## 1 - Motivation

Anticipatory thinking (AT), the deliberate and divergent exploration of relevant possible futures [1], is a key concept in formal definitions of intelligence analysis[2], government mandated deliverables[3], and the strategic foresight discipline. They all rely on AT to evaluate the current and possible future state of the world. We also use AT in our everyday lives when exogenous events compel us to use deliberate and divergent cognitive processes to explore relevant possible futures that help us prepare. Successful AT identifies impactful future events such that opportunistic or preventative actions are taken in advance. Failures range from coming to the incorrect expectation (a failure of accuracy and/or precision) to not considering trajectories that turn out to be important (a failure of recall). Such failures limit our ability to identify and mitigate risk. Because there are many ways our preparations can succeed or fail, we are subject to surprise and exposed to risks as the future unfolds. Dwight Eisenhower perhaps said it best:

"Plans are worthless, but planning is everything"– Dwight Eisenhower

This quote throws into sharp relief that any one plan is not likely to execute as prescribed. Rather, a methodology of designing a plan harnesses our innate future-thinking ability as it requires a high degree of vigilance and deliberate exploration of divergent options and contingencies. After diligently following a plan design methodology, one emerges with the knowledge to act effectively as the future unfolds.

This symposium will focus on the intersection of cognitive systems and prospective cognition communities with specific investigations into AT challenge problems. Developing challenge problems will both extend existing methods to AT tasks and initiate development of new ones.

## 2 - Problem

Despite AT's pervasiveness in everyday decision making, there is paucity of metrics, measures, and challenge problems that could assess the efficacy of AT methods. Guilford tasks assess divergent thinking, viewed as essential to creativity, but do not have a future-oriented posture. While the Anticipatory Thinking Assessment (ANTA) is a future-oriented divergent thinking task, its transferability has not been evaluated and it does not address AT methods. DARPA's SAIL-ON program strives to generate and respond to novelty in environments, but of the kind that are truly novel and cannot be anticipated.

## 3 - Approach and Challenge Problems

Our approach builds on the success of the COGSAT 2019 Fall Symposium by addressing the lack of problems and measures that could define and assess successful AT. Our goals are:

1. Develop four challenge problems that can effectively assess the ability to initiate and replicate AT using AI cognitive systems.

2. Release initial versions of the challenge problems in advance of the symposium so participants may submit position papers (for archival) suggesting modifications or preliminary results from new or existing methods.
3. In addition to presentations of accepted papers at the symposium, facilitate 4 breakout sessions designed to discuss next steps and refinements to each challenge problem.
4. Post symposium, the organizers will investigate and publish unifying views (e.g. methods, representations) of cognitive systems ability to initiate and replicate AT.
5. Identify challenge problems suitable for a venue with focused community efforts.

### 3.1 - Challenge Problem 1: PAIR

Plan, Activity, and Intent Recognition (PAIR) is the broad challenge of predicting an observed agent's behavior (ongoing or future plans, activity, or intent) on the basis of past behavior. PAIR's use of both intention and planning, two of the four prospection modalities[4], intimately links it to AT: the deliberate and divergent exploration of relevant possible futures. PAIR research, however, focuses on unobtrusive and convergent exploration. There are therefore two open questions PAIR methods must answer to become more relevant to the AT enterprise.

1. How can we determine a range of possible future behavior in which predictions are materially divergent? Current PAIR methods cannot systematically arrive at predictions of statistically and meaningfully different future behavior on the basis of prior behavior.
2. How might an observer agent act within the observation environment to obtain more information about the actor's purported future behavior? Current PAIR methods cast the observer as "disembodied," unable to affect the dynamics of the task environment.

***Concrete Challenge 1***: Adapt PAIR methods for agents in strategy games (turn-based or real-time) to actively elicit other player goals and obscure their own. This could take the form of an assistant agent that overcomes limits of working memory and initiates AT in human players by suggesting actions to manage risk exposure to another player's capabilities. Another form could be agents who themselves actively manage imperfect information about their opponent and who can reason about how to induce ambiguous or false beliefs in their opponent.

### 3.2 - Challenge Problem 2: Goal-reasoning

Goal-reasoning agents are capable of formulating and managing their goals in response to dynamic environments. Automated planning supports goal reasoning by determining the optimal or diverse course(s) of action to achieve one or more goal(s). However, little work has been done to analyze future consequences, like risk, of pursuing a goal at the strategic or mission level, or actions at the tactical level.

The goal-reasoning challenge problem addresses these limitations by introducing two novel aspects to a domain-problem. The first is to identify key anticipatory actions. Examples of early anticipatory actions may be (1) those that allow the agent to extend the number of actions it can execute (i.e. by carrying an extra battery) at the cost of another resource (i.e. carrying extra battery reduces space for storing found items), (2) actions that reduce detrimental effects of future actions or events (i.e. packing a grappling hook enables a quick recovery from falling into a pit), or (3) actions that will improve future state observations (i.e. packing a higher quality but heavier sensor). A second aspect to the AT challenge problem is the degree of uncertainty that makes it impossible for an agent to plan for more than one mission at time. Since the outcome of any single mission is not known and will affect the agent's ability to achieve future missions. For example, the agent may have a default risk-tolerance in the first few missions yet may increase or decrease the risk-tolerance after each mission when considering its progress on its overall agenda. These two aspects are driven by two research questions:

1. How can an agent determine actions that can be executed early in the plan to aid in achieving the most number of goals over the course of a single plan execution episode?
2. How can an agent balance selecting missions with an overall agenda?

***Concrete Challenge 2***: Extend the IPC "Rover" domain. At the tactical level, operationalize the concept of a "base" where the rover can choose between different restocking supplies, deposit items to save or use later, refuel, augment its capabilities with different items that enable actions out in the world (i.e. grappling hook, portable solar recharging station, etc). At the strategic level, operationalize the concept of a "life purpose" or lifelong agenda. Successful AT goal-reasoning systems capable of planning and executing those plans in a simulator that has the base and world subdomains, as well as local missions vs. a lifelong agenda. The lifelong agenda is closer to what the agent was designed for, it's overarching purpose which is a more abstract goal such as "collect information about the Mars environment".

### 3.3 - AT Challenge Problem 3: Learning without Experience for Autonomous Vehicles

Errors in autonomous vehicles, whether fatal[1] or uncomfortable[2], are due to flaws that a human driver would never make. In fact, human drivers are able to reason about circumstances they have not seen before, while autonomous vehicles are completely dependent on their ability to learn through (previously observed) experiences. But these "difficult autonomous scenarios," are not yet characterized nor benchmarked. In fact, there is a complete void of erroneous autonomous vehicle data. The current self-driving data sets are hand-curated and perfectly labeled with challenges that encourage precise or tailored execution[3]. Developing difficult scenarios will ensure that autonomous vehicles can reason and anticipate unknown(erroneous) futures, resulting in greater safety, trust, and reliability. The AV challenge will build these types of scenarios based on the following two questions:

1. How can computational models generate scenarios that require robust solutions: solutions that can only be achieved through sophisticated AT-reasoning?
2. How does knowledge representation interact with AT reasoning to affect scenario outcomes?

***Concrete Challenge 3*:** Develop (i) the mechanisms to generate a set of self-driving error scenarios that require abstract, high-level anticipatory thinking: the same type of reasoning humans do in difficult situations.Ensure autonomous vehicles can also learn new things, without failing first by developing (ii) a flexible representation (e.g. conceptual primitives, frame-based representations, etc.) and (iii) various types of reasoning (e.g. commonsense, hypothetical [5],

---

1 Uber self-driving vehicle strikes and fatally kills a pedestrian:https://www.nytimes.com/2018/03/19/technology/uber-driverless-fatality.html

2 Consistent start/stop in self-driving vehicles leads to a "herky jerky" ride:https://www.wired.com/story/ride-general-motors-self-driving-car/

3 Carla autonomous driving challenge scenarios: https://carlachallenge.org/challenge/nhtsa/

and analogical reasoning [6]). Developing these capabilities, will result in adaptive, autonomous vehicles that can reason and address the ever-growing, long tail of errors.

### 3.4 - AT Challenge Problem 4: Cognitive Psychology

Cognitive systems often leverage cognitive models to assess the system's ability to elicit a cognitive response. However, cognitive models vary in their formalism, resulting in some that are not precise enough for cognitive systems research. This challenge problem's goal is to develop AT cognitive models and assessments in lock step with AT cognitive systems.

This challenge problem will focus on the use of counterfactuals as a method to perform AT. Counterfactuals initiate the use of our imagination, which is inherently limited by prior experiences, memory accessibility, and cognitive ability. Recent research has elucidated a deeper connection between human ability to engage in counterfactual thinking, prospection, theory of mind, and imagination. This domain-general ability to envision alternative scenarios through mental simulation poses difficult challenges for cognitive systems. This challenge problem will lay the human centric foundations for assessing cognitive systems that initiate and replicate AT using counterfactual reasoning leading us to the following research questions:

1. How can we measure the effect of cognitive sys. generated counterfactuals to elicit AT?
2. How can we measure a cognitive system's use of counterfactuals to perform AT?

***Concrete challenge 4***: Design and develop ***advanced forecasting*** for geo-political dynamics. Much like advanced chess uses human-computer teams, roughly delineating the imaginative strategic tasks and mechanical tactical tasks to humans and computers, respectively, forecasting has a similar intuitive partition. While assigning a probability to an event has been a research topic that uses both computers and humans (a tactical task), identifying the important events that are worthy of a forecast has been entirely human driven and disconnected from forecasting (a strategic task). ***Advanced forecasting*** integrating identifying events (likely in the geo-political domain), an anticipatory thinking task best done by humans with support from cognitive systems, with assigning likelihoods, best done with algorithms supported by humans.

## 4 - Format

Given our focus on collaboration from two disciplines and strong application contexts, we will offer several ways to engage. Our format will include:

- Shorter talks (15-30 minutes) for papers accepted for presentation
- Invited talks (45-60 minutes) for academic, government, and industry practitioners
- Breakout sessions (60 minutes) to refine challenge problems
- Poster sessions (90 minutes) for preliminary work, case studies, and collaboration

We invite contributions as: papers (long and short), case studies, and poster abstracts.

## 5 - Organizers

Dr. Adam Amos-Binks (aamosbinks@ara.com) is the Chief Scientist of AI Research at ARA. He brings 15 years of intelligence community experience including managing the Laboratory for Analytic Sciences' Anticipatory Thinking research program at North Carolina State University. His research combines Artificial Intelligence, specifically intention recognition and revision, with cognitive models of comprehension to investigate prospection and imagination.

Dr. Dustin Dannenhauer (ddannenhauer@navatekllc.com) is a Scientist at Navatek, LLC specializing in artificial intelligence. His research interests span goal reasoning, planning and execution, cognitive architectures, and metareasoning. He has co-organized workshops on Goal-Reasoning, Explainable Planning, and served as the Publicity Chair for ICCBR-18.

Dr. Rogelio E. Cardona-Rivera (rogelio@eae.utah.edu) is an Assistant Professor in Computing and Entertainment Arts and Engineering at the University of Utah. His research focuses on the design of interactive narrative games, systems where users direct an unfolding story by taking on dramatic roles.

Dr. Gene Brewer (gene.brewer@asu.edu) is an associate professor in cognitive psychology in the Department of Psychology at Arizona State University. His interests include behavioral and electrophys. recording of memory processes, prospective memory, source memory, recognition memory, individual differences in working memory, and emotional effects on memory.

Leilani Gilpin (lgilpin@mit.edu) is a PhD candidate in Electrical Engineering and Computer Science at MIT. Her research interests include the theories and methodologies of designing and augmenting machines that can explain themselves.

## 6 - Participant Interest

## 7 - References


[1] Geden, M., Smith, A., Campbell, J., Spain, R., Amos-Binks, A., Mott, B., … Lester, J. (2019). Construction and Validation of an Anticipatory Thinking Assessment. Frontiers in Psychology, 10(December), 1–10.

[2] Breakspear, A. (2012). A New Definition of Intelligence. Intelligence and National Security

[3] Clapper, J. (2014). The National Intelligence Strategy of the US 2014, 1–24.

[4]Szpunar, K. K., Spreng, R. N., & Schacter, D. L. (2014). A taxonomy of prospection: Introducing an organizational framework for future-oriented cognition. Proceedings of the National Academy of Sciences, 111(52), 18414–18421.

[5] Blass, Joseph A and Forbus, Kenneth D. Modeling commonsense reasoning via analogical chaining: A preliminary report. CogSci, 2016.

[6] Holmes, Dylan Alexander. Story-enabled hypothetical reasoning. Master's Thesis. Massachusetts Institute of Technology, 2017.