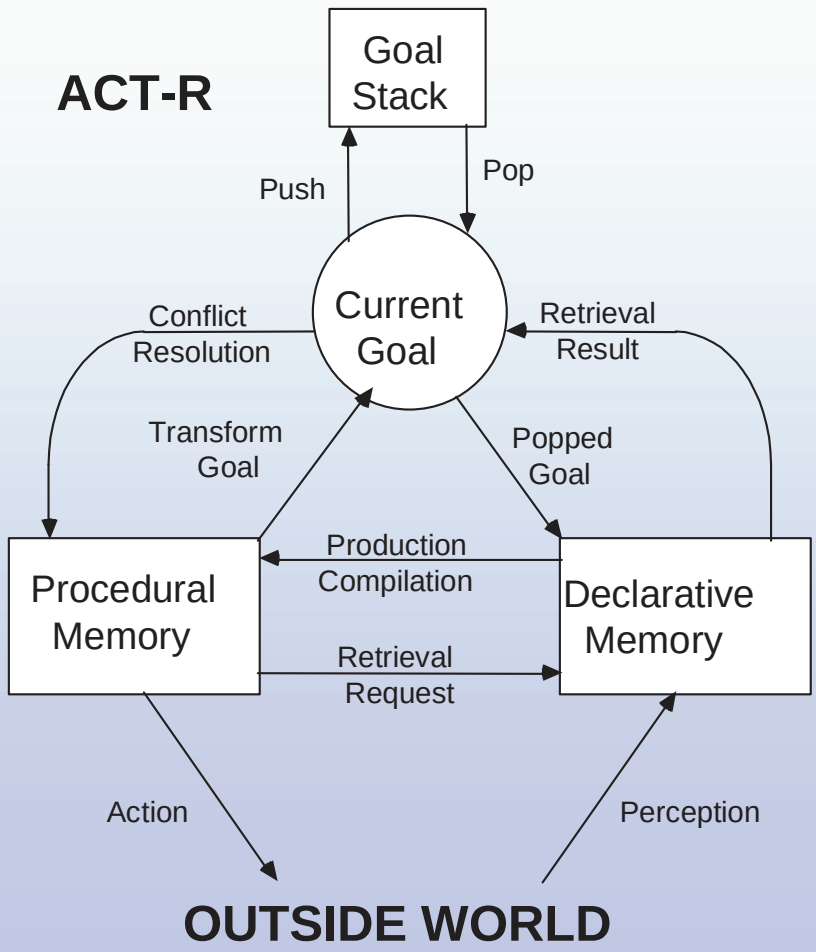
- Created by Dr. Michael Freed
  - Development continues at NASA ARC and elsewhere
- GOMS approach building on cognition, perception, and motor movement – referred to as CPM-GOMS
  - Advantage of CPM-GOMS is that models of tasks from one domain may carry over to other domains
    - Model reuse through low-level behavior template libraries.
- APEX automates the difficult CPM-GOMS task of coordinating the cognitive, motor, and perceptual resources.
  - User can perform these functions without specific expertise in cognition or CPM-GOM

Micro Analysis & Design

# Use of APEX Error Prediction

- Can be used to predict *habit capture* errors in ATC domain
  - Errors occur not because of a failure to successfully retrieve information, but rather a failure to make a retrieval attempt.
  - Occur when the human fails to allocate the minimal cognitive resources to retrieve task-relevant information from memory.
    - Without this information, decision making mechanisms rely on default assumptions, resulting in error if theses assumptions are not suitable to the situation.

# ACT-R Overview




- *Atomic Components of Thought - Rational*
- Developed at CMU
- Represent the steps of cognition by using production system theory
- Time step from 50 ms - few hundred ms
- Architecture and models experimentally grounded
- Goal stack contains the ranking of priorities for guiding behavior
- Procedural memory contains production rules in terms of a condition-action pairs
- Declarative memory contains knowledge structure elements

Micro Analysis & Design

# Use of ACT-R in Error Prediction

- Memory retrieval study demonstrated built-in mechanism for modeling two types of errors
  - Memory retrieval failure (memory loss)
  - Retrieval is successful, but the wrong information is retrieved (slip)
- Situation awareness error study
  - Soldiers with helmet-mounted displays for navigation
    - Asked questions about their location, location of friendly units, location of enemy
    - One ACT-R model per soldier per question.
    - Over 2 million lines of ACT-R code

Micro Analysis & Design

# MIDAS

- *Man-machine Integration Design and Analysis System*
- Developed for Army-NASA Aircrew/Aircraft Integration program to explore the computational representations of human-machine performance to aid crew system designers.
- 2 distinct versions of MIDAS
  - *Air MIDAS* follows the original LISP implementation.
  - *Core MIDAS* refers to the C++ version of MIDAS .
- Both versions supported by essentially same underlying theory

Micro Analysis & Design

# MIDAS

## Inputs

**CREW STATION DESIGN**
- Cockpit Geometry, Display/Control Layout
- Cockpit Equipment Functionality

**MISSION/ TASKS**
- Flight Profiles, Waypoints, Other Scenario Objects
- Planned Operator Mission Activities

**OPERATOR CHARACTERISTICS**
- Cognitive Characteristics
- Physical/Motor Characteristics

## User Interface

- Interactive Graphical Interface
- Cockpit Design Editor
- Equipment Editor
- Route Editor
- Activity Editor

## Interactive Cockpit Design Tools

- Jack Anthropometric Model
  - Vision Models

## Simulation System Models & Tools

**Symbolic Operator Models**
- Vision
- Perception, Attention
- World Representation
- Decision
- Scheduling
- Task Loading
- Mission Activities
- Motor

**World Model**
- Vehicle
- Cockpit Equipment
- Flight Dynamics
- Guidance
- Environment, Other Objects
- Terrain

## User Interface

- Interactive Graphical Interface
- Display Animation
- Runtime Data Graphical Displays
- Summary Data Graphical Displays
- 3D Graphical Display

## Outputs

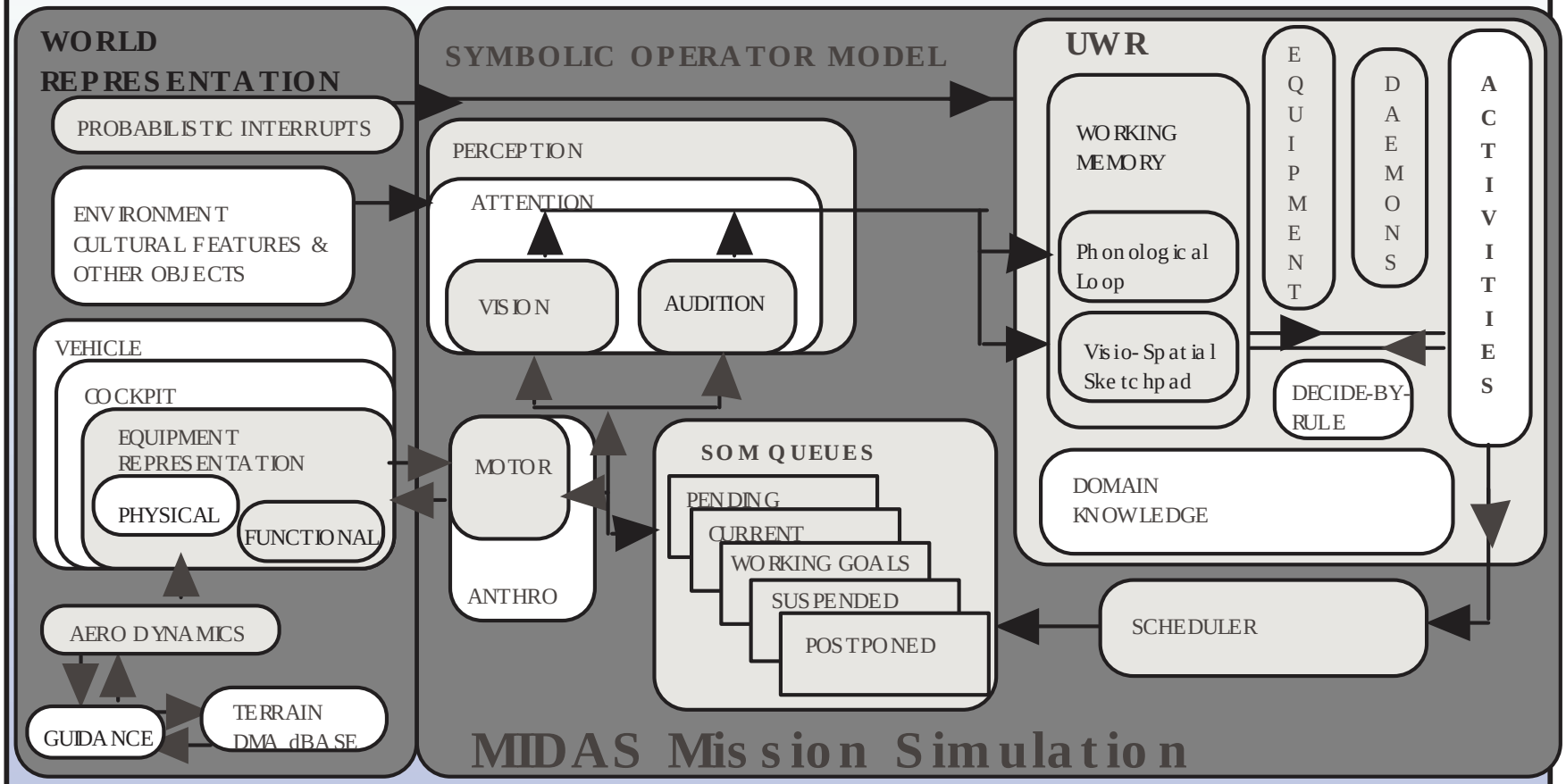
**HUMAN FACTORS ANALYSIS RESULTS**
- Reach, Fit & other MIL-STD Analyses
- Visibility & Legibility Analysis

**MISSION / OPERATOR PERFORMANCE MEASURES**
- Activity Traces
- Task Load Timelines
- Resource Conflicts
- Information Flow

**VISUALIZATION OF SIMULATED MISSIONS**
- Operator Activities
- Equipment States

Micro Analysis & Design

# Air MIDAS Architecture



Micro Analysis & Design

# Use of Air MIDAS in Error Prediction

- No research specifically aimed at error modeling has yet been completed, work underway on:
  - NASA runway incursion study
  - Large multi-agent simulation for NASA safety analysis using Georgia Tech's Re-configurable Flight Simulator
- Good potential for predicting:
  - Goal setting errors in a multi-tasking environment
  - Errors due to mismatch between task demands and resources
    - Tasks performed based on available perceptual and cognitive resources

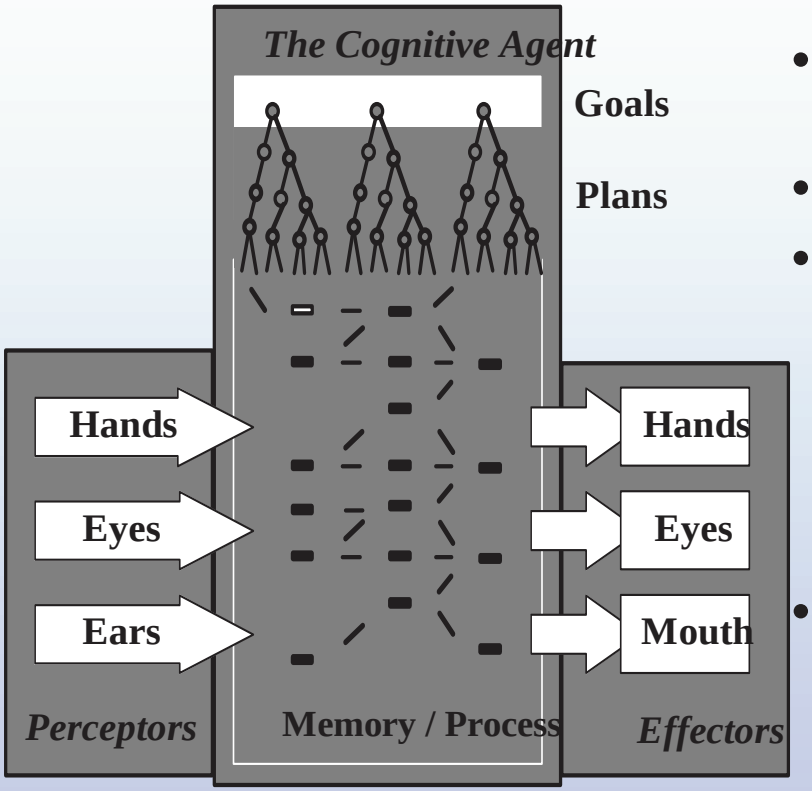
Micro Analysis & Design

# Use of Core MIDAS in Error Prediction

- No prior error prediction research
- Working memory model in 2.0 beta version had not yet been completed.
  - Ability to model memory loss and higher level errors caused by memory loss not possible at this time
- Most promising aspect of error prediction are associated with the new SA model
  - Gauges *data available from the system* versus *data known by the operator*
    - Provides a graphical output display depicting the ratio during run-time.

Micro Analysis & Design

# D-OMAR Overview




- *Distributed Operator Model Architecture*
- Developed by BBN Technologies
- Designed to facilitate the modeling of the human multi-tasking behaviors
  - Agents interacting with complex equipment in pursuing collaborative enterprises
- Not limited to any particular theoretical approach in the design of HPM.
  - Flexible architecture enables users to develop models according to their own psychology, philosophy, or inclinations

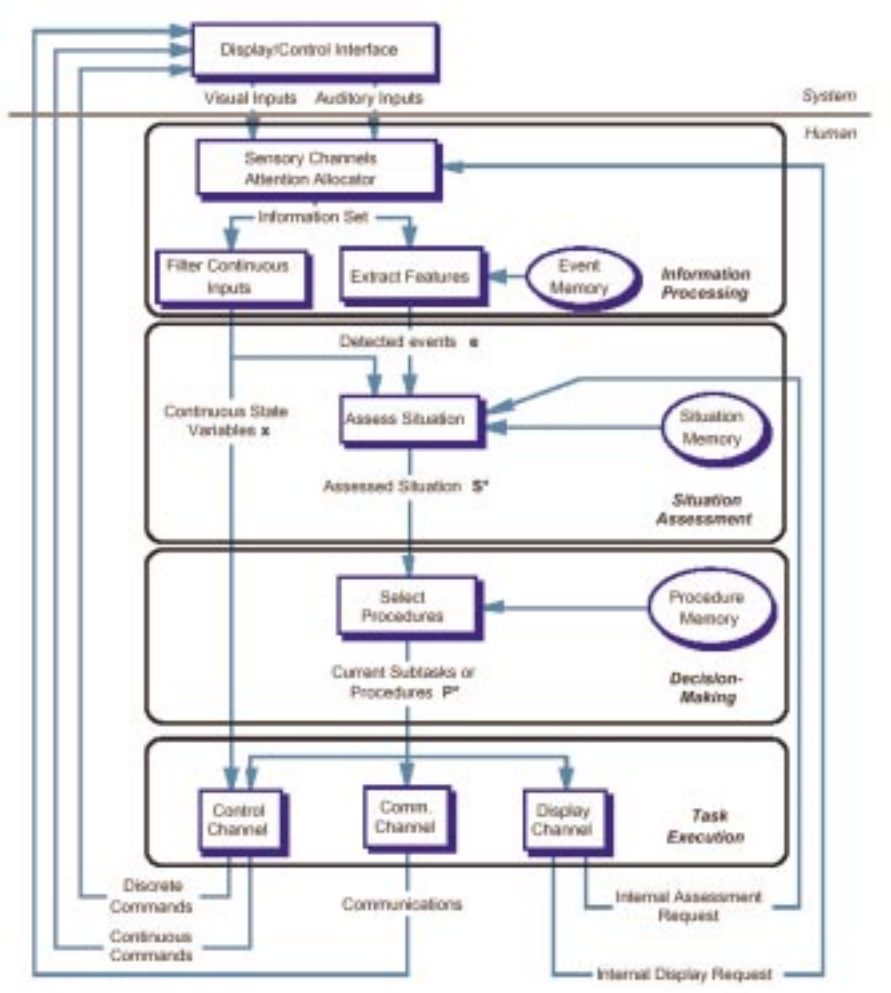
Micro Analysis & Design

# Use of D-OMAR in Error Prediction

- No prior error prediction research (except for current NASA runway incursion study)
- Potential for error prediction in these areas:
  - Policy-based approach for modeling contention between competing tasks provides a framework for exploring errors from
    - task distraction
    - task disruption
    - multi-tasking
    - attentional narrowing

Micro Analysis & Design

# SAMPLE Overview




- *Situation Awareness Model for For Pilot-In-The-Loop Evaluation*
- Developed by Charles River Analytics
- Agent-based architecture
- Emphasizes SA-centered decision-making in multi-task environment
  - Defines decision maker's view of the environment and characterizes the information needs that drive decision-making
  - *Situation assessor* employs Belief networks to represent probabilistic reasoning in the presence of uncertainty

Micro Analysis & Design

# Use of SAMPLE in Error Prediction

- No prior error prediction research
- SA-centered decision approach provide a solid foundation for investigating errors from SA Error Taxonomy
- May be able to predict skill or rule-based errors by stochastically changing thresholds in the production system

# Soar Overview

- Originally developed at Carnegie Mellon
- Began as an architecture for creating intelligent systems
- “Unified theory of cognition” that attempts to unify all intelligent behavior within a relatively small set of functional principles
  - Long term memory storage is represented by production rules
  - Working memory represents current perceptions, beliefs, goals, intentions, and actions
  - Preference memory resolves any potential resource conflicts
  - Learning mechanism can store new rules into production memory

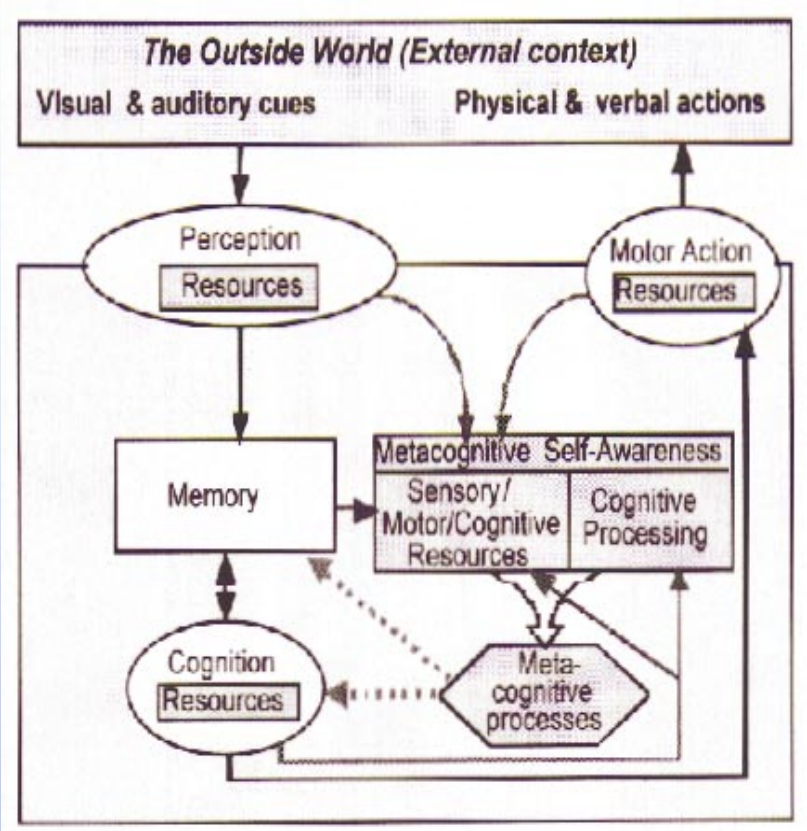
Micro Analysis & Design

# Use of Soar in Error Prediction

- Four studies that look at errors as part of the learning process
  - relevant to pilots since some pilots don't know how to use all modes of FMS
- Cognitive slowing due to fatigue
  - Failure to identify contacts leads to errors in SA
  - Not firing a weapon in time leads to compensatory actions (such as evading instead of firing)
- Potential for modeling reasoning errors, working memory decay, response to perceptual and motor errors

Micro Analysis & Design

# COGNET Overview




- *Cognitive Network of Tasks*
- Developed by CHI Systems to simulate human expertise
- Codify ways that experts in a given area might receive, process and act on info
  - Expert decision making
- Not restricted to any unified theory of human memory or memory processes.
  - Instead, allows for different theories represented

Micro Analysis & Design

# Use of COGNET in Error Prediction

- No prior error prediction research
- Does not currently mimic flawed human decision making, only expert actions
- Potential for predicting errors related to goal selection and attention

Micro Analysis & Design

# Vision Models Overview

- Use computational algorithms to simulate human visual processing of an image.
- Images usually consist of a cluttered background and a target image (usually a military vehicle) within it.
- Likelihood that attention is directed at an object is called the probability of fixation ($P_{fix}$).
  - ORACLE and the Georgia Tech models determine fixation
  - OptiMetric model must be told where to fixate

# Use of Vision Models in Error Prediction

- Error prediction capability is based on discriminating target from other clutter objects in the areas of the image that have been fixated.
  - Probability of detection ($P_d$) for accurately identifying a target
  - Probability of false alarm ($P_{fa}$) for incorrectly indicating that a fixated object is a target.
- All three vision models predict performance for target detection ($P_d$).

Micro Analysis & Design

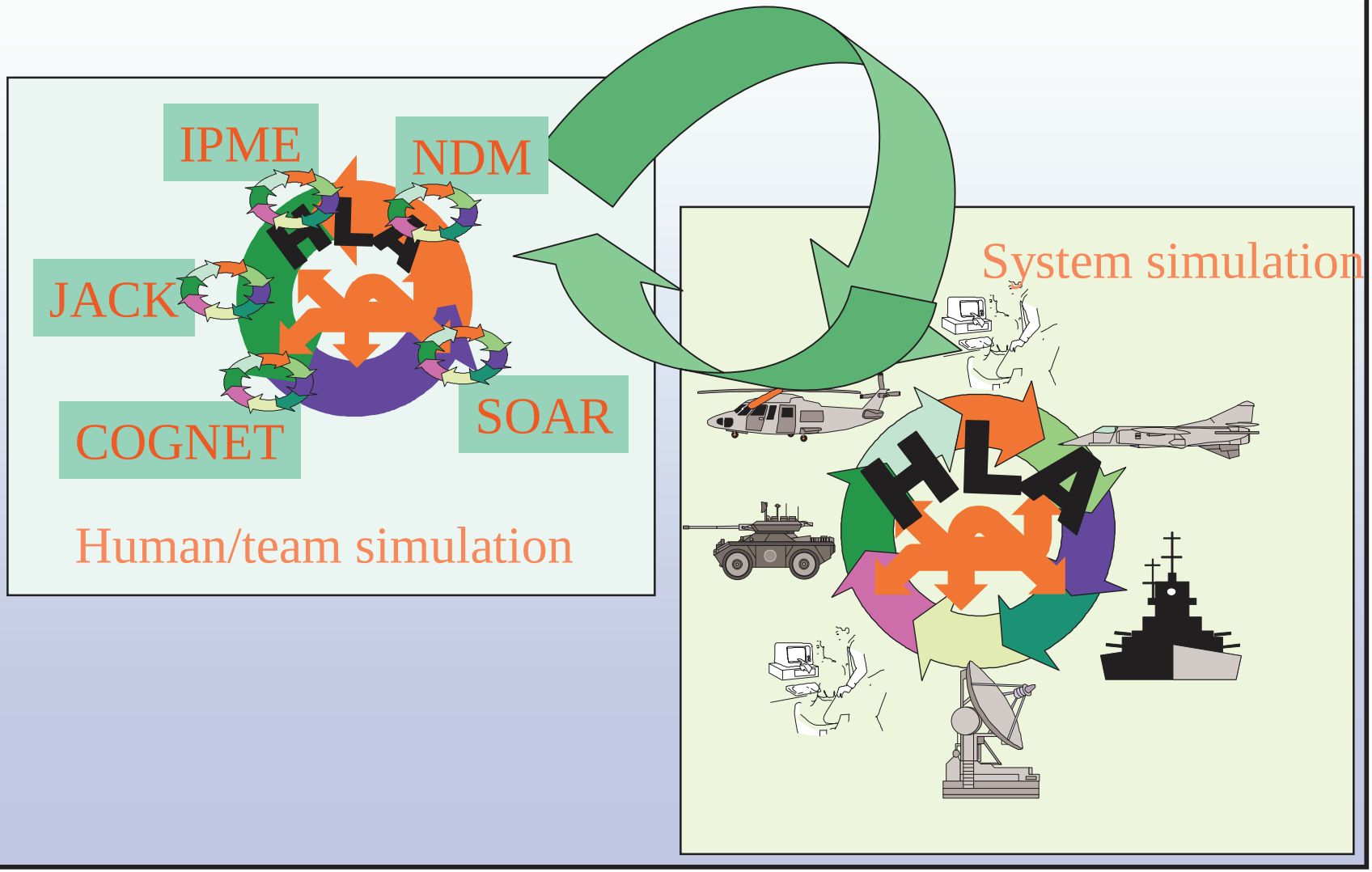
# Overview

- Introduction
- Approach
- Taxonomies of human error
- Using HPM to predict human error
- Review of HPM architectures for error prediction
- **Architecture Recommendations**

Micro Analysis & Design

# Conclusions and Recommendations

- Human error modeling is a hard modeling problem that takes time with the tools today
- No single HEM for all human error modeling questions
- None of the HPMs have proven themselves definitively in modeling human error
- Don’t ignore practicalities when making choices
- Integration with hardware software models will be needed
- Hybrid modeling approaches offer promise

Micro Analysis & Design

# An Integrative Architecture