Proceedings of the Twenty-Ninth International
Florida Artificial Intelligence Research Society Conference

Designing an Authorable Scenario Representation for Instructor Control over Computationally Tailored Narrative in Training

J.T. Folsom-Kovarik, Angela Woods, Robert E. Wray
Soar Technology, Inc.
jeremiah@soartech.com, angela.woods@soaretch.com, wray@soartech.com

Abstract

Training scenarios, games, and learning environments often use narrative to manipulate motivation, priming, decision context, or other aspects of effective training. Computational representations of scenario narrative are useful for computer planning and real-time tailoring of training content, but typically define how to display narrative in the scenario world. The training rationales and the impacts of narrative on trainees are not typically accessible to the computer. We describe a computational representation that lets instructors explicitly author the training goals and impacts in a narrative. The representation captures both causality in the simulated world and instructional intent. Furthermore the streamlined representation enables nontechnical authoring of sophisticated interactions between instructional goals when a computer tailors training material to individual learners. The narrative representation has the potential to increase instructor acceptance, understanding, and control over computer tailoring, thereby making training more effective.

Introduction

Training scenarios, games, and learning environments often include narrative elements such as persistent characters and internally consistent events. Compared to realistic but isolated simulation events that do not fit into a larger story, narrative can enhance the demonstration, practice, and evaluation of trainees' skills. For example narrative can put a trainee's actions in contexts that change the likely or correct course of action (Kray and Galinsky 2003), make learning more engaging and personally meaningful (Rowe et al. 2009), or tend to increase transfer of training to real world performance (Schumacher and Gentner 1988). Narrative can be *tailored* – adapted to support or challenge individual learners (Wray et al. 2009).

A common difficulty facing instructors is a mismatch between how instructors think about narrative for training and how narratives are represented computationally (Wray et al. 2015). In order to drive a simulation environment, computational representations often must focus on the mechanics of *how* to control characters and events in order to present a simulated world. However instructors must think about *why* to arrange events and interactions – the resulting training must achieve instructional goals.

Computational narrative representations in a persistent simulation world typically focus on how to animate characters and assemble events that play out a story. Such narrative representations typically capture the outcomes of authors' tailoring choices but not the reasoning. For example, a script with branches might let a computer first test learners and then choose different events based on their performance. But since the underlying reasons for the choices are not part of the computer's understanding, the tailoring that results cannot be extended to new settings or more complex interactions of many instructional goals.

In contrast, training rationales are of vital concern to instructors. When instructors *author* narrative – that is create or modify characters, events, or whole stories – they embed knowledge of why the narrative is present and how it is expected to affect training outcomes. Making this knowledge explicit in a computational representation would allow a computer to reason about it when controlling a training scenario or learning environment.

In the authors' experience, instructors often need to control scenarios so that training reflects their expert knowledge, new content in an evolving state of the art, or the needs of different trainee groups. Instructors often report that computer choices are hard to understand or change. But abstraction from events to training rationales typically is lacking in narrative representations for computer systems. Thus, it may be that some difficulty instructors report stems not from lack of technical expertise but from a mismatch in desired abstractions to reason about.

At the same time, giving authors new abstractions but requiring them to manipulate a general planning language and describe instructional goals and impacts in computer terms undeniably introduces technical barriers to use.

Therefore, a narrative representation ideally should enable a nontechnical user to express exactly those kinds of knowledge instructors are likely to require. Limited expressivity in the representation is hypothesized to *increase* the sophistication of instructional knowledge that instructors may author (Folsom-Kovarik, Wray and Hamel 2013).

With information about training goals and impacts integrated in a streamlined computational scenario representation, we hope to make effective training scenario narrative more accessible for instructors to author and to increase the acceptance, usability, and efficacy of the resulting training.

This paper identifies requirements for an instructor-Authorable Scenario Representation that enables tailored narrative in training. We describe an implemented design of such a representation, *ASR*, and observations and lessons from implementing it and using it to create and tailor narrative in a training scenario.

Related Work

Our work builds on the rich history of generic or specialized planning algorithms that make narrative choices such as selecting actions for characters or changing a story in response to user actions. For example, the Automated Story Director (Riedl and Stern 2006) controlled events in a cultural training scenario so that a central planner devised repairs when the trainee inputs would tend to move away from the desired scenario storyline. This system used STRIPS and later PDDL planning frameworks (Riedl 2009) to represent author's intent for the story as preconditions and postconditions of character actions. By contrast, ASR uses a much simpler representation of event effects on the simulation world but adds instructional impact.

Shareable Content Object Resource Model (SCORM) and its successors (e.g., Durlach and Johnson 2014) define semantics for expressing instructors' knowledge about learning experiences so that computers can select and sequence those experiences. These are typically created at the level of whole modules, whereas we seek to apply knowledge to individual events and details within a module, enabling narrative tailoring.

The Cultural Meeting Trainer (Wray et al. 2009) carried out similar tailoring of a learning experience. The present work builds on it by adding an explicit representation of narrative progression and adding richer attributes on each available event, so that narrative tailoring can change details of events based on more pedagogical factors (simplicity, predictability) and practical factors (plausibility).

Tailor (Carpentier, Lourdeaux and Thouvenin 2013) offers a capability to select between narrative events based on models including the events possible in the current simulation state and their impact on learning. However, the instructor control over events only encompasses selecting pre-built events and training objectives. Instructors need technical personnel to author or change tailoring choices.

Scribe (Medler and Magerko 2006) aimed to expose scenario narrative to instructor authors. Scribe presented a powerful, programmer-like control over events via a friendly user interface. ASR is much simplified in comparison because experience suggests nontechnical users, given this kind of low-level control over system internals, may make incorrect changes or even feel discouraged from using the system (e.g., because they do not understand it).

In focusing on the tension between expressivity of a scenario representation and its usability or authorability, ASR attempts to expose specifically those kinds of narrative control that instructors who present the training are likely to require. Thus, the representation is likely to align with instructors' understanding of training and is expressive enough to describe diverse and responsive scenarios.

Requirements for an Authorable Scenario

ASR is presently grounded in a military simulator for training cognitive and perceptual skills (Schatz et al. 2012). The requirements the simulator places on scenario design will also apply to a range of simulation training settings that share its broad characteristics.

As one example simulation, a trainee might play a small-unit leader arriving near a foreign village to provide security and engineering support. Over the course of days in the simulation, computer-controlled characters in the village display behaviors and interactions. Trainees must attempt to understand and interpret these scenario events, reporting the events they observe that are relevant to mission goals (Sieck, Smith and Rasmussen 2013). For example, interspersed with other events several different villagers might visit and talk respectfully with a somewhat well-dressed gentleman – the trainees should infer over time that he is a local leader. The overarching mission of deciding what support the village might need and which characters to work with provides narrative focus not present if trainees were simply assigned to interpret such scenes in isolation.

ASR allows an author to capture the process of deciding how to select options within the simulation narrative. This is in contrast with prior work, which merely encoded the *outcome* of tailoring decisions in the narrative. For example, information about training goals and rationales is missing if an instructional expert scripts a scenario so that: *at a certain point, if the trainee does not shift attention for two minutes, then trigger an argument near the leader's house.* The underlying reasoning – that fixed attention implies distraction or boredom and these conditions may be countered with an attention-grabbing event – is not represented.

Instructionally relevant information that ASR represents can include many facts that an instructor is likely to reason about. ASR describes skills the simulation should teach and the current estimated trainee proficiency in such skills, as well as trainee knowledge of story events.

Finally, ASR must represent the predicted impact of each narrative event on any of these instructional facts to enable planning scenario changes or tailoring. With the anticipated impact on learning encoded in the scenario representation, the system is able to carry out scenario tailoring that might involve complex tradeoffs in a manner similar to a human instructor. For example, the need for an attention-grabbing event could generalize beyond the one situation where it was directly specified in the previous example. Or, the system could determine that based on estimated trainee needs at a certain time, it is currently less appropriate to distract the trainee and better to wait further.

Authorable Scenario Representation (ASR)

In order to meet the requirements of instructor authoring and alignment with training requirements, we created an authorable scenario representation (ASR) intended to both reflect and guide instructors' thought processes in defining training scenarios. As argued above, we have deliberately limited the expressiveness of this representation to help make it authorable for the instructors who will use it.

ASR encodes a string of pearls representation (Majewski 2003). Pearls on the string describe possible scenario variations, available choices for both the computer system and the trainee. However, there is a single path or string of gateways through a training scenario and at each one all variations must converge.

This structure both allows and limits scenario variation. Each variation starts from the same point and ends with the same invariants as every other path in the same pearl. Each authored, alternative sequence of events forms a notional *arc* along the surface of the pearl (Figure 1). Arcs are the unit of planning. An arc describes the tailorable parameters of presenting event sequences and their different expected impacts on learning (see next section).

Arcs, or alternative paths through a scenario, in ASR are each described by their value in up to three scalar dimensions for each skill being trained: helpfulness, simplicity, and predictability (collectively, HSP). These three dimensions are designed to let instructors describe their broad instructional goals (Wray and Woods 2013). High *simplicity* makes it more straightforward for trainees to apply a skill (while reducing simplicity adds complexity, such as added steps or factors to consider). High *predictability* makes the outcomes of correct or incorrect performance more obvious to the trainee (or, to challenge with low predictability, makes outcomes more ambiguous or contradictory). High *helpfulness* describes any direct interventions such as instruction or reminders from characters. For example, under high helpfulness, a teammate might ask the trainee for a status report when the trainee forgets one. Utterances or actions with low helpfulness can challenge trainees by distracting them or adding to cognitive load.

The system does not differentiate details of how an arc is simple, helpful, etc. Labeling arcs with HSP impact on skills helps instructors move away from a strict encoding of if-this-then-that decisions in a training scenario, which is important in order to enable reasoning about many available paths and possibly conflicting instructional impacts.

The string of pearls strictly limits combinatorial explosion of narrative variations. Branches that change the story from its overall path are explicitly disallowed. However, in ASR, we elaborate the highly constrained string of pearls representation with two extensions: multiple parallel strings and floating events.

First, multiple strings represent independent, parallel narrative sequences that can be tailored individually. This parallelism enables ASR to express variation efficiently. For example, if the scenario should include two separate village leaders and each has four possible activities to display in their own homes, we can select those activities separately on parallel strings rather than explicitly describe all sixteen combinations. Multiple parallel strings let ASR efficiently describe background characters, events in two different locations, people who do not meet, etc.

Second, floating events do not fit on the paths of a pearl. They add flexibility when a scenario event may take place at any time without disrupting the flow of an arc. For example, a virtual character who is not visible to trainees may speak to the trainees by radio, which is plausible to occur at any time regardless of other scenario events. ASR allows defining the same timing constraints, in-simulation triggers, and impact on learning for floating events. With this uniformity in representation, the tailoring system selects floating events via the same decision process as selecting arcs. Floating events are especially useful for encoding scenario events designed to distract or direct the trainee's attention and for presenting helpful prompts.

HSP values for an arc or floating event are defined in relation to each of the many skills a scenario trains. In Figure 1, a simple example suggests how ASR parsimoniously represents many possible narrative variations and their possibly conflicting impacts on training. A village leader is engaged in a meeting with some suspicious characters, which the trainee should be able to notice and recognize as important to the mission. The example arcs describe multiple ways the computer can make the suspicious behavior more or less challenging for trainees with different skill sets to recognize. Despite the capability to create and edit many narrative paths and instructional goals, the design of ASR ensures the overall authoring tasks remain at a level of abstraction instructors can use to control effective computer tailoring of the narrative.

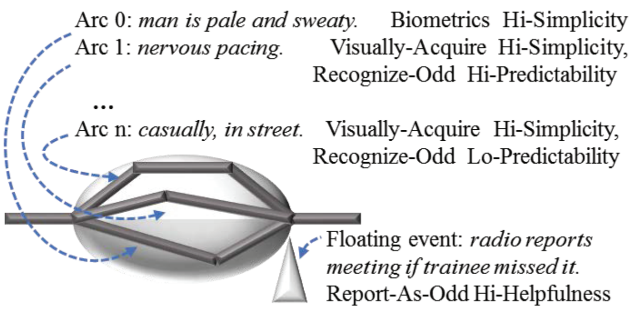

Figure 1: Example arcs for tailoring a training narrative.

Tailoring Scenarios to Improve Training

ASR is readable by a real-time plan execution system that tailors simulation scenarios, the Dynamic Tailoring System (DTS). For a description of DTS processes, see (Wray and Woods 2013, Folsom-Kovarik et al. 2014). The DTS is able to estimate trainee competence in multiple target skills and select the most appropriate scenario variation based on the helpfulness, simplicity, and predictability values that align with the trainee proficiency and training goals.

Since HSP values for a path may be defined in relation to each of the many skills a scenario trains, they provide an optimization surface which is frequently underspecified or overspecified. During scenario execution, the DTS uses a belief-desire-intention (BDI, Bratman 1987) decision formulation to select the most appropriate arc for each trainee. The DTS maintains a belief about a trainee's proficiency in each skill and a desire to increase various proficiencies corresponding to different instructional strategies. Desires result in an intention to tailor the scenario by selecting or sequencing available arcs or floating events.

We illustrate by continuing with our suspicious meeting example. The simulation environment reports a trainee visually acquired the village leader moving in a copse of trees, but then failed to report the leader's presence there as out of the ordinary. The DTS forms a belief the trainee has high proficiency in visual scan skills but low proficiency in cultural understanding skills. Therefore, the DTS decreases the simplicity of visual scan tasks and increases helpfulness for understanding. Then, the next time the leader behaves sneakily he moves on a path obscured by trees that is harder to visually discriminate, but once he is seen a simulated character speaks up to mention the event is unusual.

Within the DTS desires are expressed by target values in each training skill for all three HSP dimensions. The target values are set based on beliefs about trainee skill, and the desired HSP values taken as a whole correspond to instructional strategies such as supporting or challenging a trainee. Each might be valid ways to increase skill: for example most beginners might need support to notice a meeting, but for others challenging complexity followed by a helpful last-minute save might provide a valuable wake-up call. Instructional strategies are general across domains and are configurable, but they are not expected to change frequently and are not exposed to instructor control. In the future, we could let instructors select instructional strategies to match personal preference.

Finally, DTS carries out its intent by selecting available arcs that most closely match its HSP goals. The outputs that DTS has available (event sequences, contained in arcs through a pearl) must align with the capabilities of the simulation environment. A detailed, immersive simulation world will often be limited to mostly linear scripts that describe movements of individual characters, vehicles, or groups. In such a constrained environment, arc pre- and post-conditions (e.g. locations of characters at the start and finish) must be given some direct design consideration to avoid transitions trainees could experience as continuity errors. While such constraints are strict, they do not necessarily limit the kinds of tailoring that are useful for training. More automated (or less detailed) simulation environments able to include non-script definitions of events within arcs offer greater flexibility and could further reduce authoring effort with ASR in the future, because the simulation environment could take more responsibility for fulfilling the specifications of scenario events.

Observations and Lessons

By observing various authors use the scenario representation to create training content, we began to identify successes and limitations that arose during its initial use. We found that the ASR made tailoring choices easy to summarize with descriptions that were clear to the computational DTS. However, we also found that authors frequently requested additional control over certain aspects of tailoring that the ASR was designed to abstract away.

As expected, ASR enables instructor authoring of tailorable training. First, instructors found defining arcs to be easy and relatable to training variation they perform now. The HSP inputs that describe different arcs are easy to sum up with a spreadsheet, and the simple mapping from skill estimates to tailoring decisions appeared sufficient to describe authors' intent. The spreadsheet could then be translated directly into DTS configuration files. Second, HSP values are useful for displaying explanations of DTS decisions, so that instructors could understand system behavior and successfully go back to change it in an iterative manner. Third, the ASR also allowed working backwards to identify needed content: if no arc through a pearl offered the desired tailoring impact, that fact could be made apparent and could then drive the creation of new scripts in the simulation environment.

Any representation makes some ideas easier to express than others. We discovered some authors wanted to use ASR differently than its current design allowed. Additional research is needed to improve the author-facing metaphors.

First, authors often wanted direct control over event choices rather than letting DTS choose based on HSP values. Generally speaking, in the current state of the art the typical processes that create training (such as writing a story, a movie, or a video game script) include direct control over the viewer experience. The HSP dials metaphor did not align with this expectation. For example, authors wanted to say that whenever a trainee makes a specific mistake, the system should respond with a specific intervention. But with more than very simple branches, such direct authoring would become very burdensome. We accommodated such requests with a system of triggers. A solution more in line with the HSP dials might be automated balancing of dials to achieve targeted responses.

Second, our simulation environment's requirement to define arcs as unchanging scripts limited authors' ability to nest choices or make changes that would vary script durations such as looping scripts. Authors also requested the ability to create branches that break the pearl metaphor by failing to converge or by transferring information between pearls, such as a memory of which arc was chosen in an earlier event. To accommodate this, we would need to generalize pearl threads. We could retain some of the simplicity advantages of our existing thread structure by introducing a concept of narrative thread inheritance or subtyping. A variant branch could inherit most of the narrative in the main branch, with only certain content overridden, so that the author can reduce the work created by authoring two chains of pearls that might never converge. However, a hard problem would still remain in ensuring that large numbers of branches all produce the desired training.

As suggested previously, a different simulation environment that does not strictly script pearl contents would offer improved authoring abstractions. Examples might include directing that the village leader moves to a location without specifying his path and the amount of time he needs, or requiring that half the characters in a crowd display a behavior without needing to know which characters will be present. In effect, greater intelligence in the simulation environment could let us relax the strict precondition matching currently required in ASR while retaining its ability to guarantee post-conditions on pearls. In order to take advantage of content abstraction, ASR will require either runtime staging automation or authoring-time validation of content. Not simply a convenience factor for authoring, these represent the completion of the ASR vision.

If a future ASR were allowed to simply state that a character is at a location without specifying how to go there, then an automated staging approach would calculate at runtime when that character needs to depart in order to reach the destination in time. We note that automated staging of many characters, which is more about their precise locations at a specific time and less about their realistic motivations in arriving there, may require a different kind of narrative planning than has been explored to date.

On the other hand, an authoring validation approach would instead examine all paths through the relevant pearls to ensure that the character will always be able to meet the precondition imposed by the one path that includes the movement command. We suggest that additional automation of validation would be possible in ASR by applying a more formal semantics to the narrative representation. For example, event start and end points could be marked with times, either absolute or relative to other events. Authors would need to specifically mark a limited number of event dyads with relationships that fit within a folk accounting of causality, such as ends-before, starts-during, and so on. Current work suggests that such relations would suffice to describe a large number of training scenarios.

Next Steps and Future Work

ASR has been tested in an informal evaluation using military personnel expert in instruction. Two instructors used ASR to define narrative tailoring for observational training. A scenario was encoded in an existing first-person immersive simulator without tailoring, and ran for approximately 20 minutes of training time. The instructors then authored ASR, with tailoring, for the same scenario.

When complete, DTS used the authored ASR to make real-time tailoring choices judged similar to those a human instructor might make. The initial evaluation also found the authoring task took 60 minutes for one user and 78 minutes for the other, times which were less than expected for a prototype and judged suitable for operational use. In addition, a researcher on the team took 52 minutes to accomplish the same task, suggesting that the nontechnical users did not experience a large slowdown. An evaluation with larger sample size is required to enable tests for outcome significance. However, our team next intends to speed the authoring process further because our experience in the field suggests the optimal authoring time should be a few seconds to minutes, rather than an hour. Some of the changes in the previous section may help speed authoring.

First, narrative planners may be appropriate tools to carry out the planning we identified in the previous section and improve the authoring experience. We will examine ways to either encode the information needed for planning in parallel to the ASR, so it works without requiring direct authoring, or to expose specific subsets of the planning knowledge in a way that technical expertise is not required to manipulate. The success of existing approaches (e.g., Riedl et al. 2008) suggests a related planning approach may help accomplish interstitial *repair*, or detection that the trainee has departed a desired progression of events and selection or manipulation of events to return the trainee to the track and avoid wasted time or premature end states.

An interesting research question surrounds the handling of some unexpected errors, or errors whose repair need not be explicitly encoded. In order to carry out repair of unexpected digressions, the range of expected trainee actions could be encoded using the same representation as narrative events. If successful, the DTS could determine when trainees leave the expected envelope, predict how the digression will affect the scenario, and choose the most appropriate repair without needing the author to spell out explicit triggers in ASR.

Another extension would incorporate *spatial* planning such as staging characters in order to display coordinated events despite unpredictable movement times and other challenges. Spatial planning might let authors easily describe reusable, hierarchical events or story beats that could be presented at variable times, as did Mateas and Stern (2003), rather than needing to specify all the movement that leads up to each encounter. The result would be somewhat like removing the strings between pearls.

Finally, we look forward to new and innovative ways to generate instructional goals and impacts without authors directly specifying them. These could be learned from examples or represented in a different abstraction that in turn creates the HSP metaphor used here.

Acknowledgement

This work is supported in part by Office of Naval Research (ONR) contract N00014-11-C-0193. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of ONR or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

Bratman, M. (1987). "Intention, plans, and practical reason."

Carpentier, K., D. Lourdeaux and I. M. Thouvenin (2013). Dynamic Selection of Learning Situations in Virtual Environment. 5th International Conference on Agents and Artificial Intelligence (ICAART).

Durlach, P. and A. Johnson (2014). "Advanced Distributed Learning Initiatives 2014." Journal of Advanced Distributed Learning Technology **2**(6): 35-48.

Folsom-Kovarik, J. T., C. Newton, J. Haley and R. E. Wray (2014). Modeling Proficiency in a Tailored, Situated Training Environment. 23rd Conference on Behavior Representation in Modeling and Simulation (BRIMS).

Folsom-Kovarik, J. T., R. E. Wray and L. Hamel (2013). Adaptive assessment in an instructor-mediated system. 16th International Conference on Artificial Intelligence in Education, Memphis, TN.

Kray, L. J. and A. D. Galinsky (2003). "The debiasing effect of counterfactual mind-sets: Increasing the search for disconfirmatory information in group decisions." Organizational Behavior and Human Decision Processes **91**(1): 69-81.

Majewski, J. (2003). Theorising video game narrative, Bond University.

Medler, B. and B. Magerko (2006). Scribe: A tool for authoring event driven interactive drama. Technologies for Interactive Digital Storytelling and Entertainment, Springer**:** 139-150.

Riedl, M. O. (2009). Incorporating Authorial Intent into Generative Narrative Systems. AAAI Spring Symposium: Intelligent Narrative Technologies II.

Riedl, M. O. and A. Stern (2006). Believable agents and intelligent scenario direction for social and cultural leadership training. Proceedings of the 15th Conference on Behavior Representation in Modeling and Simulation.

Riedl, M. O., A. Stern, D. M. Dini and J. M. Alderman (2008). "Dynamic experience management in virtual worlds for entertainment, education, and training." International Transactions on Systems Science and Applications, Special Issue on Agent Based Systems for Human Learning **4**(2): 23-42.

Rowe, J., B. Mott, S. McQuiggan, J. Robison, S. Lee and J. Lester (2009). Crystal Island: A narrative-centered learning environment for eighth grade microbiology. Workshop on Intelligent Educational Games at the 14th International Conference on Artificial Intelligence in Education, Brighton, UK.

Schatz, S., R. E. Wray, J. T. Folsom-Kovarik and D. Nicholson (2012). Adaptive Perceptual Training in a Virtual Environment. 56th Human Factors and Ergonomics Society Annual Meeting (HFES), Boston, MA.

Schumacher, R. M. and D. Gentner (1988). "Transfer of training as analogical mapping." IEEE Transactions on Systems, Man and Cybernetics **18**(4): 592-600.

Sieck, W. R., J. L. Smith and L. J. Rasmussen (2013). "Metacognitive Strategies for Making Sense of Cross-Cultural Encounters." Journal of Cross-Cultural Psychology **44**(6): 1007-1023.

Wray, R. E., J. T. Folsom-Kovarik, A. Woods and R. M. Jones (2015). Motivating narrative representation for training cross-cultural interaction. 4th International Conference on Cross-Cultural Decision Making, Las Vegas, NV.

Wray, R. E., H. C. Lane, B. Stensrud, M. Core, L. Hamel and E. Forbell (2009). Pedagogical experience manipulation for cultural learning. Conference on Artificial Intelligence in Education, Brighton, England.

Wray, R. E. and A. Woods (2013). A Cognitive Systems Approach to Tailoring Learner Practice. 2nd Advances in Cognitive Systems Conference, Baltimore, MD.