

How to Keep Smart Agents Smart in Off-Normal Situations

Keith Daum Ph.D.

Ron Boring Ph.D.

Idaho National Laboratory

Human Factors, Controls, and Statistics

August 15, 2012

www.inl.gov



5th International Symposium on Resilient Control Systems

Decision Making in Complex Systems

Failures ?

- Love Canal
- Chernobyl
- Space Shuttle Challenger
- BP Oil Leak

Successes

- Apollo 13
- US Air 1549



Multi-Agent Systems

- An **agent** is an entity that carries out a task or a service
- An agent may be a **human** who is responsible for making decisions and acting upon them—the agent for a particular task
- Alternately, a **software** agent is a computer program that acts on behalf of a human to complete certain tasks
- In **Multi-Agent Systems**, agents of both types have interactions, usually with more than one other agent

Why Do Human Agents Fail in Decision Making?

- Ignorance of the situation
 - There are some things an individual Agent does not know
- Inexperience
 - There are situations where an Agent has the knowledge but has no experience in how to use the knowledge correctly
- Agent acts on an initial and incorrect judgment
- *“The volume and complexity of what we know has exceeded our individual ability to deliver its benefits correctly, safely, or reliably.”*
- Atul Gawande
- *“We are often an observer to what our System 1 Thinking delivers.”*
- Daniel Kahneman
- *“Long-range planning does not deal with future decisions. It deals with the future of present decisions.”* – Peter Drucker

A Failure in How We Think?

- System 1 Thinking
 - Automatic and quick
 - Based on previous outcomes
 - Biases and heuristics
 - Answers an easier question
 - Fails in new/different situations
- System 2 Thinking
 - Analytical problem solving
 - Logic and statistics
 - Experts as part of a team
 - Requires a specific effort to initiate
 - Systematic Planning
- Daniel Kahneman – “Thinking Fast and Slow”

Multi-Agent Control Systems

- **Problem** – CAS - agents use System 1 thinking in off-normal situations
 - Poor analysis, poor information and data, limited agent-agent transfer of information, and poor individual and group decisions
- **Hypothesis**
 - System 1 thinking leads to poor resilience
 - System 2 thinking is a **requirement for resilience**
- **Strategy** to overcome the potential biases and heuristics
 - Move agents into System 2, analytical decision making and resilience through better individual agent and multi-agent decisions
 - How - Checklist Manifesto
- Requires integrating across information and decision science, human factors, control systems, power systems

What is Resilience in Multi-Agent Systems?

- Agent-human interactions are important
 - Even with increasing cyber and decreasing human role
 - Agents must have the ability to respond to events for which there is no history and no example
 - In off-normal situations, human decision process drives decision making
- A resilient control system design will
 - Extend the normal reliability and functionality of a traditional control system
 - Implement situational awareness and control to the system by fully analyzing for abnormal conditions and providing a timely response to such conditions

Framework for Multi-Agent Decision Making

- Agents have more than one role
 - Individual (Autonomous)
 - Team
 - Evolutionary process to make system level group decisions
 - Share information, and evaluate, promote and defend ideas
- Guidance is needed at the agent level to determine:
 - When individual and when part of team decision?
 - Individual - how do they select and evaluate information and when do they need to actively seek additional information?
 - Team - how to present information fairly and honestly and how to evaluate the information presented by other agents?
- In either setting, it is important to understand:
 - What are an agent's trusted lines of communication?
 - Why is each trusted and why are others not trusted?
 - What to do when two trusted sources of information conflict?

Multi-Agent Information “Competition”

- In team decisions, information from agents competes in a three phase evolutionary type decision making process.
 - A **random phase**, in which agents contribute observations, ideas, and innovative insights
 - A **selection and growth phase**, in which agents engage in open discussion and consider alternative viewpoints; participants examine issues both in depth within a discipline and also across disciplines
 - An **emergent or organization and amplification phase**, with group discussions to develop insights from the two earlier phases into a decision
- Competition is meant to be **complementary**

Information Transmission

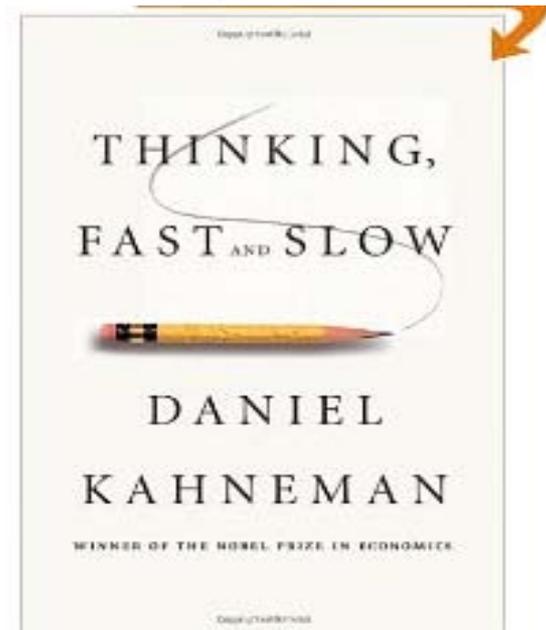
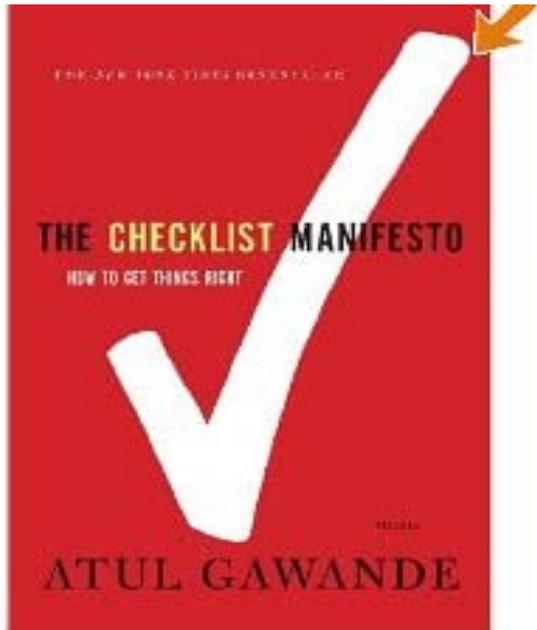
- Content and Transmission are important
- Transmitter of the message
 - Prepare the right message with right content
 - Transmit the message with 100% efficiency
- Receiver
 - Receive the message
 - Decode the message correctly
 - Understand the content of the message as intended
 - Implement the message as intended
 - All of this should also be done with 100% efficiency.
- A more useful message
 - Put reliability of the information first and then the story
 - Prevents receiver from keying in on a potentially biased “story”

Off-Normal Situations

- During an off-normal event, success is the result of team work and adherence to strict discipline and process
- When decisions have to be made quickly, getting the right steps accomplished is critical
- Under conditions of complexity, Agents need room to act and adapt
- In off-normal conditions, Agents require a seemingly contradictory mix of freedom while having an expectation to co-ordinate and measure progress towards common goals
- Communication is an integral part of recognizing and dealing with unanticipated problems

We Need a Different Strategy

- A strategy that takes advantage of the knowledge that people have while making up for human inadequacies
- The How
 - *The Checklist Manifesto*
 - By Atul Gawande
- The Why
 - *Thinking Fast and Slow*
 - By Daniel Kahneman



Two Kinds of Checklists

- Read-Do
 - Like a recipe or some Standard Operation Procedures
 - Useful for routine and repetitive operations that do not change
- Do-Confirm
 - Don't tell you what to do, but rather are a guide to how to think, especially in challenging situations

Do-Confirm Checklist Approach

- A Guide to Thinking Checklist: A Simple Tool For
 - Reduce human errors and open opportunity for creative thinking
 - Assure that easy tasks get done when needed
 - Provide reminders of the most important and critical steps
- Right conversations, right people, right content, at right time
 - Each expert in a team has critical information for overall “story”
 - Getting right experts to function as a team is critical for success
- Leave room for craft, judgment, and responding to the unexpected
- Precise, short, and practical so useful during off-normal situations
- Pause points - Information complete and adequate before next step
- Transparent and defensible
- The ticking the boxes in check list is not the ultimate goal
- **Embracing a new culture of team work and discipline is the goal**

Defects in Thinking Checklist

- **Agent-Human Self-Check Questions**
 - Self-Interest Bias
 - Affect Heuristic
 - Groupthink
- **Other Agent-Human Check Questions**
 - Saliency Bias
 - Confirmation Bias
 - Availability Bias
 - Anchoring Bias
 - Halo Effect
 - Sunk-Cost Fallacy
- **Decision Quality Check Questions**
 - Overconfidence, Planning Fallacy, Optimistic Biases, Competitor Neglect
 - Disaster Neglect
 - Loss Aversion

Successful System 2 Agent-Human Decisions

- Agents
 - Separate System 1 and System 2 thinking modules
 - Recognize normal and off-normal situations
 - Empowered to decide and act autonomously (also locally)
 - Decisions with network-wide consequences use team decision making
 - Communicate with each other and with human information gatherers and decision makers
 - Easily reconfigured to analyze new information and perform what-if walkthroughs

Successful System 2 Agent-Human Decisions

- Communications
 - Be useful for making decisions
 - Include reliability and uncertainty information **first**
 - Include identification and trustworthiness of the sources
 - Include off-normal assumptions and realities such as failed sensors or system functionality
 - Timely, and in-progress information is communicated so that it may be considered when available

Successful System 2 Agent-Human Decisions

- Decision making
 - Looks for a complete set of useful information from multiple agents and/or human sources
 - Analyses the information wisely and completely and has an unbiased judgment
 - Overrides individual agent autonomy only to the extent that decision making does not cause an impasse or otherwise slow resolution of the off-normal event
 - Allows a “time out” to let agent act autonomously when needed
 - Uses “pause points” at critical steps so that agents function as a team to make sure that information is complete and adequate before passing to the next step
 - Follows checklist criteria to ensure proper transition from System 1 to System 2 thinking

Conclusions

- Resilience can be based on letting agents be smart
- Smart comes from an independent and impartial evaluation of the information in independent and team decisions
- During an off-normal event, there is a point where smart agents move from using protocols to making their own decisions. This is the point where the checklist is most important, to guide the agent into System 2 thinking and minimize the use and impact of System 1
- The checklist gets the agent to focus on relevant information and serves as a basis for the identification of corrupted information and corrupt information sources. The checklist also ensures that human decision makers do not fall back on System 1 thinking that may prove irrelevant to the off-normal event
- Ensure that human decision makers also are guided to the smartest and most successful course of action to resolve the off-normal event.

Backup Slides

Decision Scenarios

- You are the Decision Maker
 - You suffer the consequences or get the rewards
- You provide support to the Decision Maker
 - You are authorized to analyze and recommend, but you are not the Decision Maker
 - Decision Maker risk tolerance may not be known
- You direct decision support to the Decision Maker
 - Frame the problem and the risk factors
 - Limit the alternatives to be considered
- A committee has responsibility for the decision
 - There are often hidden agendas, including individual risk tolerance
- Collective Decisions
 - Knowledge Ecosystems – an open, bottom-up process for cultivating knowledge to make decisions
 - Mitigates hidden agendas
 - Fosters useful weak signals
 - A way of overcoming individual analyst or data provider shortcomings
 - Consensus decisions

Decisions

- A Decision is often about an allocation of resources
- A “Decision Maker” is the person who has authority to allocate resources
- An “Objective or Goal” is what the Decision Maker hopes to achieve with the resources
 - Prioritization of alternatives may be needed
- A good decision has clarity of action
- One measure of success is how satisfied the Decision Maker is with the use of resources to achieve the objective

Systematic Planning

Elements

- Define the problem
- ID decision & decision maker
- ID cost of incorrect decision
- Quality and quantity of data needed to make the decision
 - Data Analysis Techniques
 - Decision Limits
- Are data already available?
- Statistical plan for additional data
- Specify QA/QC activities
 - To assess performance
 - To provide defensibility