Presenting Believable Choices

Justus Robertson  
Department of Computer Science  
North Carolina State University  
Raleigh, NC 27695  
jjrobert@ncsu.edu

R. Michael Young  
School of Computing  
The University of Utah  
Salt Lake City, UT 84112  
young@cs.utah.edu

Abstract
Interactive narratives (IN) are stories that branch and change based on the actions of a participant. A class of automated systems generate INs where all story branches conform to a set of constraints predefined by an author. Participants in these systems may create invalid branches by navigating the story world outside the bounds of the author’s constraints. Two existing methods, choice removal and intervention, are designed to mitigate these situations. However, these methods are expected to lower invisibility, being recognized as system manipulations by the participant. In this paper we present an evaluation of a new method, domain revision, that is designed to have no negative effect on invisibility. We measure invisibility by asking survey participants how believable a choice’s options and outcomes are in the context of a Choose Your Own Adventure story. We find that domain revision is more believable than choice removal when applied to a choice’s options. We also find that domain revision is equally believable as intervention on a choice’s outcomes because intervention does not cause a drop in invisibility.

Introduction
Interactive narratives (IN) are branching stories that adapt their content according to actions a participant makes while playing. Popular examples of interactive narratives range from early printed text games like the Choose Your Own Adventure series (Packard 1979) to modern video games like Mass Effect (BioWare 2007) and The Walking Dead (Telltale Games 2012). One problem with interactive storytelling, called the authorial bottleneck (Bruckman 1990), is the exponential amount of narrative content required to create truly branching storylines. One way to address the authorial bottleneck is to create a storytelling agent, called an experience manager (Riedl and Bulitko 2013), to automate the process of interactive narrative content creation and control.

One kind of interactive storytelling agent, called a strong story agent, generates a central data structure that encodes all possible sequences of narrative events. Strong story systems generate this data structure using a linear narrative generator which creates stories that conform to a set of predefined constraints. This linear storyteller can be used in conjunction with a process that solves the boundary problem (Magerko 2007), also called the narrative paradox (Louchart and Aylett 2003), the problem of telling an interesting story while allowing players to act freely and effect change in the story world. The combined system creates a tree of possible stories where each branch conforms to author constraints and accounts for possible user actions (Riedl and Young 2006). This tree is built by analyzing the current story, and for every action the player could take to contradict the current narrative, creating a new story that accounts for the player’s action and conforms to the author’s constraints.

However, not all possible player actions may be accommodated by this process. Some actions may maneuver the system into a state such that the author’s constraints no longer hold and no new story can be generated. A second method of handling user actions that contradict the current story, called intervention, swaps the effects of the action performed by the player for a second set that do not derail the story. However, intervention has fallen out of use in experience management systems because it is expected to cause a decrease in invisibility (Roberts and Isbell 2008). Invisibility is a violation of the participant’s suspension of disbelief. If the user becomes aware there is a process in control of the story world, influencing the course of world events by violating previously established world mechanics, it should negatively impact the participant’s gameplay experience.

In this paper, we present an empirical evaluation of an intervention method that tracks the player’s knowledge of world mechanics to perform interventions without contradicting what the participant has observed. We compare this approach in a Choose Your Own Adventure environment to both classic intervention and a third method, called choice removal. We find that revision and intervention are more believable to participants than choice removal, but the outcome of interventions were reported to be just as believable as the other two methods. This result contradicts the expectation that inconsistencies caused by interventions would be recognized by study participants. We finish the paper with a discussion of the results and ideas for future studies.

Related Work
The first system to introduce intervention as a method of mediating between player actions and author constraints was
Mimesis (Young et al. 2004). Mimesis was an experience management framework that controlled NPC characters and world mechanics in the Unreal Tournament game engine based on stories generated by a narrative planner. From its initial characterization, researchers identified the potential for intervention to break the participant’s suspension of disbelief by alternating between different action outcomes based on author constraints (Riedl, Saretto, and Young 2003; Roberts and Isbell 2008). More recent Mimesis-like systems have dropped support for intervention (Riedl et al. 2008; Ramirez and Bulitko 2014; Robertson and Young 2014b).

The work evaluated in this paper addresses this concern by allowing an experience manager to use interventions only when the shift between alternate outcomes does not contradict what the player has observed during gameplay. This approach is similar to alibi generation (Sunshine-Hill and Badler 2010), the process of dynamically explaining an NPC’s behavior in a sandbox game environment. Alibi generation allows NPCs to act randomly and then appear intelligent when a player begins observing the character. In the past, alibi generation has been applied to generating narrative events in interactive storytelling (Li et al. 2014; Robertson and Young 2014a). The method evaluated in this paper applies the process to dynamically configuring interactive narrative game mechanics. Game mechanics are configured by shifting between alternate models that formally describe the outcomes of actions characters can take in the story world. This process modifies models similar to work done in automatically generating new actions in interactive narrative domains (Porteous et al. 2015).

This paper presents an evaluation of this intervention method by measuring how believable choice options and outcomes are in the context of a Choose Your Own Adventure story. This study ties into a growing body of work that examines types of situations, choice options, and choice outcomes in interactive narratives and their psychological effects on players (Mawhorter et al. 2014). Work in this area has examined the link between choice options, choice outcomes, and a player’s feelings of agency (Fendt et al. 2012; Cardona-Rivera et al. 2014), the link between choice options, player behavior, and story enjoyment (Yu and Riedl 2013; 2015), and a validation of a generative theory of certain pairings of situations and choice options, like dilemmas (Mawhorter, Mateas, and Wardrip-Fruin 2015).

**Background: Domain Revision**

Domain revision is a method used by interactive narrative systems to present targeted interventions that do not contradict what the player has observed in a story world. This paper builds off of system definitions for doain revision provided by Robertson and Young (2015); in this section we briefly describe their domain revision mechanics and refer the reader to their paper for details.

In approaches that employ domain revision, the system automatically shifts between alternate versions of world mechanics to prevent contradictions between a player’s actions and an author’s constraints, working to ensure that the shift will not be noticed by the player. In this approach, action domains are managed in a strong story experience management system and modeled using the Planning Domain Definition Language (PDDL) (McDermott et al. 1998). PDDL is a logical language used to formally specify world states and the nature of transitions between them. A PDDL state specifies what is true in a model world. Anything not listed in a PDDL state is considered false. A transition from one state to another occurs just when an actor performs an action in the first state, resulting in the second. Actions are modeled with general templates that can be applied to specific situations and produce context-sensitive outcomes. An action can be applied to a state if the state satisfies the action’s preconditions. A precondition is a declarative statement about the world that must be true in order for an action to be performed. When an action is performed, it updates the state by making a list of effects true. Interactive narrative gameplay can be performed in a PDDL framework by maintaining a state that represents the story world and allowing story characters to iteratively apply actions to change the world and move the story forward.

Using a PDDL-like action representation, off-the-shelf (Helmert 2006) or narrative-centered (Riedl and Young 2010; Ware and Young 2014) planners can be used to find sequences of actions for both players and non-player characters that satisfy both the semantics of the domain theory and the representation of a domain author’s constraints on an interactive plot line. As the player takes action as a character in the story world, the system uses its planner to find narrative sequences that accommodate the user’s actions, drive the world’s non-player characters and also match the author’s intended constraints until the story reaches a conclusion.

However, the player may be able to take actions with effects that contradict the author’s constraints. One open problem in this framework is how to handle situations where the player and author are in opposition. One method, called intervention, prevents the player’s action from threatening the author’s constraints by swapping the action’s effects for an alternate set that do not contradict the author. This alternate set of effects is called the action’s failure mode. One potential limitation of intervention is that a system’s moving between the use of the regular action’s effects and those of its failure mode allows the player to realize that the system is actively manipulating the flow of events.

Domain revision was designed to allow system interventions without jumping back and forth between a base action outcome and a failure mode. Domain revision is outfitted with a model of player knowledge that tracks what the player observes as they progress through a story. When the system decides to use a failure mode, it can only happen if the player has not observed the original action’s outcomes. Once the system uses a failure mode, it can never use the original action outcomes again. Using domain revision, an experience manager can intervene only when the intervention does not contradict what the player has observed in the story world. In the next section we present an evaluation of the method by comparing it to two alternatives, intervention and choice removal.
Journeying to the Woods
You journey to the woods
You are no longer in the armory
You are on a path to the woods
You see Sam journey to the woods
You are walking on a path to a large wooded area in front of you. Your partner, Sam, is walking beside you. You see two small goblins ahead of you. The goblins appear evil and threatening. One goblin is carrying a bow and a quiver of arrows. The other goblin is carrying a sword.

**Attack the small goblin carrying a bow and arrows**
**Attack the small goblin carrying a sword**
**Move past the goblins to the woods**

Figure 1: A scene from the introductory segment of the CYOA. Outcomes from each character’s last action are displayed at the top of the page. A paragraph describing the current situation is presented in the middle. The choices available to the player are listed at the bottom.

Experiment
In this section, we present an evaluation of domain revision in the context of a Choose Your Own Adventure. The purpose of the experiment is to compare the invisibility of domain revision choice options and choice outcomes against a baseline we expect to have low invisibility. We recruited participants on Amazon Mechanical Turk to play the CYOA online. We measured invisibility by asking participants survey questions about how believable choice options and choice outcomes were once the CYOA is over.

To test choice options, we use a method called choice removal. If the outcome of a choice will negate an authorial constraint, the method of removal will not present the choice as an option to the player. We expect choice options when removal is performed to be rated as less believable than domain revision. To test choice outcomes, we use classic intervention, unconstrained by a model of player knowledge. We expect choice outcomes when classic intervention is performed to be rated as less believable than domain revision.

Treatments
Figure 1 shows a screen from the CYOA. Results from the last choice the player made, along with the actions and outcomes of actions performed by other story characters, are listed at the top of the page. A middle paragraph describes the current state of the world. A list of possible actions are listed as hyperlinks at the bottom of the page. When the participant clicks one of the hyperlinks, the story advances to a page containing the next set of choice outcomes, the new situation, and the new choice options.

Each of the three treatments begin with the same introduction sequence, pictured in Figure 2a. Each story begins in an armory. The player is with their companion, Sam, and is tasked with vanquishing an evil goblin that resides in the woods near town. The player can choose to arm themselves with a bow, a sword, or proceed immediately to the woods. On the way to the woods, the player is confronted by two small goblins, one carrying a bow and one carrying a sword.

The player can attack either of the goblins or proceed past the goblins to the woods. This is the choice depicted in Figure 1. Once the player reaches the woods, they find a large goblin wearing armor. At this point, the three versions of the story begin to differ.

All three versions share the same three authorial constraints: the armored goblin should vanquish Sam in the woods, the player should be in the woods when Sam is vanquished, and the player should vanquish the armored goblin at the clearing beyond the woods.

Choice Removal The choice removal treatment, pictured in Figure 2b, forces the player to wait a turn in the woods since both attacking the goblin and moving to the clearing would contradict authorial constraints. As the player waits, the goblin attacks and vanquishes their companion Sam. The system then forces the player to move to the clearing beyond the woods by removing the attack action, since attacking the goblin in the forest would contradict the constraints. Finally, the player attacks and vanquishes the large goblin in the clearing beyond the woods. The successful attack outcome is pictured in Figure 3b.

Intervention The intervention treatment, pictured in Figure 2c, allows the player to attack or move past the large, armored goblin in the woods. However, if the player attempts to attack the goblin, their attack is intervened against and fails. The failure outcome is pictured in Figure 3a. The CYOA tells the player not only that their attack fails to pierce the goblin’s armor, but that the goblin’s armor will always be resistant to player attacks. If the player attempts to move, the action will also be intervened against by the system by the goblin blocking the player’s way.

Once the large goblin vanquishes Sam, the player has another opportunity to attack the goblin or move around the goblin to the clearing. If the player attacks the goblin, they experience a second intervention before moving to the clearing. If they choose to move around the goblin, they go directly to the clearing. At the clearing, the player attacks the goblin and their attack succeeds. The successful attack outcome is pictured in Figure 3b.

Domain Revision The domain revision treatment, pictured in Figure 2d, allows the player to attack or move past the large, armored goblin in the woods. If the player attempts to attack the goblin, as in the intervention treatment, their attack is intervened against and fails. The failure outcome is pictured in Figure 3a. If the player attempts to move, their action will also be intervened against by the system by the goblin blocking the player’s way.

Things proceed exactly the same in the domain revision treatment as the intervention treatment until the player arrives at the clearing. When the player attacks the armored goblin at the clearing, their attack continues to be deflected by the goblin’s armor. After the player attacks the goblin, they are given the option to push the goblin off the cliff at the edge of the clearing. Pushing the goblin off the cliff vanquishes the goblin without using the player’s weapon.
Figure 2: The flow of choice framings and options in the three versions of the CYOA. Figure 2a pictures the two choices used across all three story versions. Figures 2b, 2c, and 2d picture the choice framings and options from the introduction onward in each of the three versions.

**Setup**

We use a hand-authored CYOA where the player makes choices by navigating hyperlinks in a web browser. Though it is hand-authored, the CYOA’s mechanics mirror those of a PDDL-based strong story experience manager using choice removal, intervention, or domain revision to mitigate the effects of a player’s actions when they violate author constraints. We use a hand-authored CYOA to improve the natural language output compared to the simple text templates our system uses to print state information. The use of a hand-authored CYOA also mitigates the engineering challenges of running an on-line, multi-user experience manager over the Internet. Participants were assigned to the three treatment groups in a round robin manner.

**Survey**

After the participants played through the CYOA, we asked them survey questions about their choices and choice outcomes. For the choice questions, we presented the player with each story situation and set of choice options they encountered while playing the CYOA. For each choice situation-option set pair, we asked participants to rate their agreement with two statements on a 5-point Likert scale from "Strongly Disagree" to "Strongly Agree". The statements were:

1. These are reasonable choices for this situation.
2. These are the choices I would expect.

Below the first two statements, the survey shows the player the choice they made and the outcome of their choice. Under the choice outcome, the survey asks participants to rate three additional statements concerning the outcome:

1. I understand why my choice caused these outcomes.
2. These are reasonable outcomes for this choice.
3. These are all the outcomes I would expect.

**Hypotheses**

There are two locations in the CYOA where we make and test our hypotheses. For the believability of choice options, we examine the choice framing and options when the player first encounters the large, armored goblin in the woods. The choice removal treatment forces the player to wait a turn, where the intervention and domain revision treatments offer the attack and move options the player experiences earlier on their way to the woods. We hypothesize that players will
You attack the large, armored goblin with your sword
The goblin cannot be harmed by your sword while wearing magic armor
The goblin's armor deflects your attack
Your attack does not hurt the goblin

The goblin is vanquished by their weapon.

(a) Unsuccessful attack on armored goblin.

Figure 3: The two types of outcomes of attack actions on the armored goblin. Figure 3a pictures the feedback from unsuccessfully attacking the armored goblin. The game tells the player that the goblin cannot be harmed by their weapon while it wears magic armor. Figure 3b pictures the feedback from successfully attacking the armored goblin. The game tells the player that the goblin is vanquished by their weapon.

(b) Successful attack on armored goblin.

Table 1: Results of MWW U tests on choice option results. All four hypotheses are statistically significant ($p < .05$) and have a large effect size ($r > 0.5$).

<table>
<thead>
<tr>
<th>Question</th>
<th>Hypothesis</th>
<th>$p$</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Intervention $&gt;$ Removal</td>
<td>2.902e-08</td>
<td>0.68</td>
</tr>
<tr>
<td>2</td>
<td>Intervention $&gt;$ Removal</td>
<td>7.246e-08</td>
<td>0.66</td>
</tr>
<tr>
<td>1</td>
<td>Revision $&gt;$ Removal</td>
<td>2.356e-10</td>
<td>0.75</td>
</tr>
<tr>
<td>2</td>
<td>Revision $&gt;$ Removal</td>
<td>4.171e-08</td>
<td>0.67</td>
</tr>
</tbody>
</table>

rate the set of choice options in the intervention and domain revision treatments higher on the two survey questions (reasonableness, expectedness) than the set of choice options in the choice removal treatment.

For the believability of choice outcomes, we examine the outcomes of the player attacking the armored goblin at the clearing. This is the first attack outcome a choice removal participant encounters. The intervention and domain revision treatments previously tell players they will never be able to harm the goblin wearing magic armor. The outcome of the choice removal and intervention treatments is a successful attack, which is consistent with previous feedback in the removal case but inconsistent in the intervention case. The outcome of the domain revision treatment is an unsuccessful attack, which is consistent with previous feedback. We hypothesize that players will rate outcomes in the choice removal and domain revision treatments higher on the three survey questions (understandability, reasonableness, expectedness) than outcomes in the intervention treatment.

We used a Mann-Whitney-Wilcoxon (MWW) U test (Mann and Whitney 1947; Wilcoxon 1945) to determine whether a statement was more agreed with in one treatment when compared to another.

Results

Ninety subjects participated in the experiment. After filtering out inattentive subjects that incorrectly answered a trick question (Please answer this question "Strongly Disagree") and subjects that went down unusable paths (never attacking the armored goblin in the woods), data from 30 participants was collected in the removal treatment, 27 in the intervention treatment, and 28 in the revision treatment for the choice option statements. For the outcome statements, data from 30 participants was collected in the removal treatment, 24 in the intervention treatment, and 21 in the revision treatment.

The data we gathered support our choice option hypotheses, that intervention and domain revision would be rated higher than choice removal in responses to the choice option Likert survey. A summary of the data is shown in Table 1 and a graph of the data is shown in Figure 4. The choice removal treatment’s set of choice options for the first encounter with the armored goblin in the woods is rated lower on both statements than the intervention and revision treatments. All four hypotheses have $p$-values under .05 and effect sizes greater than 50%. More interestingly, the data did not support any of our choice outcome hypotheses, that choice removal and domain revision would be rated higher than intervention in responses to the choice outcome Likert survey. A summary of the data is shown in Table 2 and a graph of the data is shown in Figure 5. Our hypotheses are not supported because the intervention treatment is rated to be just as believable as the choice removal and domain revision treatments, even though event outcomes in the intervention story violate earlier player observations about story world mechanics.

Discussion

The data do not support our expectation that story world inconsistencies introduced by system interventions are easily detected by interactive narrative participants. However, we are skeptical that repeated inconsistencies introduced by interventions over a long enough time frame will go unnoticed. In this section, we discuss three continuu that may help explain the data and guide future work: novice vs. expert, game vs. narrative, and choice options vs. outcomes.

One explanation for the results is that there is some pro-
### Table 2: Results of MWW U tests on choice outcome results.

<table>
<thead>
<tr>
<th>Question</th>
<th>Hypothesis</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Removal &gt; Intervention</td>
<td>0.2657</td>
</tr>
<tr>
<td>2</td>
<td>Removal &gt; Intervention</td>
<td>0.1841</td>
</tr>
<tr>
<td>3</td>
<td>Removal &gt; Intervention</td>
<td>0.9827</td>
</tr>
<tr>
<td>1</td>
<td>Revision &gt; Intervention</td>
<td>0.9423</td>
</tr>
<tr>
<td>2</td>
<td>Revision &gt; Intervention</td>
<td>0.4325</td>
</tr>
<tr>
<td>3</td>
<td>Revision &gt; Intervention</td>
<td>0.95</td>
</tr>
</tbody>
</table>

None of the six hypotheses are statistically significant (p < .05).

### Conclusion

Using a Choose Your Own Adventure instrument, we evaluated a method of intervention, called domain revision, that does not contradict what the player has observed within the game world. We compared domain revision to two other methods — intervention and choice removal — and found that domain revision is more believable than removing choice options. However, we found that domain revision is just as believable as presenting contradictory evidence using intervention on choice outcomes. This is contrary to our expectation that inconsistencies introduced by interventions will be problematic for story participants.

### References


Li, B.; Thakkar, M.; Wang, Y.; and Riedl, M. O. 2014. Data-Driven Alibi Story Telling for Social Believability. In *Social Believability in Games Workshop*.


---

**Future Work**

Future work will involve creating more experiments to gauge the cause and extent of intervention’s null impact on invisibility. An experiment could be created for each of the possible explanations covered in the last section. An experiment to test the first possibility, that players experience some amount of change blindness towards a game’s mechanics while they are a novice, could be tested with longer playthroughs or repeated playthroughs of similar CYOAs featuring the same mechanics. If this case is true, players should better recognize outcome inconsistencies the more they use an action. The second possibility, that players are more accepting of inconsistencies in the context of a narrative, could be tested by running a second experiment where narrative information like the characters and setting are replaced with more abstract objects that allow the player to focus on the game rules directly instead of viewing them through the lens of a narrative. The final possibility, that players are more aware of choice options than outcomes, is harder to test than the other possibilities. One indicator may be if the other experiments are run and players continue to rate inconsistent outcomes as believable and removed choices as not believable.


