

# Anticipatory Thinking: A Metacognitive Capability

Adam Amos-Binks<sup>1</sup> and Dustin Dannenhauer<sup>2</sup>

<sup>1</sup>Applied Research Associates  
Raleigh, North Carolina 27614  
aamosbinks@ara.com

<sup>2</sup>Navatek LLC  
Arlington, Virginia 22203  
ddannenhauer@navatekllc.com

## Abstract

Anticipatory thinking is a complex cognitive process for assessing and managing risk in many contexts. Humans use anticipatory thinking to identify potential future issues and proactively take actions to manage their risks. In this paper we define a cognitive systems approach to anticipatory thinking as a metacognitive goal reasoning mechanism. The contributions of this paper include (1) defining anticipatory thinking in the MIDCA cognitive architecture, (2) operationalizing anticipatory thinking as a three step process for managing risk in plans, and (3) a numeric risk assessment calculating an expected cost-benefit ratio for modifying a plan with anticipatory actions.

## Introduction

Anticipatory Thinking (AT) is the deliberate and divergent analysis of relevant future states that is a critical skill in medical, military, and intelligence analysis (Geden et al. 2018). It differs from predicting a single correct outcome in that its goal is to identify key indicators or threatening conditions so one might proactively mitigate and intervene at critical points to avoid catastrophic failure. This uniquely human ability allows us to learn, and act, without actually experiencing. AI systems with this robust capability would support the autonomy and contextual reasoning needed for next generation AI.

However, AI systems have yet to adopt this capability. While agents with a metacognitive architecture can formulate their own goals or adapt their plans in response to their environment (Burns and Ruml 2012; Cox 2016) and learning-driven goal generation anticipates new goals from past examples (Pozanco, Fernández, and Borrajo 2018), they do not reason prospectively about how their current goals could potentially fail or become attainable. Expectations have a similar limitation, they represent an agent’s mental view of future states and are useful for diagnosing plan failure and discrepancies in execution (Muñoz et al. 2019) but do not critically examine a plan or goal for potential weaknesses or opportunities in advance. Duff et al. 2006 use an

explicit knowledge representation in the domain to ensure that maintenance goals do not fail while proactively achieving maintenance goals when they do not conflict with existing achievement goals. This is a similar goal to anticipatory thinking, but its computation of goals is more akin to prediction than anticipation. At present, agents do not analyze plans and goals to reveal their unnamed risks (e.g. such as actions of another agent) and how they might be proactively mitigated to avoid execution failures. Calls to the AI community to investigate *imagination machines* (Mahadevan 2018) highlight the limitations between current data-driven advances in AI and matching human performance in the long term.

To address this limitation, we take a step towards *imagination machines* with a contribution that operationalizes the concept of anticipatory thinking, a cognitive process reliant on an ample supply of imagination, as a metacognitive capability. We propose this capability as a kind of solution formulation method, a post-planning step that analyzes a solution plan for potential weaknesses and modifies the solution plan to account for them. This approach is in contrast to problem formulation, a pre-planning step that analyzes a problem for efficient search strategies, as well as online risk-aware planning processes (Huang et al. 2019). Our first step of AT identifies properties of a plan that are prone to failure. These include concepts such as atoms needed throughout a plan but are only achieved in the initial state. As a second step, we extend goal-reasoning agent expectations to include anticipatory expectations, a kind of expectation derived from a plan’s relevant states that identifies exogenous sources that could potentially introduce failures. Finally, we define anticipatory reasoning to proactively mitigate the potential failures. An agent reasons over the conditions in the anticipatory expectations, generating anticipatory actions to be executed at specific times, foiling an exogenous source of failure. To exercise this new capability we use a simple example and define metrics for evaluating an agent’s anticipatory thinking.

## Previous Work

Our contributions are based on three related areas of work. Prospective cognition is a fledgling field in cognitive psy-

chology the goal of which is to understand human ability to reason about and imagine the future. We discuss some prospection modalities. The second area, goal-reasoning agents, is a type of agent that adapts to and formulates their own goals in response to their environment. We highlight some of the overlap between prospection modalities and the agent’s methods for formulating and achieving goals. Finally, investigations into metacognition’s role in decision making and behavior draws a close tie with autonomy. We detail some of the existing capability to frame anticipatory thinking’s role.

## Anticipatory Thinking

Anticipatory thinking is an emerging concept in psychology (Geden et al. 2018) that captures the cognitive processes in use when preparing for the future. The deliberate consideration of a diverse set of possible futures which aggregates imaginative, divergent, and prospective processes and is more than any of the individual processes alone. Imagination is a mechanism to reason about what is outside our immediate sensory inputs. More than an artist’s creative reservoir, imagination drives the creativity in complex sciences from engineering to finance. Imagination is used to reason about details in problem-solving, such as what might have happened in a mystery novel, as well as generating novel ideas through methods such as counterfactual reasoning. Calls to the AI community to investigate *imagination machines* (Mahadevan 2018) highlight the gap between current data-driven advances in AI and matching human performance in the long term.

Divergent thinking is often used to assess individual differences in creativity and has been part of scientific studies on creativity since the 1960’s (Guilford 1967). Assessing divergent thinking asks subjects to perform divergent thinking tasks, the scores and measures of which are still the focus of numerous studies (Silvia et al. 2008). Physical limitations such as working memory and recall from long-term memory have been the source of inspiration for developing methodologies to counteract them (e.g. structured analytic techniques (Heuer 2008)).

Lastly, the emerging field of prospection is the ability to reason about what may happen in the future. Szpunar et al. 2014 provide a taxonomy of prospection that covers four modalities (planning, intention, simulation, prediction) in both syntactic and semantic spaces. Several AI research communities have investigated the methods that, at least in name, overlap with the modalities but have lacked the unifying taxonomy to characterize them in prospective cognition.

## Goal Reasoning

One approach to mitigate risks is to encode mappings from states to goals, such that when an agent is in a state, it should pursue the corresponding goal. Thus, if risks are known at design time, an agent can be given mappings from risky states to mitigating goals. MADBot (Coddington et al. 2005) investigated goal formulation via motivator strategies within, and external to, the planning process. An example of a motivator function is the following: a rover robot may have

the motivator function that when its battery level drops below a threshold, the agent will generate a goal to have a fully charged battery. This would then be achieved by a plan for the rover to navigate to the power source and plug itself in. As shown in Coddington (2005) these motivator functions can be either (1) encoded into the plan operators as constraints (i.e. every action has a precondition that the battery level is above a threshold) or (2) a separate goal formulation process which runs outside the planner and generates a new goal when motivator functions trigger.

Other approaches to mitigating risk with planning systems include rationale-based monitors (Veloso, Pollack, and Cox 1998), perceptual-based plan monitors (Dannenhauer and Cox 2018), and contingency planning (Hoffmann and Brafman 2005). Prior work on mitigating risk during plan execution has considered monitoring rationales for goals (Dannenhauer 2019). In MADBot and other work on goal motivator strategies (Muñoz-Avila, Wilson, and Aha 2015), goal motivator functions are known at the design time of the agent. The primary difference of the approach presented here is that goal formulation strategies are identified automatically at runtime by anticipatory thinking approaches using the plan solution as a source of information.

## Metacognition

Metacognition refers to processes that reason about cognition in some form or another (Cox, Raja, and others 2011). We use the Metacognitive Integrated Dual-Cycle Architecture (MIDCA) to discuss anticipatory thinking processes. A primary benefit of MIDCA is its explicit separation of cognitive and metacognitive processes. Cognitive processes (see Figure 1) are those that are more directly concerned with the world (goals are world states, plans are sequences of actions that act on world states, etc). Metacognitive processes (see Figure 2) are those that are more directly concerned with cognitive processes and states (identifying and resolving issues such as impasses that arise in various cognitive processes). One of the core assumptions here is that the agent’s mental state is separate from the world state (otherwise reasoning about world states would also be reasoning about cognitive states).

At a general level it seems that AT could be considered a cognitive process since the objective of AT is to prevent risk that arises from various world states in order to achieve some goal that is a world state. When considering specific AT processes (presented in the next section) we argue that AT is truly a metacognitive process since it is concerned with meta goals such as *achieved(g’)* where *g’* is a cognitive level goal. AT is also concerned with decision making on resource trade-offs (a type of metareasoning) for risk mitigation (i.e. spending X extra actions to mitigate Y potential risks). Additionally, if AT processes were to take into account an agent’s likelihood of succeeding at a task, than AT processes are making use metacognition self-prediction mechanisms.

MIDCA is currently under active development, and until recently most work has consisted of implementation at the cognitive level. Prior work on the metacognitive level includes monitoring capabilities that maintain a cognitive trace and control actions capable of switching planning al-

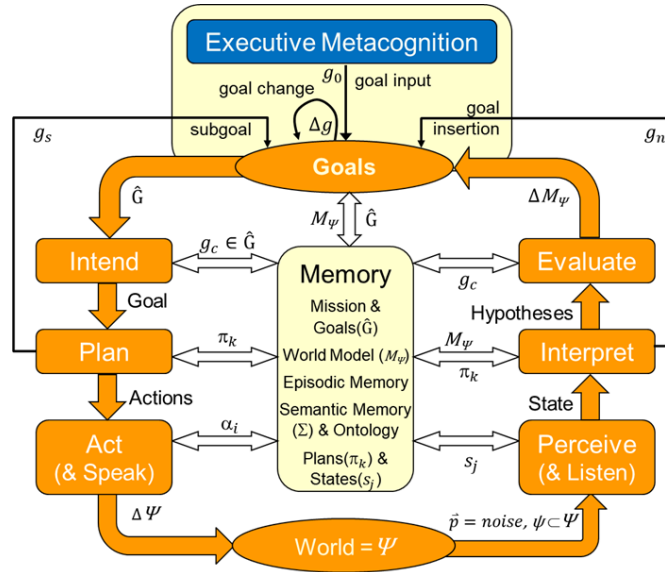


Figure 1: MIDCA cognitive level.

gorithms at runtime (Cox, Dannenhauer, and Kondrakunta 2017) and domain independent expectations<sup>1</sup> of cognitive processes (Dannenhauer, Cox, and Muñoz-Avila 2018). The primary contributions from these works have mostly focused on the Monitor, Interpret<sup>2</sup>, and Control phases of the metacognitive layer. The AT process we describe in this paper proposes additional new methods to the Intrepret, Plan, and Control phases of the metacognitive level.

## Anticipatory Thinking as Metacognition

Our approach to operationalizing anticipatory thinking begins with the concept as explained in (Geden et al. 2018) as "deliberate, divergent exploration and analysis of relevant futures to avoid surprise". We define three steps that operationalize the (i) deliberate, (ii) divergent, and (iii) relevant components of the above AT concept.

First, in Section 3.1 we identify goal vulnerabilities as an example method in the deliberation step. This step reasons over a plan's structure to identify properties that would be particularly costly were they not to go according to plan. A second step, failure anticipation, identifies sources of failure for the vulnerabilities in Section 3.2. Sources of failure can range from unknown environment states to other agent's interfering goals. Finally, in Section 3.3 we detail a failure mitigation step that modifies an existing plan, reducing the exposure to the sources of failure and creating an anticipatory expectation. We capture these steps as a process in Table 1

<sup>1</sup>It is worth noting that while expectations at the cognitive and metacognitive levels have analogous roles, their definitions at each level are different.

<sup>2</sup>The Interpret phase generally includes discrepancy detection, explanation/diagnosis, and goal formulation. Of these, discrepancy detection is the only one with prior work at the metacognitive layer of MIDCA.

and demonstrate these steps through an *NBeacons* running example from (Dannenhauer, Cox, and Muñoz-Avila 2018).

Table 1: This method takes as input a plan and returns a modified plan that includes anticipatory actions.

	METHOD: Anticipatory Thinking INPUT: A plan $\pi$ OUTPUT: A modified plan $\pi'$
1	identify goal vulnerabilities in $\pi$
2	identify failure anticipation in $\pi$
3	edit $\pi$ with failure mitigations to get $\pi'$
4	return $\pi'$

To illustrate these ideas, it is useful to consider following example. An agent must generate plans to reach beacons and activate them. If the agent ever passes through a sandpit square, they must take three actions to dig out. The wind may blow in a known direction at a known speed after every agent action. The wind pushes an agent a number of squares further (equivalent to the speed) in the wind's direction and can result in an agent passing over a sandpit and getting stuck. In our example, the wind is blowing West at a speed of five making an agent's plan vulnerable to any sandpit that lies within five squares West of their location.

We use the Partial Order Causal Link (POCL) representation (Penberthy and Weld 1992) for an agent's plans. The main advantage to using POCL over other plan representations is that *causal link threats* can explicitly represent potential failures from external events. We use the typical definitions for the POCL representation from (Penberthy and Weld 1992) where a POCL plan consists of steps (S) that are ground actions from the domain model, bindings (B) that map free variables to literals, step orderings (O) that con-

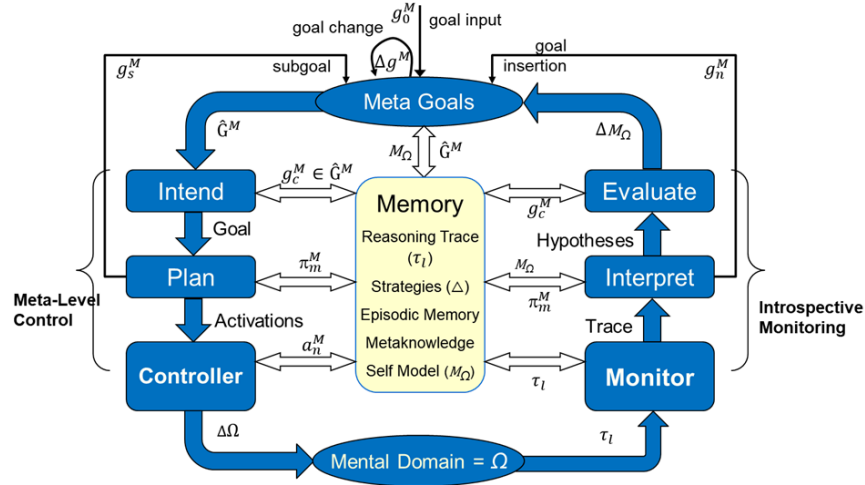


Figure 2: MIDCA metacognitive level.

strain when steps must execute relative to one another, and causal links (L) that connect steps to one another when an effect of one step instantiates a precondition for a following step. While the above plain English definitions of the individual aspects of the POCL representation will suffice for those who are familiar with planning, we provide a formal definitions for causal link threats as they are key to our choice of POCL.

**Definition 1 (Causal link threat)** A causal link threat occurs when a causal link  $s \xrightarrow{p} u$  between steps  $s$  and  $u$  for literal  $p$ , and some other step  $w$  has effect  $\neg p$  and could be executed after  $s$  but before  $u$ . Executing  $w$  in this interval establishes  $\neg p$  making the precondition  $p$  of  $u$  no longer satisfied by the effect  $p$  of  $s$  and thus  $u$  will not execute.

### Goal Vulnerabilities

Identifying a plan's vulnerable structural properties is the first step in proactively mitigating its failure. We define a single vulnerability, precondition strength, of what could be numerous vulnerable properties of a plan. Precondition strength, is a measure of how many times a precondition is established and used in plan. The fewer times a precondition is established and the more it is used increases the vulnerability of a plan to failure. It is useful in AT to identify literals that have a weak precondition strength, as their failure can require many repairs to a plan or eliminating an area of the solution space entirely.

**Definition 2 (Precondition Strength)** Precondition strength of a plan  $\pi$ ,  $\text{PRESTRENGTH}(\pi)$ , is a set of tuples  $\langle a, k, e \rangle$  where  $a$  is a literal,  $k$  is the number of steps in  $\pi$  that use  $a$  as a precondition, and  $e$  the number of times  $a$  is an effect before it is first used as precondition.

Our example in Figure 3, A is the agent's location, ~ are the sandpits, and 1 is the beacon location, and the optimal path from the agent's location to the beacon is in orange. The

optimal path comprises of eight move actions all of which have the (canMove) precondition. This precondition is only established once in the initial state and so its entry in  $\text{PRE-STRENGTH}(\pi)$  is  $\langle \text{canMove}(\text{agent}), 8, 1 \rangle$ .

### Failure Anticipation

Vulnerabilities are not by themselves indicative that a plan is at risk of failure. Risk of failure requires some means to exploit the vulnerability, what we will call conditioning events. We approach identifying these events as a kind of prefactual reasoning (future-oriented counterfactual reasoning), where we take the negation of the most vulnerable preconditions (identified in Section ) and identify actions with them as effects.

**Definition 3 (Conditioning Event)** The conditioning events  $\text{CE}(\pi)$  of a plan  $\pi$  are actions in the domain model such that one or more effects of each action is the negation of a precondition of a step in  $\pi$ , introducing a causal link threat.

In our *NBeacons* example, wind may blow after each agent action and gives rise to a conditioning event. If the wind blows at any point where a sandpit is within five squares West of the agent, a causal link threat is introduced by the effect  $\neg \text{canMove}(\text{agent})$ . This results in four potential conditioning events, one after each of the agent's first four moves, indicated by the red CE notation in Figure 3. After finding themselves in a sandpit, an agent must spend three actions digging out of the sand pit in order make  $\text{canMove}(\text{agent})$  true and be able move again. We term these three dig actions as the impact of a conditioning event.

**Definition 4 (Impact)** The impact of a conditioning event,  $\text{impact}(\text{CE})$ , is the cost of the actions needed to remove the causal link threat introduced by the conditioning event.

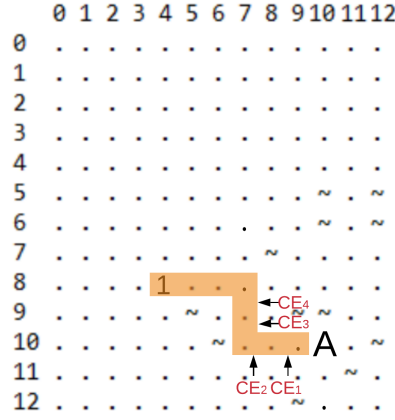


Figure 3: An *NBeacons* example where A is the agent’s location, ~ are the sandpits, and 1 is the beacon location, and the optimal path from the agent’s location to the beacon is in orange. Wind may blow after each of the agent’s actions and gives rise to the four conditioning events that could take place and are indicated by  $CE_{1-4}$ .

### Failure Mitigation

The final step to operationalizing anticipatory thinking is to define what the relevant property means for a goal-reasoning agent. We term this step failure mitigation where the agent reasons over conditioning events to identify actions that reduce a plan’s risk exposure to these conditioning events. These actions are anticipatory actions.

**Definition 5 (Anticipatory Actions)** The anticipatory actions  $ANT(\pi)$  of a plan  $\pi$  is a set of tuples  $\langle A^{ANT}, CE \rangle$  where  $A^{ANT}$  is an action sequence,  $a_1, a_2, \dots, a_n$  added to  $\pi$  such that at least one effect reduces the impact of the conditioning events CE in  $CE(\pi)$ .

To mitigate the unpleasantness of being blown about by the wind and getting trapped in a sandpit, our agent has the option to outfit itself with a grappling hook. A grappling hook allows an agent to move out of a sandpit in a single action. However, adding the grappling hook adds action costs of one for the *buy* and *pack* steps that need to be executed. We add these two anticipatory actions to the agent’s plan before the journey begins, see  $\pi'$  in Figure 4.

Adding the grappling hook to the plan creates an expectation within an agent that risk exposure to the wind conditioning event has been reduced. We refer to this new type of expectation as an anticipatory expectation and define it as:

**Definition 6 (Anticipatory Expectations)** An Anticipatory Expectation is the action cost reductions expected from introducing anticipatory actions to mitigate conditioning events.

### Anticipatory Thinking in MIDCA

We now put forth an anticipatory thinking approach as a metacognitive process in MIDCA, highlighting the role of each phase of the metacognitive layer:

**Monitor:** Obtain observations of the cognitive level components, including the current plan  $\pi$  and the current goal  $g$ .

**Interpret (as composed of the following three steps):**

**Discrepancy Detection:** Flag the current plan  $\pi$  from the cognitive level Plan phase (see Figure 1, cognitive layer, left side) as potentially risky,  $risk\_level(\pi, HIGH)$ .

**Explanation / Diagnosis:** Assess the risks associated with the plan  $\pi$  using anticipatory thinking approaches, such as those described in Section using the notions of *prestrength*. The results of the analysis would be vulnerabilities  $V$  of the plan  $\pi$ .

**Goal Formulation:** Formulate the goal to transform  $\pi$  into a new plan  $\pi'$  with a safer risk level, while maintaining that the current goal of  $\pi$  is achieved. The new goal would then be  $\{risk\_level(\pi', LOW) \wedge achieves(\pi', g)\}$ .

**Evaluate:** Drop any meta goals if they have been achieved.

**Intend:** Commit to achieving the newly formulated meta goal  $\{risk\_level(\pi', LOW) \wedge achieves(\pi', g)\}$ .

**Plan:** Take current mental state containing  $risk\_level(\pi, HIGH)$  and vulnerabilities  $V$  and search for a set of new actions, meta\_plan  $m_p$ , consisting of add or delete edits from plan  $\pi$  in order to achieve  $\pi'$  such that  $\{risk\_level(\pi', LOW) \wedge achieves(\pi', g)\}$ .

**Control:** Carry out the sequence of plan edits in  $m_p$  resulting in a new  $\pi'$  such that  $risk\_level(\pi', LOW)$  and  $\pi' \models g$  where  $g$  is the original goal of  $\pi$ .

The primary effort occurs in the Plan phase which we speculate could be modeled as a search process such that nodes are plans and their associated risk\_levels and edges between nodes are anticipatory actions that are added to (or possibly removed from) the plan. The search process would terminate when a goal node is reached that meets a low risk\_level for the plan inside the node. This example through the metacognitive phases serves as one possible realization of AT in MIDCA. We leave more concrete implementation details for future work.

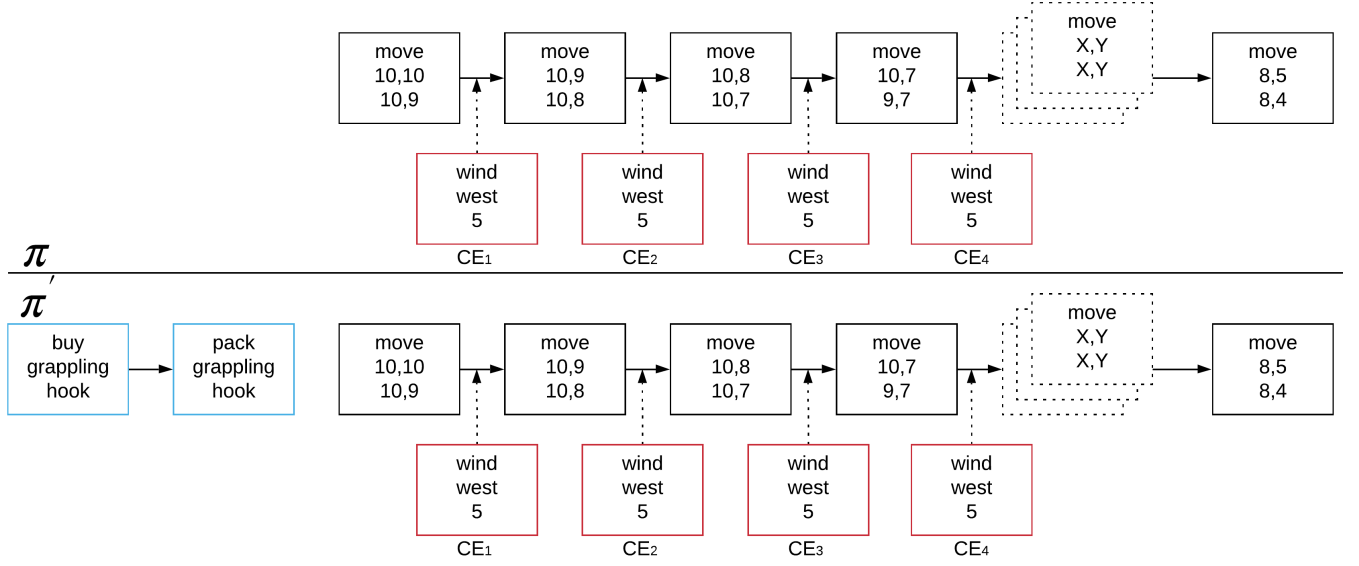


Figure 4: The original plan ( $\pi$ ) on the top with conditioning events and modified plan ( $\pi'$ ) with two anticipatory actions (blue) on the bottom

## Evaluation Framework

Anticipatory thinking is concerned with identifying possible worlds that affect desirable outcomes and taking action to mitigate them. This differs from typical future-oriented analysis centered around prediction that is focused on identifying a single likely outcome. As such, to appropriately evaluate anticipatory thinking we require alternative measures than those used in prediction.

### Successful Anticipatory Thinking

Conceptually, anticipatory thinking's goal is to have a high recall rate. More specifically, it is to ensure that the events that ultimately occur are accounted for in a set of possible futures. However, calculating recall does not capture the cost of adding anticipatory actions or the cost of identifying con-

ditioning events. To address this limitation we develop an assessment of anticipatory thinking that accounts for the cost in relation to the potential benefits.

An additional challenge is to avoid coupling anticipated outcomes to the actual outcomes. AT mitigates, rather than predicts, failure. Therefore AT assessment should only assess the potential payoff from mitigating, not whether any individual future comes to pass.

In Figure 5, we represent anticipatory thinking as a plan's identified conditioning events in the green circle. Before anticipatory thinking, conditioning events are unknown to our agent and reside in the blue area. Successful anticipatory thinking is the set of identified conditioning events where anticipatory actions are taken to mitigate their impact and reside in the yellow area. We assess successful anticipatory thinking as

$$AT^{assess}(\pi) = \frac{|CE(ANT)|}{|CE|} \times \left( 1 - \frac{|A^{ANT}(ANT)|}{\sum_i \left( \frac{|CE(ANT_i)|}{\sum_j impact(CE_j)} \right)} \right), \quad (1)$$

where  $|CE|$  is the number of conditioning events identified and  $|CE(ANT)|$  are the mitigated conditioning events. Their ratio represents how many conditioning events were mitigated. A second ratio calculates the potential benefit of mitigation. In the numerator, we calculate the cost of all anticipatory actions with  $A^{ANT}(ANT)$  where we assume an action cost to one. For each anticipatory action set, ANT, we sum the impact of each mitigated conditioning event mitigated,  $impact(CE_j)$ .

### Example

Applying equation 1 to our NBeacons conditioning events, we have four wind conditioning events,  $|CE| = 4$ , and each one is mitigated,  $|CE(ANT)| = 4$ . This ratio of 1.0 (4/4) is best possible case in that every identified conditioning event was mitigated. Conversely, plans where many identified conditioning events that have few mitigations would have a ratio closer to zero and may benefit from the use of robust search algorithms. Our next ratio assesses the potential mitigation



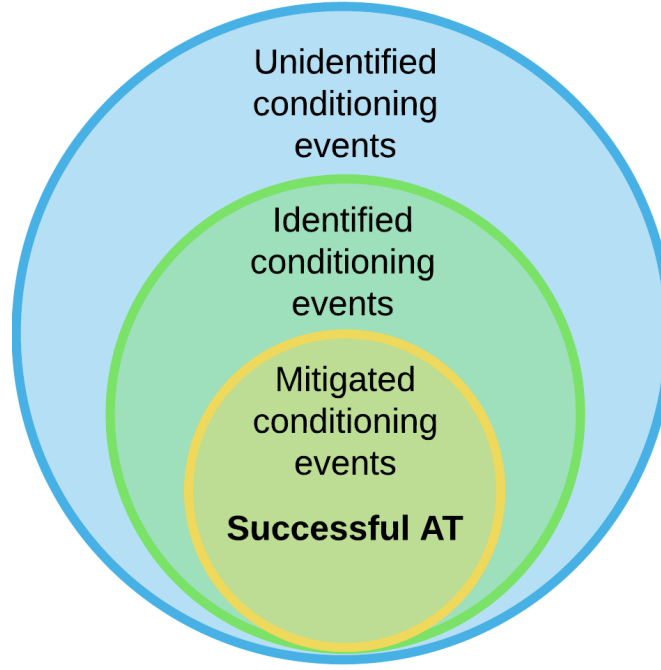


Figure 5: Successful anticipatory thinking identifies conditioning events that can be mitigated with anticipatory actions. In our *NBeacons* example, before the agent has done any anticipatory thinking, the *wind* conditioning events are *unidentified* and are in the blue area. After the the goal vulnerability and failure anticipation steps (Sections , ), *wind* conditioning events are now in the green area. Once anticipatory actions have been added to the plan (Section ), the *wind* conditioning events are mitigated and in the yellow area.

benefit. Mitigating the wind event requires buying and packing the grappling hook, each with an action cost of one for a total of two,  $|AA(ANT_i)| = 2$ . The sole anticipatory action sequence,  $i = 1$ , is expected to save the agent three dig actions for each of the four wind events,  $j = 4$ , resulting in a potential mitigation of twelve actions, and a resulting mitigation ratio of 0.83 ( $1 - 0.17$ ). Again, plans with not so favorable benefits from mitigations would have a lower expected payoff from their actions and would have a ratio closer to zero. Together these two ratios result in an  $AT^{assess}(\pi')$  of 0.83 ( $1.0 \times -0.83$ ), this calculation is reflected in Equation 2.

$$AT^{assess}(\pi') = \frac{4}{4} \times \left( 1 - \frac{2}{\sum_i (3 + 3 + 3 + 3)} \right), \quad (2)$$

## Conclusion and Future Work

Anticipatory thinking is a complex cognitive process for assessing and managing risk in many contexts. It allows humans to identify potential future issues and proactively take actions in the present that will manage their risks. We have defined how an artificial agent may perform anticipatory thinking at a goal reasoning level, so they may receive the same benefits and enable further autonomous capability.

Our approach made three contributions. First we defined anticipatory thinking in the MIDCA cognitive architecture

as a goal reasoning process at the metacognitive layer. Specifically, Section highlights the role of AT in each phase of the metacognitive layer of MIDCA shown in Figure 2. Second, we operationalized the anticipatory thinking concept as a three step process for managing risk in plans. Goal vulnerabilities, failure anticipations and failure mitigation identify weakness of a plan, their potential failure sources (conditioning events), and failure mitigations (anticipatory actions) to reduce the impact of the failure sources. Finally, we proposed a numeric assessment for successful anticipatory thinking. Key to the assessment are a ratio of identified conditioning events to mitigated ones and an expected cost-benefit ratio for the anticipatory actions.

We expect two immediate areas of future work. First, we are planning to integrate our anticipatory thinking definitions into an existing MIDCA implementation. From there, we will be able to perform experiments on existing domains. A second area is to develop more methodologies for each of the three anticipatory thinking steps. Expanding the failure sources beyond the failure inducing step (e.g. an action sequence) to identify the most parsimonious mitigation and extracting some benefit from unmitigated conditioning events are promising avenues of investigation.

## References

- Burns, E., and Ruml, J. B. W. 2012. Anticipatory On-Line Planning. *ICAPS* 333–337.
- Coddington, A.; Fox, M.; Gough, J.; Long, D.; and Serina,

- I. 2005. MADbot: A motivated and goal directed robot. In *Proceedings of the National Conference on Artificial Intelligence*, volume 20, 1680. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.
- Cox, M. T.; Dannenhauer, D.; and Kondrakunta, S. 2017. Goal operations for cognitive systems. In *Thirty-First AAAI Conference on Artificial Intelligence*.
- Cox, M.; Raja, A.; et al. 2011. Metareasoning: an introduction. *Metareasoning: Thinking about thinking* 3–14.
- Cox, M. T. 2016. MIDCA: A Metacognitive, Integrated Dual-Cycle Architecture for Self-Regulated Autonomy. *Aaai* 3712–3718.
- Dannenhauer, Z., and Cox, M. T. 2018. Rationale-based perceptual monitors. *AI Communications* 31(2):197–212.
- Dannenhauer, D.; Cox, M. T.; and Muñoz-Avila, H. 2018. Declarative Metacognitive Expectations for High-Level Cognition. *Advances in Cognitive Systems* 6:231–250.
- Dannenhauer, Z. A. 2019. *Anticipation in Dynamic Environments: Deciding What to Monitor*. Ph.D. Dissertation, Wright State University.
- Duff, S.; Harland, J.; and Thangarajah, J. 2006. On proactivity and maintenance goals. *AAMAS* 1033.
- Geden, M.; Smith, A.; Campbell, J.; Amos-Binks, A.; Mott, B.; Feng, J.; and Lester, J. 2018. Towards Adaptive Support for Anticipatory Thinking. In *Proceedings of the Technology, Mind, and Society*, TechMindSociety '18, 11:1—11:1. New York, NY, USA: ACM.
- Guilford, J. P. 1967. Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior* 1(1):3–14.
- Heuer, R. J. 2008. Taxonomy of Structured Analytic Techniques. *International Studies Association*.
- Hoffmann, J., and Brafman, R. 2005. Contingent planning via heuristic forward search with implicit belief states. In *Proc. ICAPS*, volume 2005.
- Huang, X.; Hong, S.; Hofmann, A.; and Williams, B. C. 2019. Online risk-bounded motion planning for autonomous vehicles in dynamic environments. In *Proceedings of the 29th International Conference on Automated Planning and Scheduling (ICAPS 2019)*.
- Mahadevan, S. 2018. Imagination machines: a new challenge for artificial intelligence. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- Muñoz-Avila, H.; Wilson, M. A.; and Aha, D. W. 2015. Guiding the ass with goal motivation weights. In *Goal Reasoning: Papers from the ACS Workshop*, 133–145.
- Muñoz, H.; Muñoz-Avila, M.; Dannenhauer, D.; and Reifsnnyder, N. 2019. Is Everything Going According to Plan?—Expectations in Goal Reasoning Agents. In *AAAI*.
- Penberthy, J., and Weld, D. 1992. UCPOP: A Sound, Complete, Partial Order Planner for ADL. In *International Conference on Knowledge Representation and Reasoning*.
- Pozanco, A.; Fernández, S.; and Borrajo, D. 2018. Learning-driven goal generation. *AI Communications* 31(2):137–150.
- Silvia, P. J.; Winterstein, B. P.; Willse, J. T.; Barona, C. M.; Cram, J. T.; Hess, K. I.; Martinez, J. L.; and Richard, C. A. 2008. Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts* 2(2):68.
- Szpunar, K. K.; Spreng, R. N.; and 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.
- Veloso, M. M.; Pollack, M. E.; and Cox, M. T. 1998. Rationale-based monitoring for planning in dynamic environments. In *AIPS*, 171–180.