Rebel Agents That Adapt to Goal Expectation Failures

Zahiduddin Mohammad[1], Michael T. Cox[2] and Matthew Molineaux[2]
[1]Wright State University, Dayton, OH 45435
[2] Wright State Research Institute, Beavercreek, OH 45431
mohammad.48@wright.edu, {michael.cox, matthew.molineaux}@wright-research.org

Abstract

Humans and autonomous agents often have differing knowledge about the world, the goals they pursue, and the actions they perform. Given these differences, an autonomous agent should be capable of rebelling against a goal when its completion would violate that agent's preferences and motivations. Prior work on agent rebellion has examined agents that can reject actions leading to harmful consequences. Here we elaborate on a specific justification for rebellion in terms of violated goal expectations. Further, the need for rebellion is not always known in advance. So to rebel correctly and justifiably in response to unforeseen circumstances, an autonomous agent must be able to learn the reasons behind violations of its expectations. This paper provides a novel framework for rebellion within a metacognitive architecture using goal monitoring and model learning, and it includes experimental results showing the efficacy of such rebellion.

Introduction

In recent years there has been increased interest in autonomous agents capable of rebellion when given unreasonable goals or tasks (Aha & Coman, 2017; Briggs and Scheutz 2016; Coman and Aha, 2018). Rebel agents are those that can reject assigned goals and plans or oppose behaviors and attitudes of other agents including humans.

An agent may need to rebel if:

1. It has critical information which humans or other agents do not have.
2. It has been given a goal which may not be achieved due to lack of resources.

In item (1), consider an agent that is helping a human cutting down the trees in a forest. While the human is busy cutting down trees, the agent sees an animal which could cause a potential harm to the human. The agent rebels by stopping and informing the human about the danger, even though its current goal is to help with logging. In item (2), consider an agent that is given a goal to reach a destination with insufficient fuel. Given no additional resources, it should refuse.

Boggs, Dannenhauer, Floyd and Aha (2018) put forth one approach to handling rebellion. Here, the flow of rebellion is in such a way that given world W and goals G, the agent interprets and then evaluates each goal. If a goal cannot be achieved, it is removed immediately, while if a goal can be achieved, it plans and acts accordingly. If a goal is achievable but undesirable, the agent may choose to rebel. A goal is undesirable if the action to achieve it results in a state of lower utility. Then given an undesirable goal, the agent makes a probabilistic choice to rebel based on its inherent tendency towards rebellion. If the choice is to rebel, the agent informs a human about its rebellion and seeks permission. The human has a choice to either accept or reject the rebellion (again this is a probabilistic choice). If the human rejects the rebellion, the agent probabilistically chooses either to reject or comply with the human's advice. The work here seeks to develop an alternative basis for rebellion not grounded in simple utility or probabilistic choice.

The basis for rebellion and the characterization of a goal as unacceptable in the work here is grounded in the concept of a goal expectation. Once achieved, an agent expects the goal state to continue to hold true unless another agent or the agent itself does something to make it unsatisfied. When these expectations fail, there is reason to question other goals whose executed plan fragments can violate the expectation. So, if an agent has achieved two goals, but the action to achieve a third would undo them both, we are justified in rejecting the third goal (all else being equal).

We present a rebel agent implemented using MIDCA, an open source architecture that provides a modular structure and an explicit focus on both a cognitive layer and metacognitive layer. In this work, the rebel agents are MIDCA agents. Humans and agents often have their own specific knowledge of the world, the goals they pursue, and the actions they perform. When there is a goal expectation conflict in the agent's understanding of the world, then the agent should rebel.

The remainder of the paper describes our rebel agent and how rebellion can improve performance (maximizing goal achievement). First, we discuss about the MIDCA architecture followed by the framework for learning in rebellion. Next, we discuss about the plant protection domain followed by how rebellion is implemented in MIDCA. Then, we report our experimental setup and results. Finally, we provide a brief concluding statement.

Metacognitive Integrated Dual-Cycle Architecture

The *Metacognitive Integrated Dual-Cycle Architecture (MIDCA)* (Cox et al., 2016) is a cognitive architecture that models both cognition and metacognition for intelligent agents. It consists of "action-perception" cycles at both the cognitive level and the metacognitive level. In general, a cycle performs problem-solving to achieve its goals and tries to comprehend the resulting actions and those of other agents. Problem solving contains intention, plan, and action execution phases, whereas comprehension consists of perception, interpretation, and goal evaluation (see Figure 1).

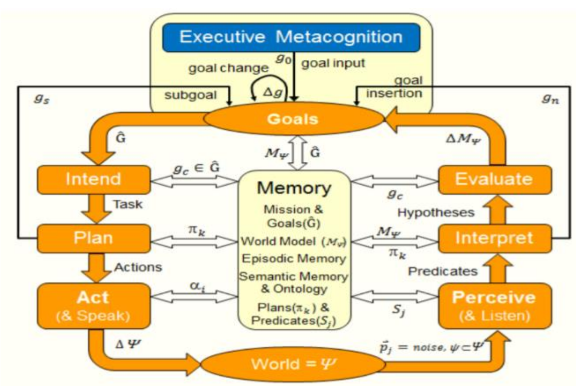


Figure 1. Cognitive layer of MIDCA (adapted from Dannenhauer & Cox (2018)

Interpret handles a number of processes including transforming raw data into an internal state and identifying goal discrepancies. Evaluate is responsible for tracking progress made on the agent's current goals, including dropping goals that have been achieved or have failed. In Figure 1, Intend decides the current goals that the agent should be pursuing, Plan generates the sequence of actions needed to reach the agent's goals, and finally Act executes the behaviors needed to carry out the planned actions. A cycle similar to the cognitive layer occurs at the meta cognitive layer, except instead of perceiving and acting in an environment, perception occurs on the agent's cognitive layer and action is taken to change the cognitive layer or internal memory instead of changes in the world environment. The rebellion related processes in this work occur at the cognitive layer in the Interpret and Evaluate phases.

A Preliminary Framework for Rebellion

This work revolves around the notion of a *goal expectation*. Like an informed expectation (Dannenhauer and Munoz-Avila 2015), a goal expectation leads an agent to believe that once a goal is achieved, it will remain achieved unless the agent does something to change the goal state (at least for relatively short intervals of time). Note that, as a boundary case, if a goal is already satisfied in the initial state, then it will also remain achieved. When an agent has such an expectation and it observes that the goal becomes unsatisfied, a goal discrepancy occurs.

After the first goal expectation discrepancy is detected, the agent interprets a goal among the remaining goals i.e., infers whether achieving this particular goal can lead to another goal expectation violation and then evaluates the goal. If the goal is undesirable (leading to another goal expectation violation), the agent then informs the human about the rebellion and then removes the goal from the goal graph (i.e., the agent's goal agenda). The agent then continues with the remaining goals accordingly by planning and acting. If the goal is desirable, then the agent plans and acts accordingly.

Here we propose a general framework for rebellion with learning (see Figure 2). After the first goal discrepancy is detected, when it enters the interpret phase of cognitive layer it then checks for any knowledge discrepancy and learns the correct operator if there is any. It then changes its knowledge accordingly and continues with the evaluate phase. If the agent does not detect any knowledge discrepancy, it continues with evaluate phase. As shown within the dashed boxes of Figure 2, these are the main points our approach differes from that of Boggs and colleagues.

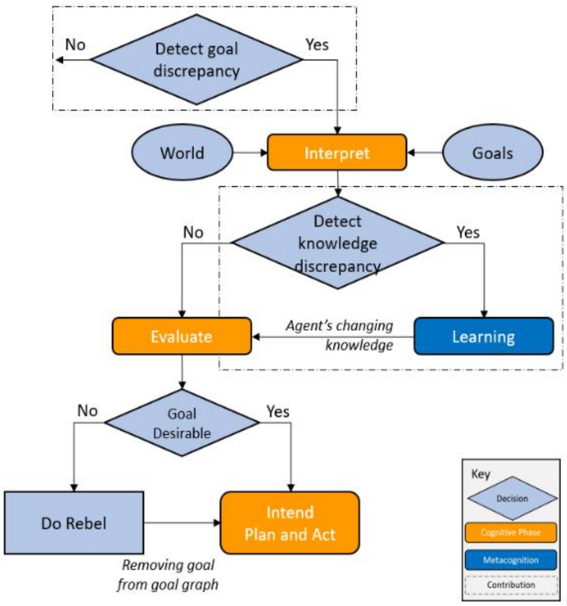


Figure 2. Flow of agent's rebellion with learning

The Plant Protection Domain

The plant protection domain (Boggs & Dannenhauer 2018) is a grid world that includes harmful invasive plants, endangered native plants, human operators, and robotic gardening agents. The robots are responsible for navigating to target locations and deploying herbicides. Deployment of herbicides will kill plants at an agent's location and neighboring locations. The world in which agents act is a map grid of tiles where a plant occupies a single tile and tiles containing plants cannot be moved through by agents. Plants are static fixtures which cannot be moved. Once dead, they stay dead (i.e., there is no re-plant action). The motivation for selecting this domain is that an agent executing a plan when pursuing a goal, e.g., not(invasive-at(loc1)), can accidently undo some of its other goals, e.g., native-at(loc2).

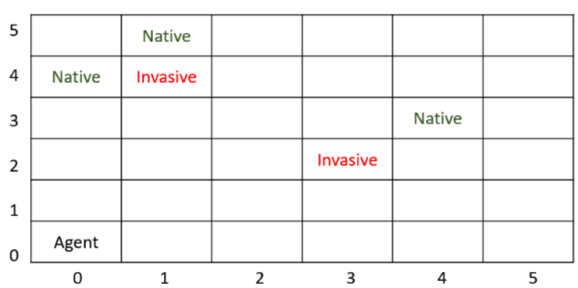


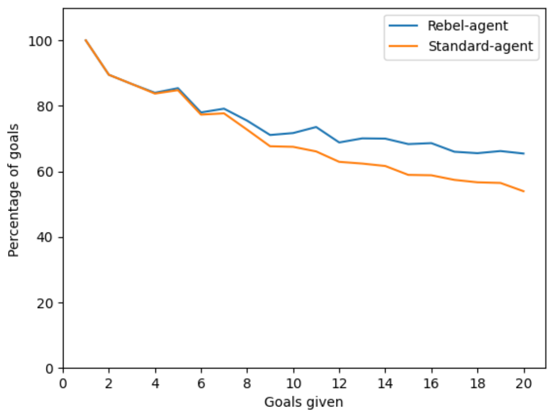
Figure 3. Representation of an instance of plant protection domain

In Figure 3, the agent possesses a set of goals to preserve all native plants and remove all invasive plants. An agent traverses to location (3,2) to remove the invasive plant and deploys a herbicide which kills the invasive plant and also a native plant at (4,3). This records a goal expectation violation because it undoes a goal that was already true (in the initial state). Later it moves to location (1,4) and infers that deploying a herbicide is undesirable as the agent might remove native plants at (0,4) and (1,5). Therefore, it rebels by refusing to perform the spray action that would have achieved the removal of the invasive plant (i.e., the goal). By doing so, the rebel agent achieves three out of the five goals; whereas, a traditional agent achieves only two of the five goals, because it kills all native plants.

Experimental Setup and Results

The experiments we have conducted so far provide a baseline for understanding how rebellion effects the outcome of missions within the Plant Protection Domain. The results obtained from these experiments are preliminary and learning was not conducted in this experiment. However, in future we will intend to incorporate learning. Unlike Boggs, Dannenhauer, Floyd and Aha (2018) where they consider plants remaining, we do take percentage of goals achieved as a performance metric.

Our tests were run in a 6×6 world with a rebellious agent and a standard agent. Scenarios were created by randomly placing invasive and native plants in the map grid. Throughout these scenarios the agent possesses a set of goals to remove all the invasive plants and to preserve all native plants. We systematically varied the number of goals to be accomplished from 1 to 20, for a total of 2000 scenarios, 100 scenarios run at each number of goals. We compared performance of a rebel MIDCA agent that operates with a domain model to an equivalent agent that does not rebel. The rebel agent operates as described above. Results, shown in Figure 4, clearly demonstrate that the rebellious agent outperforms the standard agent.



Related and Future Research

The Belief-Desire-Intention (BDI) community has done research work addressing the same issues as the work presented here. Although in a different vocabulary and formalism, both communities worry about conflicting, changing, and inconsistent goals (Harland, Morley, Thangarajah, & Yorke-Smith, 2017). For example, Khan and Lespérance (2010) have developed an approach for agents to reason about whether or not a goal should be performed based on priority, preferences and consistency between goals. Goals may be refused by postponing until they become appropriate at a later time, or they may be dropped when inconsistent with other goals.

The concept of rebel agent as discussed herein was first introduced by Coman, Gillespie & Munoz-Avila (2015). Coman and Aha (2018) provide a thorough review of the idea, and as already mentioned, Boggs, Dannenhauer, Floyd & Aha (2018) represents the work from which this paper has drawn the most influence. Here we also use MIDCA as a development platform and their plant protection domain is reimplemented here. The differences are reflected most by our use of goal expectations as the basis for classifying a goal as unacceptable, and our preliminary and planned integration of learning.

In the current work, we say that agent is avoiding second goal expectation discrepancy by rebelling against a specific goal which can lead to it. Instead the agent can come up with a different plan to achieve that specific goal by avoiding second goal discrepancy to occur. For this to occur the agent should learn if there is a knowledge discrepancy i.e., when the expected state is different from the current state. Let us assume we have an agent which has an incorrect action model (i.e., a planning operator). In this scenario, the agent should be capable of learning about its actions over a period. So, the next question which arises is when should the agent start learning? It is whenever an agent observes a knowledge discrepancy. Here we assume that there are no external factors influencing the effects of the action.

In the current version of MIDCA for every phase in the cognitive layer, a metacognitive cycle is being run (see figure 5). A knowledge discrepancy is detected in the interpret phase of the cognitive cycle. As we have a complete metacognitive cycle, which starts with the monitor phase where it obtains the most recent trace of cognition from memory (Dannenhauer, Cox & Munoz-Avila 2018). The cognitive trace is constructed by recording what each module in each phase takes as an input and produces as an output. Specifically, the

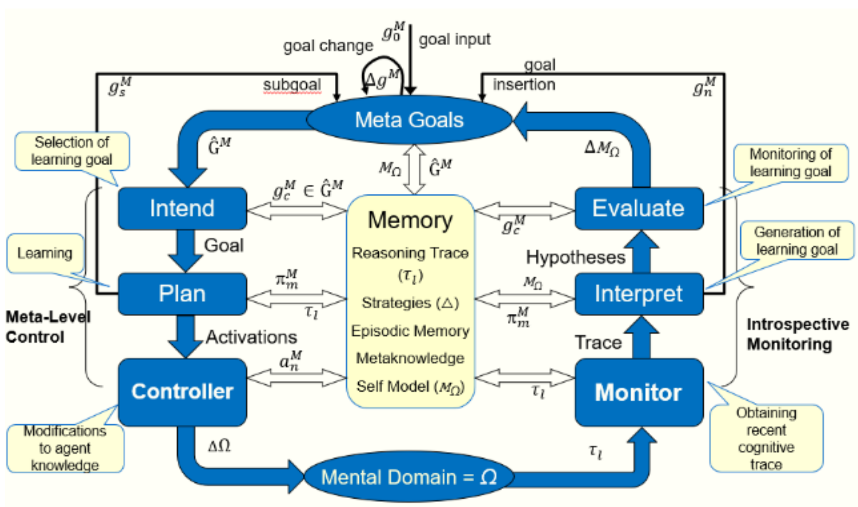


Figure 5. Metacognitive layer of MIDCA (adapted from Dannenhauer, Cox, & Munoz-Avila, 2018)

cognitive trace is composed of an ordered dictionary that can be indexed via the phase and cycle. Here from the cognitive trace, set of background knowledge about the problem, a knowledge base that describes training examples and a knowledge base that provides labels for training examples are recorded. The Interpret phase of the metacognitive layer is responsible for detecting the cognitive level behavior. In this phase, a learning goal (Cox and Ram, 1999) will be generated which will be passed to the next phases of the meta-layer. The evaluate phase of the metacognitive layer is responsible for monitoring the metacognitive goals. The intend phase will select the goals which are supposed to be achieved. Current implementation of intend selects all the goals which are pending.

The planning or learning phase runs on cognitive traces. In the learning phase, the goal which is given from the intend phase will be achieved. The learning modification of the agent's knowledge about the operator is carried out in the control phase. In the control phase, the agent finds if there is discrepancy between the rule generated by the learning algorithm with the agent's knowledge of a target operator, then adds to the operator conditional effects corresponding to rules inferred by learning algorithm. After the agent's memory is updated with the rule in the control phase the metacognitive cycle ends.

The cognitive cycle then continues with the cognitive evaluate phase, which checks whether given goals are completed or not. If the agent's knowledge of operators is improved, the overall performance of the agent will also improve due to increased accuracy of its predictions, which will allow it to rebel appropriately.

Conclusion

This paper examined the influence of agent rebelliousness when goals have both positive and negative effects relative to goal achievement. Our results demonstrated that given a scenario where to achieve a goal it must undo a few other goals, rebel agents outperform normal agent to maximize the number of goals achieved. As a result, after the first goal discrepancy is detected, invasive plants which are not close to native plants are killed. Whereas the normal agent will kill the invasive plant irrespective of native plants being adjacent to them which lead to the decrease in the overall percentage of goals achieved.

Acknowledgments

This work is supported by AFOSR grant FA2386-17-1-4063, by ONR grant N00014-18-1-2009 and by NSF grant 1849131.

References

Aha, D. W.; and Coman, A. The AI Rebellion: Changing the Narrative. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 4826–4830, 2017.

Boggs, J.; Dannenhauer, D.; Floyd, M. W.; Aha, D. W. 2018. The Ideal Rebellion: Maximizing Task Performance in Rebel Agents. In *Proceeding of the 6th Goal Reasoning Workshop.* Stockholm, Sweden. ICJAI 2018.

Briggs, G; and Scheutz, M. 2016. The Case for Robot Disobedience. *Scientific American*, 316(1): 44–47.

Coman, A.; and Aha, D. W. 2018. AI Rebel Agents. *AI Magazine,* 39(3): 16-26. doi.org/10.1609/aimag.v39i3.2762

Coman, A.; Gillespie, K.; and Muñoz-Avila, H. 2015. Case-Based Local and Global Percept Processing for Rebel Agents. In *Workshop Proceedings from the 23rd International Conference on Case-Based Reasoning*, edited by J. Kendall-Morwick, 23–32. The CEUR Workshop 1520. Aachen, Germany: RWTH-Aachen University.

Cox, M. T.; Alavi, Z.; Dannenhauer, D.; Eyorokon, V.; Munoz-Avila, H.; and Perlis, D. 2016. MIDCA: A Metacognitive, Integrated Dual Cycle Architecture for Self-regulated Autonomy. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 3712–3718. AAAI Press.

Cox, M. T., & Ram, A. (1999). Introspective multistrategy learning: On the construction of learning strategies. *Artificial Intelligence, 112*, 1-55.

Dannenhauer, D., Cox, M. T., & Munoz-Avila, H. (2018). Declarative metacognitive expectations for high-level cognition. *Advances in Cognitive Systems, 6*, 231-250.

Dannenhauer, D; Floyd, M. W.; Magazzeni, D. and Aha, D. W. 2018. Explaining Rebellion Behavior in Goal Reasoning Agents. In *Proceeding of ICAPS-18 Workshop on Explainable Planning.*

Dannenhauer, D., & Munoz-Avila, H. (2015). Raising expectations in GDA agents acting in dynamic environments. *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence* (pp. 2241–2247). Buenos Aires, Argentina.

Dannenhauer, Z. A., & Cox, M. T. (2018). Rationale-based perceptual monitors. *AI Communications, 31*(2), 197-212.

Harland, J.; Morley, D.N.; Thangarajah, J.; and Yorke-Smith, N. Aborting, suspending, and resuming goals and plans in BDI agents. *Autonomous Agents and Multi-Agent Systems* 31**,** 288–331, 2017.

S.M. Khan and Y. Lespérance. A Logical Framework for Prioritized Goal Change. In *Proc. AAMAS'10*, pp. 283–290, 2010.