Anticipatory Thinking Challenges in Open Worlds: Risk Management

Adam Amos-Binks, [1] **Dustin Dannenhauer,** [2] **Leilani H. Gilpin** [3]
[1] Applied Research Associates, Raleigh, NC 27616
[2] Parallax Advanced Research, Beavercreek, OH 45431
[3] University of California, Santa Cruz, Santa Cruz, CA 95060
aamosbinks@ara.com, dustin.dannenhauer@parallaxresearch.org, lgilpin@ucsc.edu

Introduction

Anticipatory thinking (Geden et al. 2019) drives our ability to manage risk – identification and mitigation – in everyday life, from bringing an umbrella when it might rain to buying car insurance. As AI systems become part of everyday life, they too have begun to manage risk. Autonomous vehicles log millions of miles (Kalra and Paddock 2016), StarCraft and Go agents have similar capabilities to humans, implicitly managing risks presented by their opponents. To further increase performance in these tasks, out-of-distribution evaluation has emerged as a way to characterize a model's bias, what we view as a type of risk management.

However, learning to identify and mitigate low-frequency, high-impact risks is at odds with the observational bias required to train machine learning models. StarCraft and Go are closed-world domains whose risks are known and mitigations well documented, ideal for learning through repetition. Adversarial filtering datasets provide difficult examples but are laborious to curate and static (Hendrycks et al. 2021), both barriers to real-world risk management. Adversarial robustness focuses on model poisoning under the assumption there is an adversary with malicious intent, without considering naturally occurring adversarial examples. Adversarial generation focuses at the object level (e.g. Zhao, Dua, and Singh 2018) without an open-world context where adversarial scenes occur. These methods are all important steps towards improving risk management but they do so without considering open-worlds, where new risks and new mitigations require dynamic risk management.

We unify these open-world risk management challenges with two contributions. The first is our perception challenges, designed for agents with imperfect perceptions of their environment whose consequences have a high impact (e.g. autonomous vehicles and public safety). Our second contribution are cognition challenges, designed for agents that must dynamically adjust their risk exposure as they identify new risks and learn new mitigations. Our goal with these challenges is to spur research into solutions that assess and improve the anticipatory thinking required by AI agents to manage risk in open-worlds and ultimately the real-world.

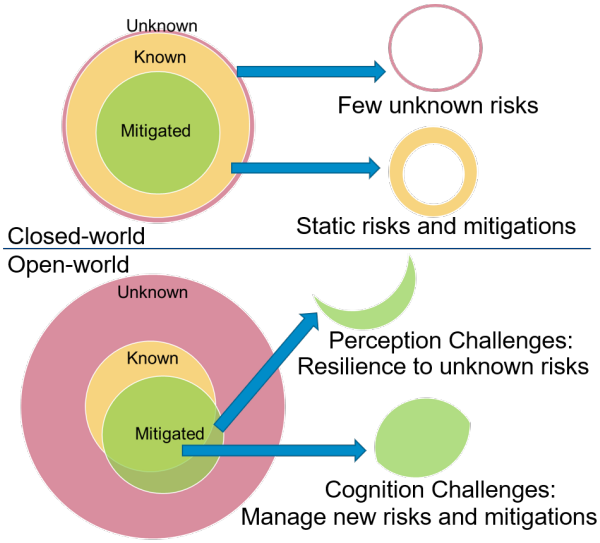

Figure 1: Managing risk in open-worlds is more complex than in closed-worlds. Agents must be resilient to unknown risks and once observed, new risks need to be balanced with existing risks by applying new mitigations.

Open World Risk Management

We view current AI systems like insurance agents that manage risks in a closed-world. Within an insurance domain (e.g. home) there are few unknown risks to insure against and the known risks are static (e.g. flood, fire, etc), see top Figure 1. Agents calculate mitigations (premiums) from an ample supply of historical likelihood and impact data. An individual manages their risk by purchasing policy premiums depends on their risk tolerance and budget. In contrast, the open-world perception challenge captures how well a perception system adapts to the disproportional (and growing) number of unknown risks in relation to known risks, see bottom Figure 1. For example, perception systems in autonomous vehicles have caused accidents by mis-classifying concrete barriers as the horizon. This epitomizes the often referred 'long-tail' of errors, the infrequent but highly impactful situations that open-world AI systems must manage the risk of. Solutions to the perception challenge will provide accurate assessments of a model's ability to identify and

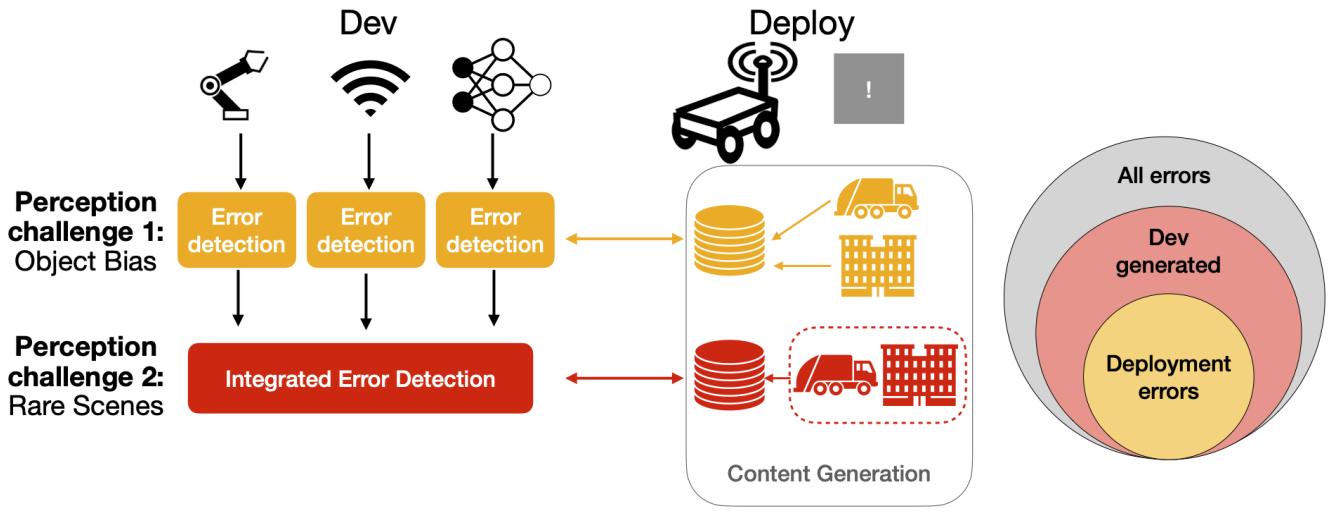

Figure 2: Flow diagram of the Perception challenges. Perception challenge 1: Object bias adds a content generation step at the perception system level. The inputs to the content generation step are generated from deployment (real, rare objects that were previously incorrectly labelled). Perception challenge 2: Rare Scenes adds a content generation step at the system level. The inputs to this content generation step are generated in development; inspired by "real" rare scenes; e.g., a truck carrying traffic lights.

manage unknown risks both at the object and scene level. Our cognition challenge assumes high quality perceptions (e.g. identifying new known risks) and focuses on balancing newly identified risks and learning new mitigations to maintain a desired risk profile. Open-world games (those with large, complex action models and catastrophic consequences such as permanent death) are ideal domains as long-range credit assignment is difficult to model and new adversaries and capabilities are constantly being discovered. The following sections detail these challenges in concrete domains and propose example high-level solutions that combine both symbolic reasoning and machine learning to operationalize risk management in open-worlds.

Perception Challenges

Autonomous vehicles are constantly perceiving their surrounding environments with perception systems, e.g., vision, LiDAR, and radar, to resolve the difference between reality and their representation of it. The vision system, which is opaque to humans, is not prepared for rare but highly impactful perception errors; making autonomous vehicles an open-world domain (Langley 2020).

One of the challenges of quantifying the risk of autonomous vehicles in the real world is that they are deployed in an open-world but tested in closed-world simulations. This juxtaposition highlights how unprepared autonomous vehicles are for naturally occurring adversarial examples. In this challenge, we suggest an alternative: generating naturally occurring adversarial examples from "real" rare cases to supplement the closed-world test simulations, mimicking the open-worlds where they are deployed. We further refine the perception challenge at the object and scene level, summarizing them in Figure 2.

Perception Challenge 1: Object Bias

The current state-of-the-art vision systems are easily fooled (Nguyen, Yosinski, and Clune 2015) by out-of-distribution inputs that exploit a model's observational bias. Some solutions include training on adversarial examples (Ilyas et al. 2019), but they only lead to improvements on this data set. Instead, we propose iterative testing where objects mislabelled in deployment are used as input to a content generation model for testing and evaluation of a perception system's observational biases. The generative model creates sets of naturally adversarial objects that can be tested during development (e.g. faded stop signs, shadow patterns).

Perception Challenge 2: Rare Scenes

A second type of imperfect perception occurs when perceiving rare scenes. While there is still a potentially problematic observational bias at the object level, we assume that the object detection system is mostly accurate. A vision system can still be "fooled" by rare scenes such as traffic lights on a truck[1]. In these cases, the perception is correct: the objects are correctly identified, but the scene is so rare and disruptive that it should be immediately recognized so the system can avoid catastrophic failures (e.g. immediately stopping on a freeway). These "naturally-occurring adversarial scenes" exemplify the problem with the long tail of errors in autonomous vehicles. An additional layer of reasoning is required to test and monitor for these types of rare scenes.

To address rare scene limitations, we propose a similar iterative process: a generative model that constructs out-of-distribution scenes with commonly occurring objects. The

[1]https://futurism.com/the-byte/tesla-autopilot-bamboozled-truck-traffic-lights

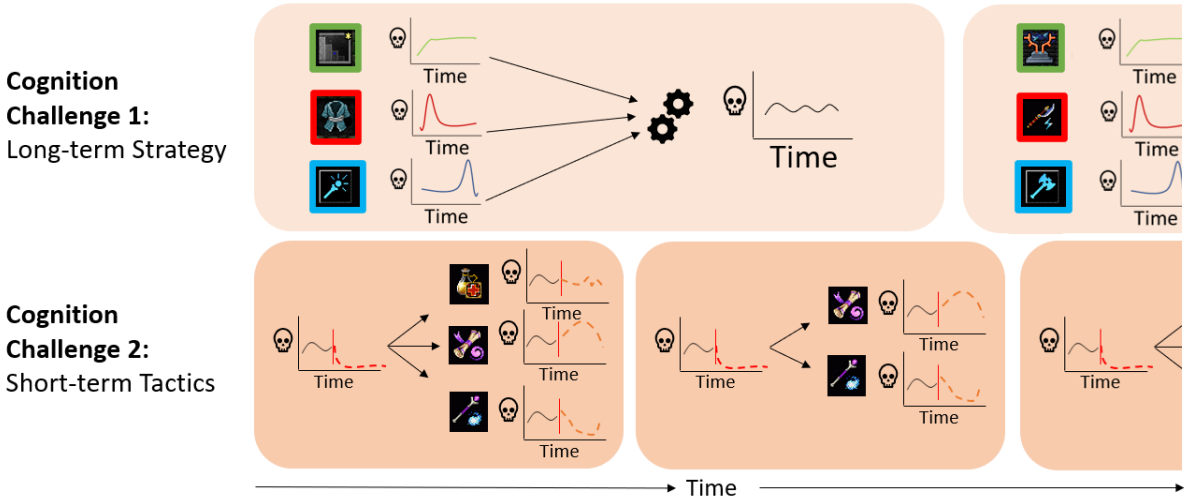

Figure 3: We operationalized cognition challenges in the Dungeon Crawl Stone Soup environment that include long-term strategy (top) and short-term tactics (bottom) risk management tasks.

generative model uses objects that are regularly seen in the data (e.g., road signs, objects, etc.) and changes their orientation (e.g., fallen signs on the ground) or environment (e.g., objects in the same scene that usually are not together).

Example Solution

Autonomy requires adapting to operational circumstances that cannot be enumerated pre-hoc. However, assessing this ability is static and contrived, relying on benchmark tasks (e.g. avoid collisions in a benchmark scenario) that systems could be gaming with non-generalizable rules. Methods such as variational auto-encoders (VAEs) and generative adversarial networks (GANs) can generate a whole family of tasks and push the boundaries of contrived task-oriented evaluations to rich attribute-oriented evaluations. We envision integrating physical world knowledge into content generation methods to ensure real-world properties are preserved and natural adversarial instances are generated. For example, static traffic indicators (stop signs, traffic lights, etc) that are now dynamic (e.g. loaded onto a flatbed truck, advertisement painted on a large truck). Performance against a family of out-of-distribution tasks would indicate a system's competency in a risk management and provide a level of insight into future performance not currently available.

Impact

Autonomous vehicles in new, open environments are susceptible to causing harm (even death). Despite high accuracy in testing, object recognition systems cannot account for rare but consequential out-of-distribution inputs. Testing and improving a perceptions system's ability to manage the risk presented by these new circumstances would make autonomous systems safer and cover a larger number of error and failure cases. This type of anticipatory, introspective, solution impacts public safety, manufacturers, and regulators.

Cognition Challenges

Our cognition challenges for anticipatory thinking in open worlds involve AI agents acting in unexplored, hostile, and dynamic environments. The first challenge: *long-term strategy* concerns issues of risk as agents move between situations and micro environments where resource availability changes. The second challenge: *short-term tactics* concerns more immediate risks that require agents to make use of the right resources at the right time and choice of enemy engagements. We use the rogue-like video game Dungeon Crawl Stone Soup (DCSS) to operationalize this challenge problem. In DCSS a player moves through a procedurally generated, partially observable, and stochastic environment to retrieve the 'Orb of Zot' while managing the risks (permanent death) of encountering thousands of monsters. DCSS is one of two rogue-like games with increased interest in recent years and remains an unsolved domain for AI (Dannenhauer et al. 2021; Küttler et al. 2020).

We consider DCSS an open-world for both human and AI agents owing to over 600 unique monsters and 100 item types that a player comes across as they delve deeper into the 100 levels of the dungeon. As the player enters new areas of the game, the chances of encountering each type of monster and item change, such that certain items and monsters cannot be found until reaching a certain depth. Therefore as a player who begins after only playing the short tutorial, there are many surprises and unknowns encountered. For this reason, DCSS can be considered an open-world for any learning agent who does not start with a complete model of the environment (such a model likely only exists in the minds of the lead developers). For use in AI research, *dcss-ai-wrapper* (Dannenhauer et al. 2021) provides an API for both symbolic and vector AI agents to interact with DCSS.

Cognition Challenge 1: Long-term Strategy

To win a game of DCSS, AI agents must mitigate risk over longer time horizons as they visit different regions of a world. Humans perform this type of reasoning frequently, such as going from home to work, going from the grocery store to a camping trip, or going from the airport lobby through security to their gate. In all cases, agents are able to take actions in the source region that may no longer be available in the destination region. Such actions could include whether to pack an umbrella and lunch box, or taking a water filter. These same challenges are present in the DCSS domain and require decision making to decide where in the dungeon to go next and choices in enhancing character capabilities. The effects of these decisions are often noticed only after significant time delays as shown in the top portion of Figure 3. We now highlight two aspects of the long-term strategy challenge in DCSS:

Capability Enhancement via Skill Point Allocation: The player earns experience when killing monsters which is permanently allocated to one or more of 33 skills. These skills impact the success of a variety of player-actions and defensive capabilities. Failure in the mid to late game often results from poor risk management of skill points.

Region Transitions via Level Choices: There are 24 themed branches (regions) of the dungeon, each with their own special characteristics. Some branches are known for an overwhelming majority of poison-based monsters (spider lair and snake pit), corrosive monsters (slime pit), monsters that will cause unhelpful mutations to your character (the abyss), or monsters with special abilities, such as locking stairs preventing escape (the vaults). Further increasing complexity, monsters move in levels independent of your movement, it is likely that as you explore a level, monsters will move in between your planned escape route from the location at which you entered the level. Entering a branch without dynamically adjusting risks and mitigations (poison resist gear, corrosion resist gear, potions of cure mutation, etc.) will likely lead to catastrophic failure.

Cognition Challenge 2: Short-term Tactics

The limited visibility and highly dynamic, stochastic nature of item and monster interactions means that there are many possible futures over short horizons consisting of a small (< 5) possible actions. Additionally, agents may never encounter uncommon situations more than once because of the procedurally generated nature of the game. Therefore short-term anticipatory thinking should seek to mitigate the absolute worst outcomes even if unable to accurately predict a single outcome. These types of decisions are shown in the bottom portion of Figure 3. We now highlight two aspects of the short-term tactics challenge in DCSS:

Engagements via Direct Combat: It is impossible to progress into the mid to late game by continuously engaging in combat with little thought to tactical reasoning. Retreating, hiding from monsters, using special abilities, and using other terrain features such as hallways, are needed to gain advantages (i.e. in a hallway you can fight monsters in smaller batches).

Resource Management via Item Inventory: The player will encounter tens of thousands of item instances from 100 item types and the player has a limited inventory space. Sometimes items may need to be left and returned to later, other items dropped to make sure the most important life saving items are kept for dire situations. Additionally, the choice to use or save a life-saving consumable item is often the difference between success and failure. These types of decisions must be made many times during gameplay.

Example Solutions

Simply learning a complete action model of an open-world is intractable and so is encoding all knowledge an agent needs for managing risks. A hybrid or 3rd wave approach would reason over the semantics of what new capabilities or assets would produce a desired risk profile (e.g. goal reasoning, meta-cognition), taking into account the small samples of newly observed risks (e.g. meta-learning, few-shot learning). This approach is just an illustrative example of the complexities required to go beyond single task performance and manage risk in open-worlds.

Impact

Developing AI agents that can solve these two challenges in DCSS will lead to approaches that can lead to better AT in a wide variety of autonomous systems. An autonomy approach that can solve the right choice of entering which branch of the dungeon could be the same approach that helps an autonomous drone decide whether it can follow a target into a broken down urban building, a forest, or a cave. An AI that is tasked with it's mission and can solve the pre-deployment problem could alert it's operating before it leaves the forward operating base that it's missing a needed component to mitigate a possible risk from adversaries that the operator may have forgotten. An AI agent that can manage risks during operation will be able to make the call that it cannot outrun a new enemy and must seek to hide instead.

Conclusion

Real-world deployments of AI systems require a level of robustness and safety best tested in open-worlds. We use a risk management approach to contribute two challenges and illustrate them in different domains. Our perception challenges characterize a systems ability to recognizing both rare objects and configurations of them. Solutions with this capability are essential as the real-world is constantly producing new permutations of objects and contexts they appear in. Our cognition challenges characterize the relationship between new risks and new capabilities to mitigate them. Dynamically balancing a risk profile with new information enables AI agents to adapt to situations like humans do and avoid needless catastrophic failures. Our example solutions focus on integrating symbolic and learning approaches together to manage risks in perception and cognition.

References

Dannenhauer, D.; Dannenhauer, Z. A.; Decker, J.; Amos-Binks, A.; Floyd, M. W.; and and, D. W. A. 2021. dcss-ai-wrapper: An API for Dungeon Crawl Stone Soup providing both Vector and Symbolic State Representations. In Planning and Reinforcement Learning Workshop. International Conference on Automated Planning and Scheduling (ICAPS).

Geden, M.; Smith, A.; Campbell, J.; Spain, R.; Amos-Binks, A.; Mott, B.; Feng, J.; and Lester, J. 2019. Construction and Validation of an Anticipatory Thinking Assessment. Frontiers in Psychology 10(December): 1–10. ISSN 16641078. doi:10.3389/fpsyg.2019.02749.

Hendrycks, D.; Zhao, K.; Basart, S.; Steinhardt, J.; and Song, D. 2021. Natural Adversarial Examples. In 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 15257–15266. doi:10.1109/CVPR46437.2021.01501.

Ilyas, A.; Santurkar, S.; Tsipras, D.; Engstrom, L.; Tran, B.; and Madry, A. 2019. Adversarial examples are not bugs, they are features. arXiv preprint arXiv:1905.02175 .

Kalra, N.; and Paddock, S. M. 2016. Driving to Safety: How Many Miles of Driving Would It Take to Demonstrate Autonomous Vehicle Reliability? Santa Monica, CA: RAND Corporation. doi:10.7249/RR1478.

Küttler, H.; Nardelli, N.; Miller, A. H.; Raileanu, R.; Selvatici, M.; Grefenstette, E.; and Rocktäschel, T. 2020. The NetHack Learning Environment. In Proceedings of the Conference on Neural Information Processing Systems (NeurIPS).

Langley, P. 2020. Open-World Learning for Radically Autonomous Agents. Proceedings of the AAAI Conference on Artificial Intelligence 34(09): 13539–13543. doi:10.1609/aaai.v34i09.7078. URL https://ojs.aaai.org/index.php/AAAI/article/view/7078.

Nguyen, A.; Yosinski, J.; and Clune, J. 2015. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In Proceedings of the IEEE conference on computer vision and pattern recognition, 427–436.

Zhao, Z.; Dua, D.; and Singh, S. 2018. Generating natural adversarial examples. 6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings (2016): 1–15.