Understanding Novelty in Reinforcement Learning-Based Automated Scenario Generation

Jonathan Rowe, Andy Smith, Randall Spain, and James Lester

North Carolina State University

INTRODUCTION

Simulations will serve a critical role in the next generation of training. A key feature of simulation-based training is the capacity to deliver scenarios that support the acquisition, practice, and assessment of domain-critical knowledge, skills, and abilities. The recently formed Close Combat Task Force, ordered by the U.S. Secretary of Defense, has called for Soldiers to take part in 25 simulated battles before facing their first real contact (Judson, 2018). To attain this level of training, devising effective methods for the creation and delivery of simulation-based training scenarios is essential. However, creating training scenarios for simulation-based environments poses significant challenges: authoring simulation-based training scenarios is often resource-intensive; it requires specialized knowledge about specific simulators and authoring tools; scenarios often support only limited reuse; and most scenarios adhere to a one-size-fits-all approach that does not support adaptivity, either with respect to the changing needs of instructors or changing needs of trainees. These issues point toward the need for automated scenario generation to increase the availability and diversity of training scenarios across a range of tasks, domains, and simulation environments.

A key criterion in evaluating the effectiveness of automated scenario generation systems is their capacity to create training scenarios that are novel. Novel training scenarios are (a) meaningfully different from previously experienced scenarios, and (b) aligned with relevant instructional objectives for training. For example, it is possible to generate a vast range of variations on a training scenario by subtly adjusting the location of a single entity in a simulated environment. However, these variations could hardly be considered novel. Similarly, generating training scenarios that are misaligned with relevant instructional objectives, or unrealistic with respect to real-world scenarios, is of little value as well. Instead, novel scenarios should differ from existing scenarios in ways that are pedagogically meaningful—for example, they modify a scenario’s difficulty or alter the format of a correct solution—in order to provide learners with new, beneficial training opportunities.

In this paper, we explore the role of novelty in data-driven automated scenario generation for simulation-based training environments. This is informed by our ongoing work to develop DEEPGEN, a reinforcement learning (RL) framework for automated scenario generation in the domain of Call for Fire (CFF) training with Virtual Battlespace 3. DEEPGEN utilizes RL techniques to induce computational models for run-time tailoring of scenarios to achieve instructor-specified training objectives (Rowe, Smith, Pokorny, Mott, & Lester, 2018). Specifically, we formalize scenario generation as an RL task that involves sequential decision making about enacting adaptations (i.e., actions) to an exemplar scenario, observing learner interactions with the generated scenario (i.e., trajectories), and using the resultant learner performance data (i.e., reward) to refine an internal decision-making model for future scenario adaptation decisions.

To guide our discussion of novelty, we draw upon the psychology literature on creativity. Kaufman and Beghetto (2009) devised the Four C Model of Creativity, which distinguishes between *Big-C*, *Pro-c*, *Little-c* and *mini-c* conceptualizations of creativity. The Big-C category refers to eminent creativity, which is understood as creative work that is historically significant, lasting, and signifying creative genius. Pro-c creativity refers to effortful progression toward, and often antecedent to, Big-C status; it is associated with professional-level expertise that exceeds novice-level creativity, but does not yet reach a level of Big-C contribution. Little-c creativity is everyday focused, referring to expressions of creativity that are performed by non-experts, such as inventive problem solving or creative endeavors undertaken as a hobby. Mini-c creativity is a signature of the learning process; it is defined as the “novel and personally meaningful interpretation of experiences, actions, and events” (Kaufman & Beghetto, 2009).

Drawing upon the Four C Model, we conceptualize novelty in automated scenario generation in terms of four categories. First, we distinguish training scenarios that are wholly new and unique, and valuable; this is akin to anticipating future scenarios that have never been encountered before. Second, we distinguish scenarios that are new to an instructor (or simulation environment), but are not necessarily novel in a universal sense. Generating scenarios in this category is a significant enhancement to the training capacity of a simulation-based training environment. Third, we distinguish training scenarios that are new to a particular learner, or group of learners, even if they were already pre-existing. Fourth, we distinguish novel experiences that a learner might have with a scenario that they have previously experienced; successfully completing a scenario for the first time would fall into this category.

Utilizing this conceptual model, we analyze a prototype version of the DEEPGEN scenario generation system in terms of the Scenario Adaptation Library that was developed for CFF training, as well as the scenarios that can be automatically generated by the prototype system. We discuss novelty in terms of several dimensions of CFF scenario adaptation, including adjustments to low-level features such as unit types, terrain, weather, and entity locations, as well as higher-level features, such as adversary behaviors, mission objectives, and training contexts. We explore how novelty can be operationalized within an RL framework for automated scenario generation. Finally, we discuss the implications for the Generalized Intelligent Framework for Tutoring (GIFT) as they relate to the integration of automated scenario generation capabilities within adaptive training systems. Developing theoretically grounded approaches to conceptualizing novelty in automated training scenario generation will be critical to meeting the mandate of next-generation simulation-based training.

RELATED WORK

We approach automated generation of training scenarios from the perspective of a related research area: interactive narrative generation (Riedl & Bulitko, 2012). Interactive narrative generation focuses on the design of computational models for dynamically generating and tailoring digital interactive experiences in which users drive an unfolding storyline through their own actions and decisions. A range of computational techniques have been investigated for interactive narrative generation, including classical AI planning (Young et al., 2013), adversarial search (Nelson & Mateas, 2005), case-based reasoning (Fairclough, 2004), and machine learning (Wang et al., 2018). Grounded in this work, we conceptualize scenarios in terms that are analogous to interactive narrative systems: scenarios consist of sequences of events that unfold within a virtual environment. A scenario specification includes the initial state of the virtual world, including its terrain, agents, buildings, weather, and overall task instructions presented to the user. In addition, the set of agent behaviors and associated triggers that define how events play out within the virtual environment are integral to the scenario. In a training context, scenarios specify a set of learning objectives to be addressed as well as assessment criteria. Finally, scenarios are completable, and they should be both coherent and internally consistent.

Recent years have seen growing interest in the use of machine learning techniques for data-driven automated scenario generation in education and training. This includes applications of dynamic decision networks (Mott & Lester, 2006), dynamic Bayesian networks (Lee, Rowe, Mott, & Lester, 2014), and reinforcement learning techniques (Rowe & Lester, 2015; Wang et al., 2018). However, much of this work has focused on devising computational models for tailoring narrative-centered learning experiences to ensure they are effective and engaging; there has been comparatively little work investigating novelty in machine learning-based frameworks for automated scenario generation. Although novelty has been investigated for scenario generation with evolutionary algorithms (Folsom-Kovarik & Brawner, 2018), to date, novelty has not been a central factor in machine learning-based approaches.

DEEPGEN RL-BASED SCENARIO GENERATION FRAMEWORK

To contextualize our discussion of novelty in automated scenario generation, we cite examples from an ongoing project in our lab that investigates the design and development of a data-driven scenario generation system, DEEPGEN. DEEPGEN formulates automated scenario generation as an instance of data-driven interactive narrative generation utilizing deep reinforcement learning techniques.

To serve as an initial testbed, DEEPGEN focuses on automated scenario generation in the task domain of artillery call-for-fire (CFF) training. Broadly speaking, a CFF mission consists of an infantry soldier requesting indirect fire on a target from supporting artillery (e.g. field artillery, unmanned aircraft). The requesting soldier, or forward observer (FO), follows a defined communication protocol to identify himself, describe the mission type, describe the target and location, and describe the method of engagement. The mission continues as the FO may choose to adjust fire as necessary based on the results of initial shots, and conclude the mission by relaying a battle damage assessment once the target has been hit. Given this general structure, there are a broad range of scenario adaptations that can be enacted to augment the difficulty of a CFF training scenario, such as changing the type of mission, modifying enemy behaviors, modifying weather and time-of-day, changing the types and locations of targets, and varying the types of equipment in the FO’s loadout.

**Table 1. Partial Scenario Adaption Library for CFF domain**

|  |  |
| --- | --- |
| **Adaptable Elements of Scenario** | **Scenario Adaptation Variants** |
| Target Type | * Infantry squad
* Transport vehicle
* Tank (T72)
 |
| Target Behavior | * Stationary
* Patrol
* Move to waypoint
 |
| Target Reaction To Fire | * No reaction
* Stop movement
* Flee to cover
* Return to base
 |

As dynamic scenario adaptation involves enacting a series of decisions about how to orchestrate training events at run-time, DEEPGEN enumerates the full range of possible adaptations in a *Scenario Adaptation Library.* The Scenario Adaptation Library defines the space of possible transformations that can be applied to example (i.e., parent) scenarios in order to produce new (i.e., child) scenarios. Thus, DEEPGEN approaches novelty through the systematic combination and application of individual scenario adaptations. In the CFF training testbed, we have defined 16 possible dimensions for scenario adaptation corresponding to more than 1,000,000 possible scenario variations that could be generated from a single example scenario. A sample of adaptations from DEEPGEN’s Scenario Adaptation Library can be found in Table 1.



Figure 1. Screenshots of prototype DEEPGEN instructor tool.

To generate scenarios using DEEPGEN, an instructor first selects a set of learning objectives and selection criteria to guide the scenario generation process, as shown in Figure 1(a). Using these inputs, DEEPGEN systematically generates a set of scenario adaptations that can be enacted with a given parent scenario, ranking these combined variations according to their match to the selected criteria, as shown Figure 1(b). Finally, the instructor can view a more detailed view of each scenario in the form of a Warning Order, as shown in Figure 1(c), before selecting one or more scenarios to be downloaded and realized within a 3D simulation-based training environment. To serve as a testbed simulation environment for CFF training, DEEPGEN interoperates with Virtual Battlespace 3 (VBS3). Developed by Bohemia Interactive Simulations, VBS3 is a 3D simulation platform that is widely used by the U.S. Army for a range of training purposes, including IED training, land navigation, route clearance, and convoy training. Furthermore, we utilize the VBS3Fires plug-in, a third-party tool created by SimCentric Technologies that provides a GUI interface and ballistics simulation engine for training CFF in VBS3. Scenario specifications generated by DEEPGEN are realised in VBS3 by modifying example VBS missions through a semi-automated compilation process that produces a full set of executable and configuration files required by VBS.

In the current prototype version of DEEPGEN, the Scenario Adaptation Library for CFF training has been hand-authored through close collaboration with U.S. Army subject-matter experts, thus guaranteeing that generated scenarios are completable and coherent. However, there are significant overlaps between many of the generated scenarios. This raises important questions related to scenario novelty: How should we understand the degree of novelty that is supported by DEEPGEN’s automated scenario generation framework? To what extent can we enhance the degree of novelty exhibited in scenarios created by DEEPGEN? And how can novelty be conceptualized to advance training objectives in task domains like CFF training?

NOVELTY IN DEEPGEN

To frame our discussion of novelty in scenario generation, we use the Four C Model of Creativity devised by Kaufman and Beghetto (2009). The Four C model extends traditional creativity research, which has historically focused mainly on “genius”-level creativity (Big-C) and “everyday” creativity (Little-c) to include two new levels, Pro-c and Mini-c. In this section we expand on each of the Four C’s, discussing how they relate to novelty in scenario generation and providing specific examples of how they are, and are not, addressed by DEEPGEN.

Big-C creativity refers to “clear-cut, eminent creative contributions” (Kaufman & Beghetto, 2009). Depending on the domain, this can refer to creative works ranging from award-winning musical compositions to scientific discoveries to Pulitzer Prize-winning novels. This category exists to distinguish exceptional creative people and works from highly competent, even professional level creators. From a scenario generation perspective, we identify Big-C level novelty as referring to scenarios that introduce fundamental, long-lasting changes to the rules or performance expectations associated with a target domain. For example, consider the case of AlphaGo, an artificial intelligence-based game-playing agent designed by Google DeepMind to automatically learn how to play the board game Go (Silver et al., 2017). AlphaGo famously performed at a level capable of beating top-rated human players, and on the 37th move of Game Two in its historic match against 18-time world champion Lee Sedol, AlphaGo performed a move that was not only effective, but that commentators later regarded as “creative,” “beautiful,” and likely to be incorporated into future matches by human players (Metz, 2016). It is important to note that the system did not just beat expert-level human players, but it did so by exhibiting a novel strategy that had not been previously encountered. Another example of Big-C novelty in scenario generation comes from the Millennium Challenge 2002 wargaming exercise (Borger, 2002). During the exercise, the commander of the “Red” force adopted a variety of novel strategies that effectively defeated the “Blue” forces despite their superior technological advantages. The strategies were so effective that the simulation was reset, and the rules of engagement were rewritten because the “Blue” force had been defeated so quickly. Although created by a human author, the Millennium Challenge example demonstrates a Big-C level contribution from a scenario due to its impact on the strategies employed in future war-gaming exercises.

For a system to produce Big-C level novelty in scenario generation, a variety of factors are required. By definition, these scenarios make a contribution that did not previously exist, and thus the scenario generator must have access to a broad range of flexibility and freedom to explore the space of possible

scenarios. This calls for direct access to robust, high-fidelity simulation environments, or game-like environments where agents can be trained through “self-play” or other strategies likely to produce emergent behavior, which contrasts with approaches designed to mimic expert human performance. Notably, it can be difficult to recognize innovative scenarios if they are not highly effective in comparison to competing options (i.e., innovative scenarios might not be selected by an imperfect optimization process). Thus, even given ideal conditions, it is not guaranteed that any system will generate scenarios at the Big-C level of novelty.

The next level of novelty we consider is Pro-c*,* or professional-level creative expertise. This level refers to creativity exhibited by individuals who have earned professional-level status in a discipline, but may not yet have transformed their field or made an eminent contribution. For scenario generation, Pro-c novelty is associated with scenario generation that produces scenarios at a level of quality, complexity, distinctiveness and unexpectedness as to offer significant value to a domain expert (i.e., instructor) possessing deep experience in the subject domain.

Pro-c level novelty is a target level for DEEPGEN, because of its promise to offer value to both instructors and advanced trainees alike. It calls for scenario generation functionalities that can operate upon virtually all aspects of a scenario within a given domain. In CFF training, this corresponds to modifying a CFF scenario’s pre-mission briefing; augmenting the types, locations and behavior of friendly and adversary forces; inserting dynamic events at run-time (e.g., weather changes, communication failures); and altering the embedded scaffolding and assessment rules at play in scenarios. Notably, Pro-c level novelty need not exclusively pertain to complex orchestration of world states and triggered events. It is possible for two scenarios with identical units, terrain, locations, and scripted events to be distinguished from one another by augmenting the mission, or related context, faced by a trainee. When a prospective target enters an area, should the forward observer call for indirect fire, or let the target pass? How does the target’s value compare to other possible targets? How much ammunition is available? What is the assigned objective for the forward observer? Questions such as these introduce meaningful forms of experiential novelty without requiring complex orchestration of events within a virtual simulation environment.

Given that Pro-c level systems are not necessarily targeted at the discovery of new military tactics or strategies, scenario generators like DEEPGEN can be designed to restrict the space of possible scenarios through deliberate authoring of the Scenario Adaptation Library, ensuring that all generated scenarios are both feasible and qualitatively different from one another. For example, several adaptable scenario elements listed in Table 1, such as changing how enemies react to being fired upon, are higher-level scenario adaptations that might typically be associated with the *Pro-c* level. This is in contrast to lower-level modifications, such as changes in weather or target type, that may be more likely to yield trivial variations between scenarios.

The next category, Little-c, refers to non-expert expressions of creativity, such as every-day problem solving or creative works. For scenario generation, we characterize this level of scenario generation in terms of creating “base-level” scenarios that enable trainees to reach basic proficiency in a domain. Little-c systems might only adjust a small number of features of a scenario, producing scenarios that are different by a small degree and do not require the level of domain knowledge or instructional expertise associated with professional-level scenario generation. In DEEPGEN, enacting scenario adaptations such as changing the weather, time of day, or type of target are all consistent with *Little-c* novelty. In some cases, this level of scenario generation may be preferable, because it is relatively inexpensive and efficient to setup and generate a large number of different, similarly structured scenarios that can be used for repeat practice of specific competencies.

The final level in the Four C model is mini-c. Mini-c describes creativity that is inherent to the learning process; it is expressed at an individual level while engaged in productive problem solving. This can be understood as a form of novelty that is experienced by novice learners as they begin to learn a new domain. For scenario generation, we define this level of novelty as consistent with generating “introductory” scenarios for a domain. These scenarios should be relatively simple from a complexity standpoint with novelty measured in relation to the concepts and competencies already achieved by a trainee.

Overall, the Four C model provides a useful framework for discussing and evaluating novelty in automated scenario generation systems. It provides a framework for formulating design requirements of these systems, and it points toward directions for evaluating the degree of novelty supported by scenario generators. Furthermore, it offers a useful ontology for describing how different characteristics of scenarios impact novelty within a given training domain.

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

Automated scenario generation will serve a key role in the future of simulation-based training because of its significant potential for reducing the cost of creating novel scenarios and expanding access to high-quality simulation-based training. Data-driven approaches to automated scenario generation hold promise for enhancing trainee learning experiences by leveraging recent advances in reinforcement learning and interactive narrative technologies. We have presented an overview of DEEPGEN, a data-driven automated scenario generation framework, which formalizes the task in terms of enacting sequential adaptations to a canonical “parent” scenario in order to generate “child” scenarios that can be evaluated to assess learning outcomes. We have described an initial Scenario Adaptation Library that was developed for the domain of Call for Fire training. To better define and evaluate the degree of novelty embedded in scenarios generated by DEEPGEN, we have adopted the Four C Model, discussing how the model fits within the context of automated scenario generation and CFF training specifically.

In future work, we plan to expand DEEPGEN’s Scenario Adaptation Library to capture a broader range of possible transformations to “parent” training scenarios, including sequential adaptations that can be performed dynamically over the course of a scenario. Further, it will be important to systematically investigate how instructors and learners interact with DEEPGEN, including the DEEPGEN Instructor Tool for configuring scenario generation, as well as generated scenarios for training a range of CFF skills. Finally, it will be critical to demonstrate how the DEEPGEN framework can be generalized to support additional domains. Integration with GIFT will be useful for enabling this line of investigation—automated scenario generation is particularly well-aligned with the Practice Quadrant in GIFT’s Engine for Management of Adaptive Pedagogy (EMAP)—setting the stage for expanding our understanding of how the Four C Model can be operationalized to measure and contextualize novelty in scenario generation across different domains.

ACKNOWLEDGMENTS

The research described herein has been sponsored by the U.S. Army Research Laboratory under cooperative agreement W911NF-18-2-0020. The statements and opinions expressed in this article do not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

REFERENCES

Borger, J. (2002). War game was fixed to ensure American victory, claims general. *The Guardian*, *21*. Fairclough, C. (2004). Story games and the OPIATE system. University of Dublin, Trinity College.

Folsom-Kovarik, J.T., & Brawner, K. (2018). Automating Variation in Training Content for Domain-general Pedagogical Tailoring. In *Proceedings of the Sixth Annual GIFT User Symposium* (pp. 75-86). Orlando, FL: U.S. Army Research Laboratory.

Judson, J. (2018, September). 25 bloodless battles: Synthetic training will help prepare for current and future operations. *Defense News*, Retrieved from https://[www.defensenews.com/smr/defense-news-conference/2018/09/05/25-bloodless-battles-synth](http://www.defensenews.com/smr/defense-news-conference/2018/09/05/25-bloodless-battles-synth) etic-training-will-help-prepare-for-current-and-future-operations/

Kaufman, J. C., & Beghetto, R. A. (2009). Beyond big and little: The four c model of creativity. *Review of General Psychology*, *13*(1), 1-12.

Lee, S., Rowe, J. P., Mott, B. W., & Lester, J. C. (2014). A Supervised Learning Framework for Modeling Director Agent Strategies in Educational Interactive Narrative. *IEEE Transactions on Computational Intelligence and AI in Games*, 6(2), 203-215.

Luo, L., Yin, H., Cai, W., Zhong, J., & Lees, M. (2017). Design and evaluation of a data-driven scenario generation framework for game-based training. *IEEE Transactions on Computational Intelligence and AI in Games*, *9*(3), 213-226.

Metz, C. (2016). In two moves, AlphaGo and Lee Sedol redefined the future. *WIRED. com*, *16*.

Mott, B. W., & Lester, J. C. (2006). U-Director: A Decision-Theoretic Narrative Planning Architecture for Storytelling Environments. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems* (pp. 977-984). Hakodate, Japan: ACM.

Nelson, M. J., & Mateas, M. (2005, June). Search-Based Drama Management in the Interactive Fiction Anchorhead. In Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (pp. 99-104), Menlo Park, CA: AAAI.

Riedl, M. O., & Bulitko, V. (2012). Interactive narrative: An intelligent systems approach. *AI Magazine*, 34(1), 67.

Rowe, J., & Lester, J. (2015). Improving Student Problem Solving in Narrative Centered Learning Environments: A Modular Reinforcement Learning Framework. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence in Education* (pp. 419-428). New York, NY: Springer.

Rowe, R., Smith, A., Pokorny, B., Mott, B., & Lester, J. (2018). Toward Automated Scenario Generation with Deep Reinforcement Learning in GIFT. In *Proceedings of the Sixth Annual GIFT User Symposium* (pp. 65-74). Orlando, FL: U.S. Army Research Laboratory.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... & Chen, Y. (2017).

Mastering the game of go without human knowledge. *Nature*, *550*(7676), 354-359.

Wang, P., Rowe, J. P., Min, W., Mott, B. W., & Lester, J. C. (2018) High-Fidelity Simulated Players for Interactive Narrative Planning. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence* (pp. 3884-3890). Stockholm, Sweden.

Young, R. M., Ware, S. G., Cassell, B. A., & Robertson, J. (2013). Plans and planning in narrative generation: a review of plan-based approaches to the generation of story, discourse and interactivity in narratives. *Sprache und Datenverarbeitung, Special Issue on Formal and Computational Models of Narrative*, *37*(1-2), 41-64.

ABOUT THE AUTHORS

**Dr. Jonathan Rowe** is a Research Scientist in the Center for Educational Informatics at North Carolina State University. He received the Ph.D. and M.S. degrees in Computer Science from North Carolina State University, and the B.S. degree in Computer Science from Lafayette College. His research is in the areas of artificial intelligence and human-computer interaction for advanced learning technologies, with an emphasis on game-based learning environments, intelligent tutoring systems, user modeling, educational data mining, and computational models of interactive narrative generation.

**Mr. Andy Smith** is a Research Scientist in the Center for Educational Informatics at North Carolina State University. He received his M.S. in Computer Science from North Carolina State University, and his B.S. degrees in Computer Science and Electrical and Computer Engineering from Duke University. Prior to graduate school Andy worked as an Underwater Robotics Engineer at SPAWAR SSC Pacific in San Diego, CA. His research is focused on the intersection of artificial intelligence and education, with emphasis on user modeling, game-based learning, and educational data mining.

**Dr. Randall Spain** is a Research Psychologist in the Center for Educational Informatics at North Carolina State University where he uses principles, theories, and methods of applied psychology (human factors, educational psychology, personnel psychology, experimental psychology, and psychometrics) to design and evaluate the impact of advanced training technologies on learning and performance. He has conducted training and human factors research for the Department of Defense and the Department of Homeland Security for the past 10 years with a focus on adaptive training, performance assessment and measurement, user modeling and human-automation interaction. Dr. Spain is a PhD graduate from Old Dominion University’s Human Factors Psychology program and serves on the editorial board for Military Psychology.

**Dr. James Lester** is Distinguished Professor of Computer Science at North Carolina State University, where he is Director of the Center for Educational Informatics. His research centers on transforming education with technology-rich learning environments. With a focus on adaptive learning technologies, his research spans intelligent tutoring systems, game-based learning environments, affective computing, and tutorial dialogue. The adaptive learning environments he and his colleagues develop have been used by thousands of students in K-12 classrooms. He received his Ph.D. in Computer Science from the University of Texas at Austin in 1994. He is a Fellow of the Association for the Advancement of Artificial Intelligence (AAAI).