Game User Interaction & Intelligence

Northeastern University

Exploratory Automated Analysis of Structural Features of Interactive Narrative

Nathan Partlan[1] Elin Carstensdottir[1] Sam Snodgrass[1]
Erica Kleinman[1] Gillian Smith[2] Casper Harteveld[1] Magy Seif El-Nasr[1]

[1] Northeastern University
[2] Worcester Polytechnic Institute

Introduce Self: Name, Advisor, University, co-authors
Introduce Topic: Lots of words, so let's talk about what that really means.

Primary Questions

Image: http://ipkitten.blogspot.com/2012/11/to-settle-or-not-to-settle-what-would.html

I'd like to start by describing the primary questions we're trying to answer with this work, and then we'll go back and talk about each section in more detail.

Image: http://ipkitten.blogspot.com/2012/11/to-settle-or-not-to-settle-what-would.html

How can we automatically encourage and support designers' creativity?

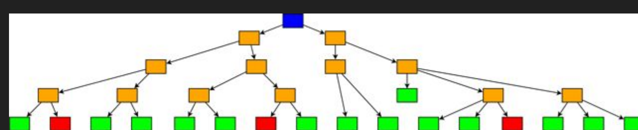
Man and his "brain children."

James Ryan. 2017. "Popular Visual Depictions of Early Generative Systems," in *Seeds* vol. 2.

Original: The Lincoln Star, Oct. 23, 1960, p. 48.a

The real, big question is: how can our tools better support and encourage game designers' creativity and aid them in reflection and iteration on their design, specifically in interactive narrative?

What representation enables us to understand an interactive narrative’s structure?

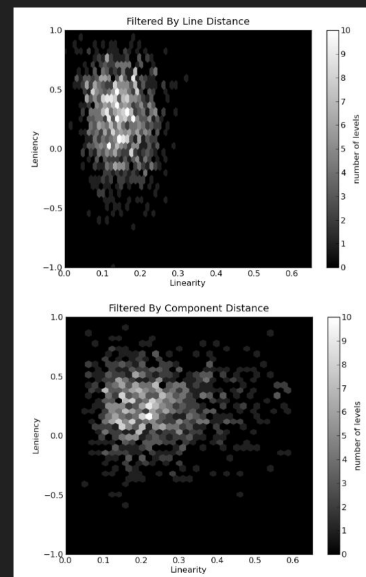


Ashwell, Sam Kabo. 2015. “Standard Patterns in Choice-Based Games.” *These Heterogenous Tasks* (blog). January 27, 2015.

But to get there, we need to computationally understand as much as possible about what the designer is building. That's a hard problem - games, and interactive narrative games, especially, are full of complex, culturally informed meaning.

But game designs do have structure, and interactive narratives do as well. We decided to start by trying to learn everything we can from the structure of the narrative, and how it interacts with players. We wanted to figure out what sort of representation would enable automated analysis and exploration of that structure.

How do metrics enable analysis of, and comparison between, scenarios?



Smith, Gillian, and Jim Whitehead. 2010. “Analyzing the Expressive Range of a Level Generator.” In *Proceedings of the 2010 Workshop on Procedural Content Generation in Games*, 4:1–4:7. PCGames ’10. New York, NY, USA: ACM.

After determining a representation, the next question is what we can measure from it. In this work, we do this through metrics, which each capture specific attributes of the scenario and enable analysis of, and comparison between, scenarios.

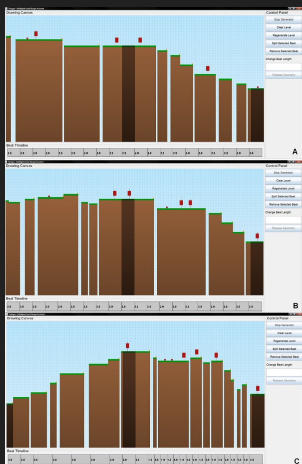
Mixed-Initiative AI

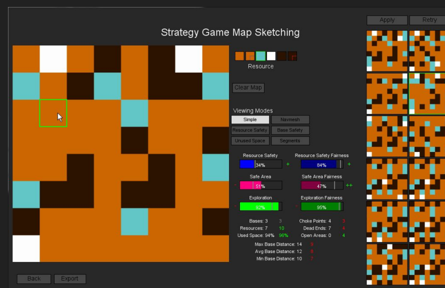
Image: Content-Aware Fill in Photoshop

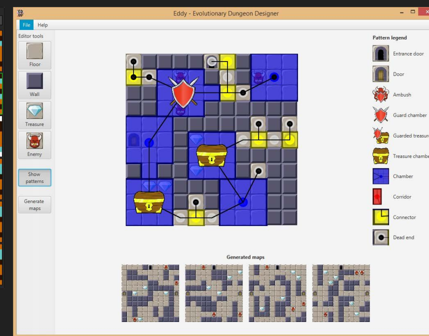
Image: https://4.bp.blogspot.com/-T454JVtuG78/T6l360ilR1I/AAAAAAAAAGo/pzssWNMAYRQ/s1600/content-aware.jpg

All of this is leading towards "mixed-initiative AI," a growing area of interest in AI and games research.

Image:
https://4.bp.blogspot.com/-T454JVtuG78/T6l360ilR1I/AAAAAAAAAGo/pzssWNMAYRQ/s1600/content-aware.jpg







AI As Partner

Baldwin, Alexander, Steve Dahlskog, Jose M. Font, and Johan Holmberg. 2017. "Towards Pattern-Based Mixed-Initiative Dungeon Generation." In Proceedings of the 12th International Conference on the Foundations of Digital Games, 74:1–74:10. FDG '17. New York, NY, USA: ACM.

Liapis, Antonios, Georgios N. Yannakakis, and Julian Togelius. n.d. "Sentient Sketchbook: Computer-Aided Game Level Authoring." In In Proceedings of ACM Conference on Foundations of Digital Games, 2013.

Smith, Gillian, Jim Whitehead, and Michael Mateas. 2010. "Tanagra: A Mixed-Initiative Level Design Tool." In *Proceedings of the Fifth International Conference on the Foundations of Digital Games*, 209–216. FDG '10. New York, NY, USA: ACM.

In a mixed-initiative system, the AI is a partner, with the AI sometimes leading the human designer, and the human designer taking the initiative at other times. In game design, most mixed-initiative systems have focused on visual level design.

This work begins to expand the domain for these tools into narrative and mechanics in games by providing a starting point for metrics and suggestions in a new domain.

Examples of this include Tanagra, by Gillian Smith et. al., Sentient Sketchbook, by Antonios Liapis et. al., and EDDY, by Alexander Baldwin et. al. Each of these tools provides options, explains patterns and metrics it sees, and enables the human designer to make informed choices and design with those suggestions in mind.

Interactive Narrative

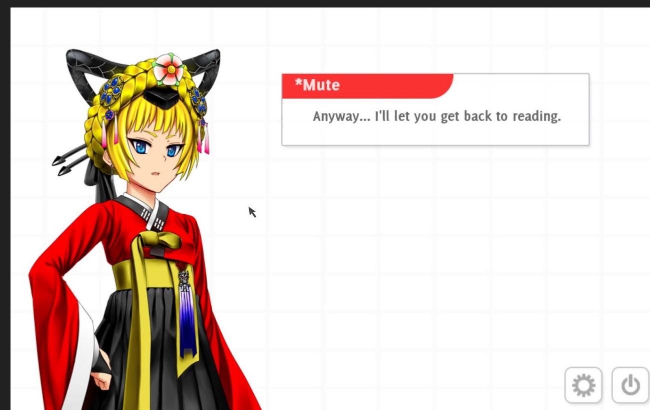
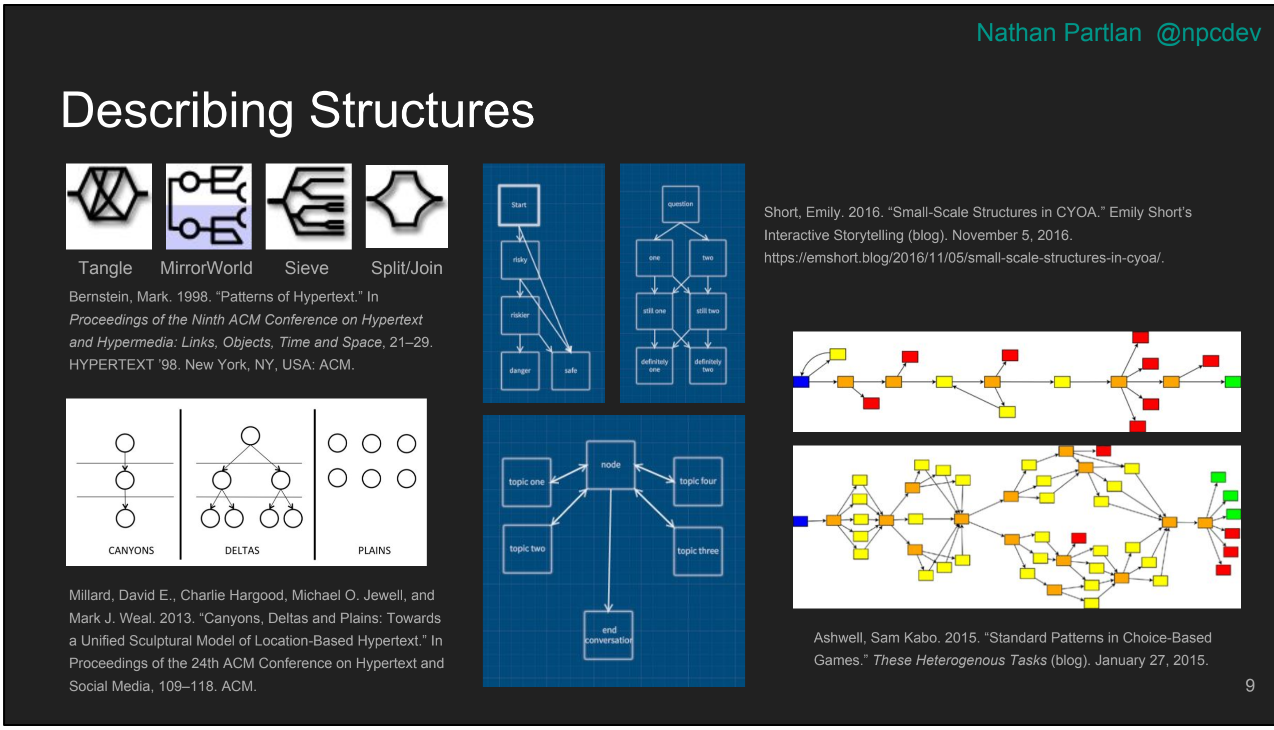


Image: Analogue: A Hate Story, Christine Love (retrieved from http://www.choicestgames.com/2013/05/analogue-hate-story-review.html)

Let’s talk about interactive narrative. It is a rich domain, with a long history and many styles.

In this work, we mostly focus on choice-based interactive narrative, such as visual novels, but we expect that our work could generalize to other games where narrative is important.

Image: http://www.choicestgames.com/2013/05/analogue-hate-story-review.html



Several researchers have worked on describing interactive narrative structures. Here are some examples!

Many of these structural descriptions point to the use of graphs for narrative understanding, but their definitions are not fully operationalized for a computational analysis of the structures they describe.

This research goes back at least to the '90s, with the roots of research in traditional narratives going back many years before that. We can't possibly do justice to all that research in this talk, but some good examples are Bernstein's "Patterns of Hypertext," Millard et. al. describing "Canyons, Deltas, and Plains," Ashwell detailing "Standard Patterns in Choice-Based Games," and Short's "Small-Scale Patterns in CYOA."

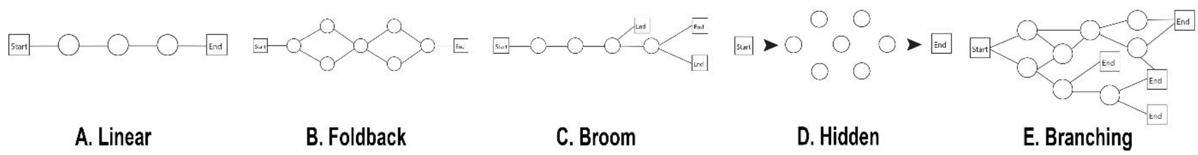


Fig. 3. Five (of six) story structures identified in the study.

Carstensdóttir, Elín, Erica Kleinman, and Magy Seif El-Nasr. 2017. "Towards an Interaction Model for Interactive Narratives." *ICIDS 2017*.

Narratives as Graphs

Recently, some of my co-authors, Elin Carstensdottir, Erica Kleinman, and Magy Seif El-Nasr, have been working to more fully operationalize the analysis of these structures and their interactions with players.

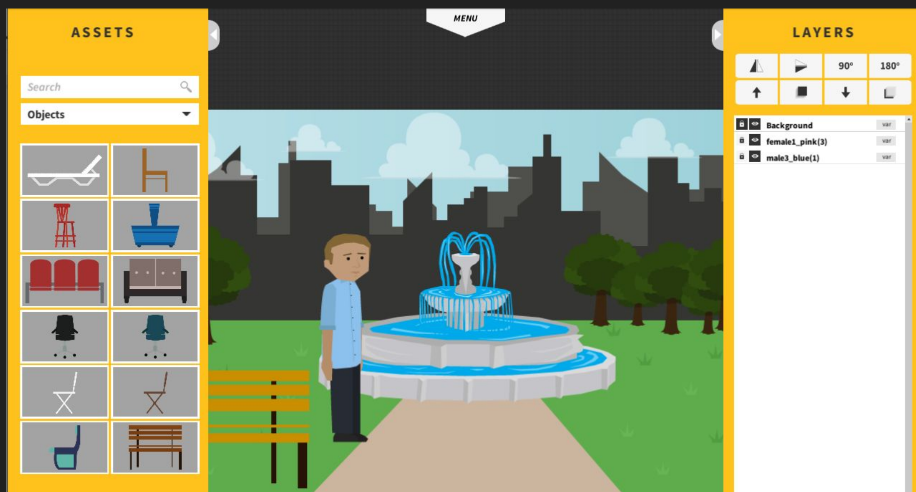
We have implemented their ideas in this work, using a graph-based representation based on Elin's "Interaction Maps." We will describe the details of interaction maps soon, but first, let's talk about the particular scenarios we are representing.

Context: StudyCrafter

We implemented our model for StudyCrafter, an existing design tool that is in continuing development.

Nathan Partlan @npcdev

2D, Playful Scenarios for Social Science Research



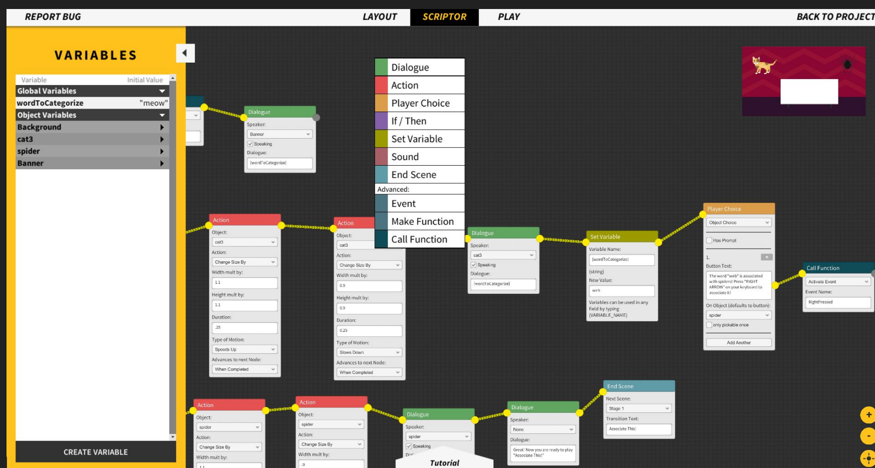
StudyCrafter is an easy-to-use tool for designing playful, 2D scenarios, often using interactive narrative.

It's often used to build social science experiments.

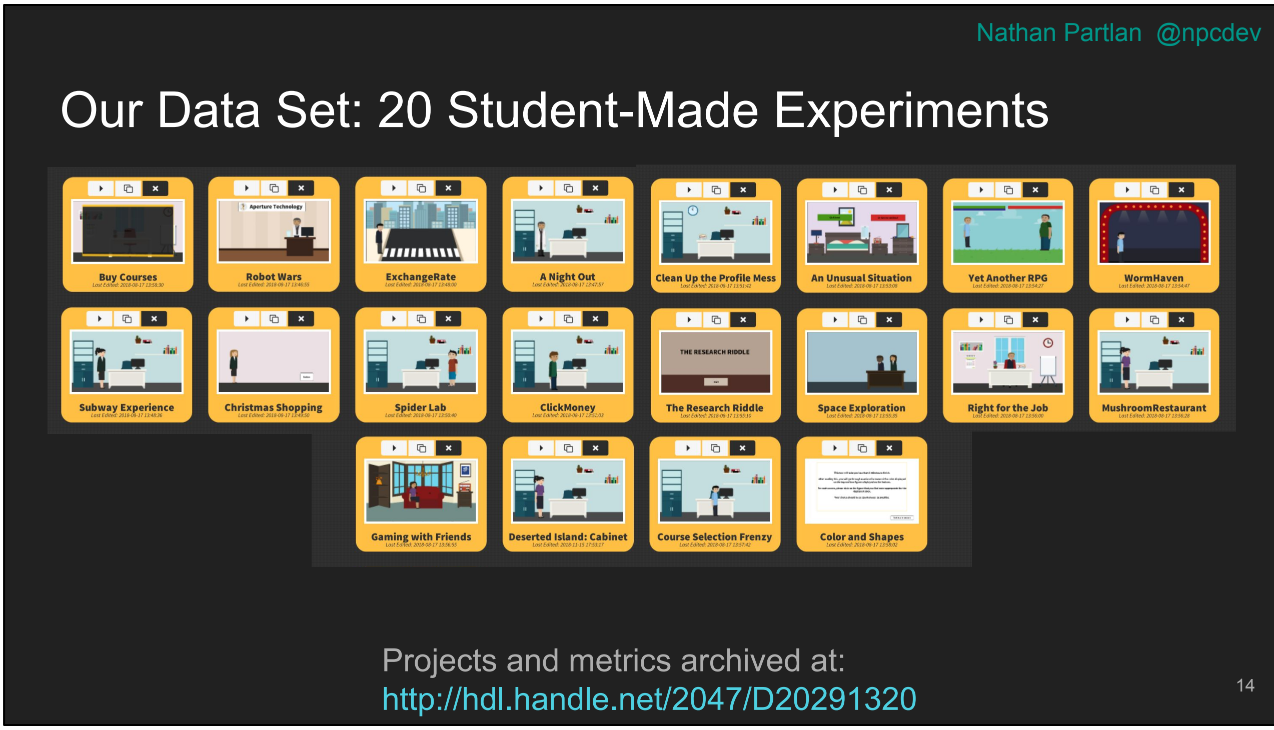
StudyCrafter is well-suited for developing visual-novel-like interactive narratives, but it can also support some additional gameplay mechanics, such as keyboard and mouse input, timed events, simple animation, and more.

Nathan Partlan @npcdev

Easy-to-Use Visual Scripting



StudyCrafter provides a visual scripting language, which makes it possible for designers with little or no programming experience to build functional scenarios.



In this case, we analyzed 20 scenarios built by game design students in a graduate-level research methods course.

They were new to research design, but had some experience designing games.

The students had only a few weeks to build their scenarios, but they produced a nice variety of games. Some had significant interactive narrative and characterization, some had little framing narrative and focused on the research question. Some provided only a few dialogue choices, a few had significant puzzles or even simple fighting mechanics.

You can find all the scenarios in the Northeastern Library archive, along with our metrics about them, if you want to explore them after the talk.

Nathan Partlan @npcdev

Example: “Deserted Island Cabinet Mystery”

One fairly representative example of a student project from this data set is "Deserted Island Cabinet Mystery." We will use this scenario as an example for some of our analysis later in this talk.

In it, the player is stuck on a deserted island.

They must repeatedly choose between sets of two items from a mysterious cabinet on the island, with each drawer locking after they have chosen a single item from it.

The scenario incorporates some framing narrative, a bit of animation and feedback for player choices, and has two "conditions," testing whether players will respond differently to two different button colors when selecting items.

Nathan Partlan @npcdev

Representation

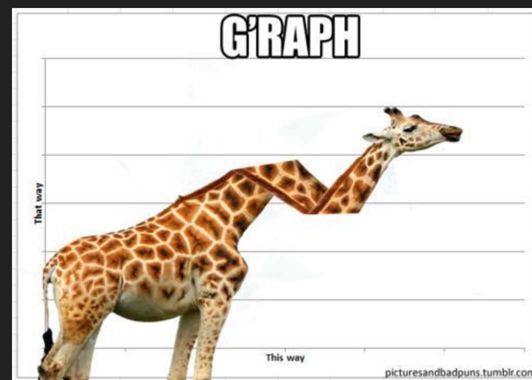


Image: http://www.graphgraph.com/tag/giraffe/

So how did we automatically analyze these scenarios? Through the clever use of graphs.

Image: http://www.graphgraph.com/tag/giraffe/

Graphs

Scene

Layout

Script

Interaction

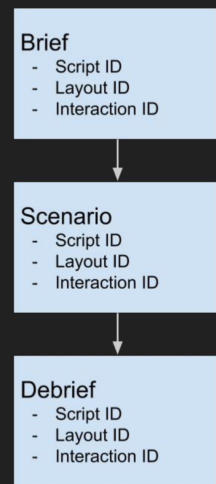
We created a multi-graph representation, with four different but related types of graph.

Image Credit: Ingo Arndt - mediadrumworld.com

Scene Map

Represents possible transitions between scenes

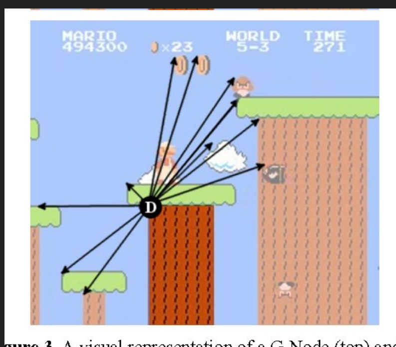
Points to IDs of other graphs for that scene

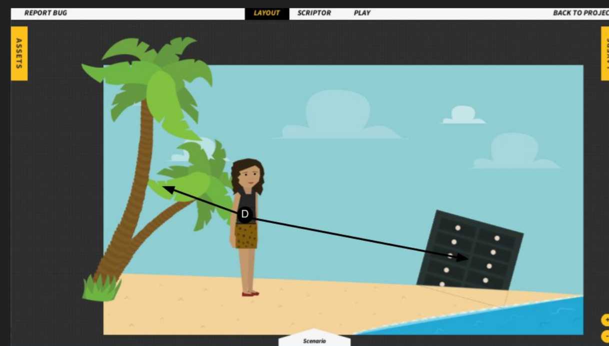


The scene map for “Deserted Island Cabinet Mystery”

The scene graph provides information about the number and transitions between scenes in the scenario. Each scene is linked by IDs to its other graphs, which were calculated per scene.

Layout Graph

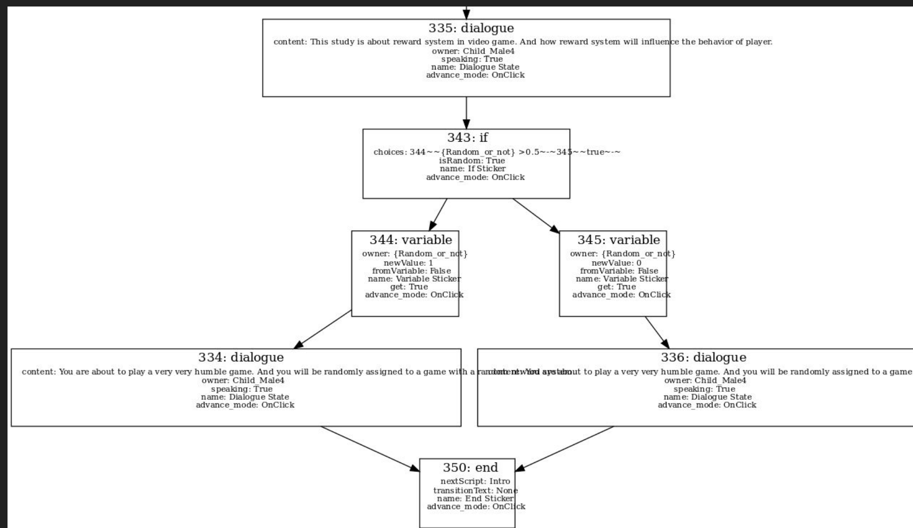




Guzdial, Matthew, and Mark Riedl. 2016. “Toward Game Level Generation from Gameplay Videos.” *ArXiv:1602.07721 [Cs]*, Procedural Content Generation Workshop 2015, February. http://arxiv.org/abs/1602.07721.

The Layout graph describes the visual design of the scene. It is inspired by the graphs used for analyzing and generating Mario levels by Guzdial and Riedl. Each node represents a single object in the scene, and edges represent the distances and directions to each other object.

Script Graph

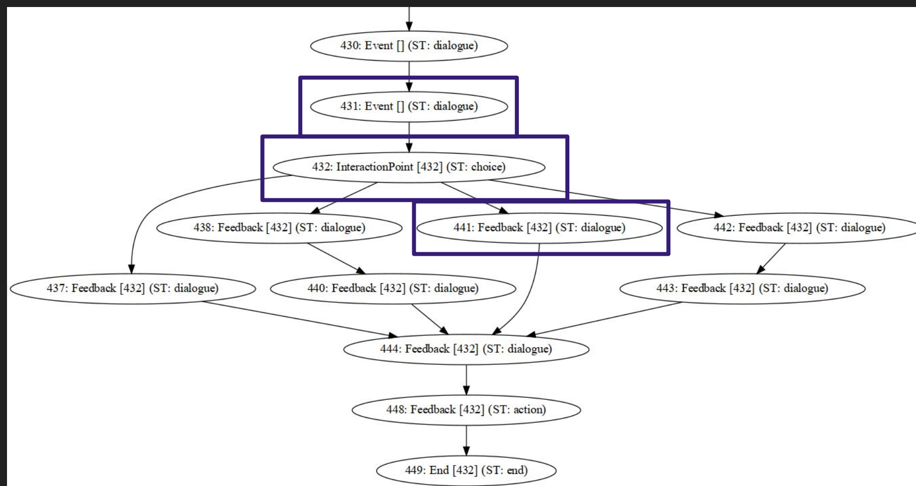


A partial script graph for a particular scene.

The script graph is based on the visual scripting language in StudyCrafter, which already has a graph-like structure. We simply formalize that, creating a true directed graph. Each node contains all the information from the original script, stored in a dynamically defined set of properties.

Nathan Partlan @npcdev

Interaction Map



A partial interaction map for a particular scene.

The interaction map is built automatically from the script graph. By analyzing how players can move through the scenario, we implement the theory developed by my co-authors Elin Carstensdottir and Magy Seif El-Nasr.

The graph contains three types of nodes: event nodes, which are things that happen regardless of player input; interaction points, where players provide input or make a choice; and feedback nodes, which are reactions to players' actions.

The edges represent possible flows of player interaction through the scenario.

Nathan Partlan @npcdev

Metrics

Image: http://www.radiolab.org/story/kg/

Now that we understand the representation, let's talk about what we can learn from it.

Image: http://www.radiolab.org/story/kg/

Playthrough-Based Metrics

Length/Duration: # of actions, amount of time per playthrough

Intra/Global Diversity: # of unique actions within one/in all playthroughs

Renewal Rate: % of new content in each playthrough

Choice Range: how many options does the player have at each choice point?

Choice Frequency: how often does the player make choices?

Choice Variability: the average ratio of new vs. repeated options

Szilas, Nicolas, and Ioana Ilea. 2014. "Objective Metrics for Interactive Narrative." In *Interactive Storytelling*, 91–102. Lecture Notes in Computer Science. Springer, Cham.

First, we can start with some prior work.

Most interactive narrative analysis is qualitative and difficult to automate.

Szilas and Ilea, however, have been working to build concrete, measurable metrics for playthrough-based analysis of interactive narratives. By calculating statistics on the length, action and choice variability and range, and more, they began to search for metrics that would provide helpful information for designers.

We incorporate their metrics in our work.

Static

Playthrough-Based

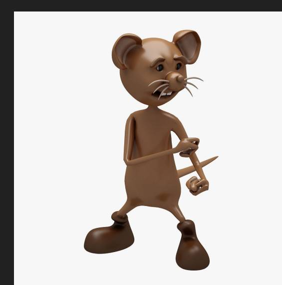


Image: https://www.cgstudio.com/3d-model/cartoon-mouse-01-02-rigged-t-pose-78975

We implemented two ways of calculating metrics: using static graph analysis techniques, and playthrough-based metrics.

The latter use the interaction map to generate "random walkthroughs" of the scenario, providing us a way of implementing Szilas and Ilea's metrics. Our walkthroughs are randomized, so a better approximation might be possible through more advanced automated playtesting AI. We leave that to future work.

Image:
https://www.cgstudio.com/3d-model/cartoon-mouse-01-02-rigged-t-pose-78975

Factors

Narrative Structure Complexity

Interactive Affordances

Choice Range

Image: http://www.funnyanimalsite.com/pictures/A_Stack_Of_Cats.htm

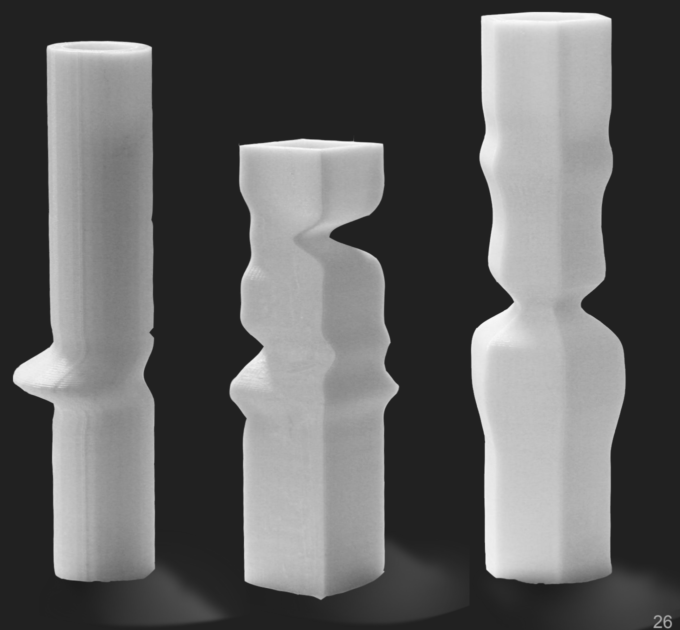
Our metrics are organized into three factors, each capturing a different aspect of how the scenario is structured and how it interacts with players.

Image: http://www.funnyanimalsite.com/pictures/A_Stack_Of_Cats.htm

Nathan Partlan @npcdev

Narrative Structure Complexity

The size and complication of the scenario.



Images: Britton Horn et. al., Visual Information Vases. http://viv.ccs.neu.edu/

Narrative Structure Complexity measures the overall size and complication of the scenario, in terms of script, visual, and interaction complexity.

Nathan Partlan @npcdev

Action Space

The number of choices available to the player, and their impact on the story.

Image: https://marswillsendnomore.wordpress.com/2012/07/27/space-action-mission-into-time/

Action Space represents the degree of choice available to the player, at each point and throughout the scenario.

Image:
https://marswillsendnomore.wordpress.com/2012/07/27/space-action-mission-into-time/

Interactive Affordances

The frequency and types of player actions, and the scenario’s responses to those actions.

Image: https://www.flickr.com/photos/b0xman/4437279497 - James Box

The Interactive Affordances factor is about how the player interacts with the scenario, how often, and what the scenario's responses to that interaction are. How much feedback does the player get when they take action?

Image credit: https://www.flickr.com/photos/b0xman/4437279497 - James Box

Results

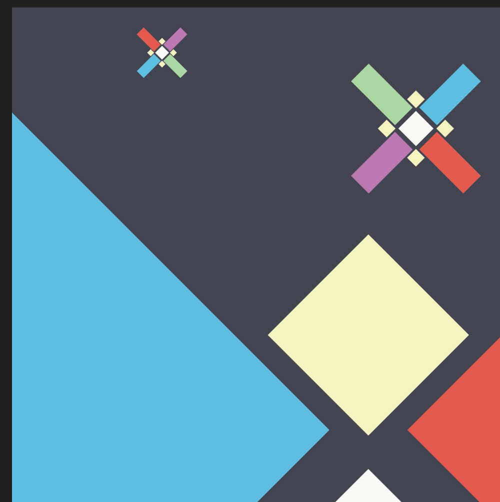
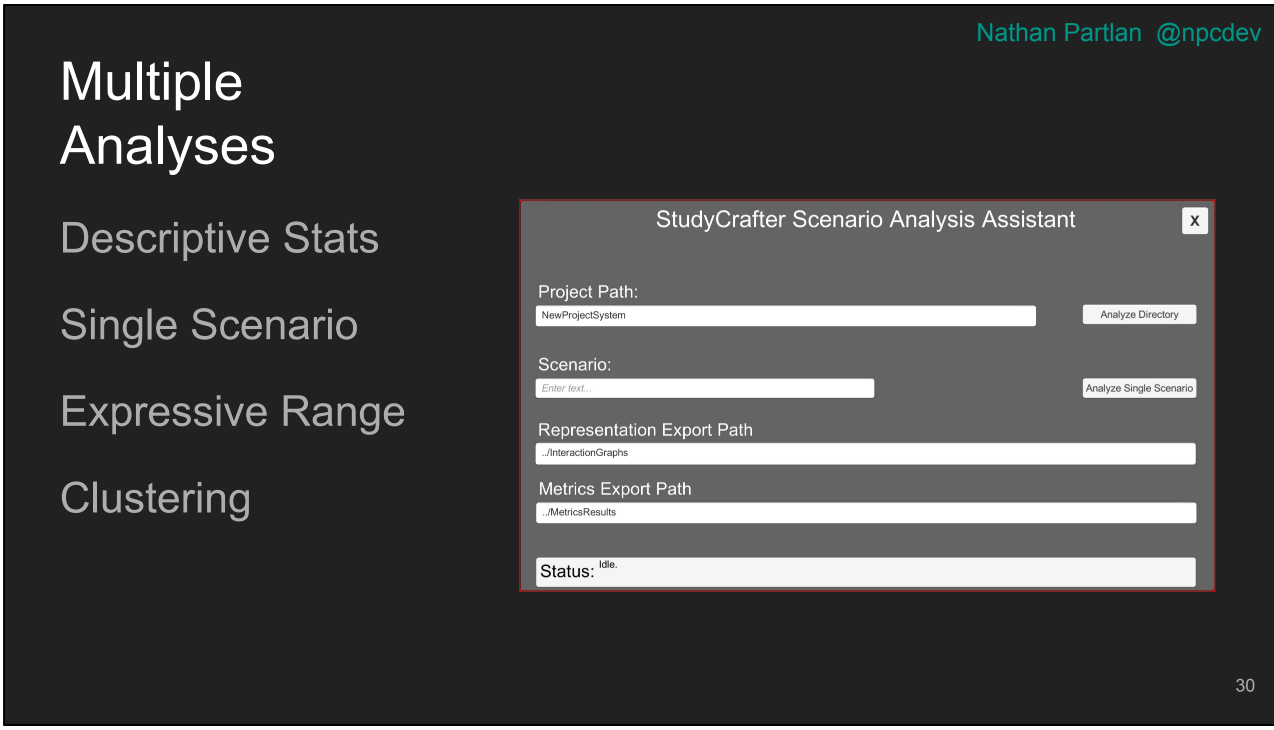
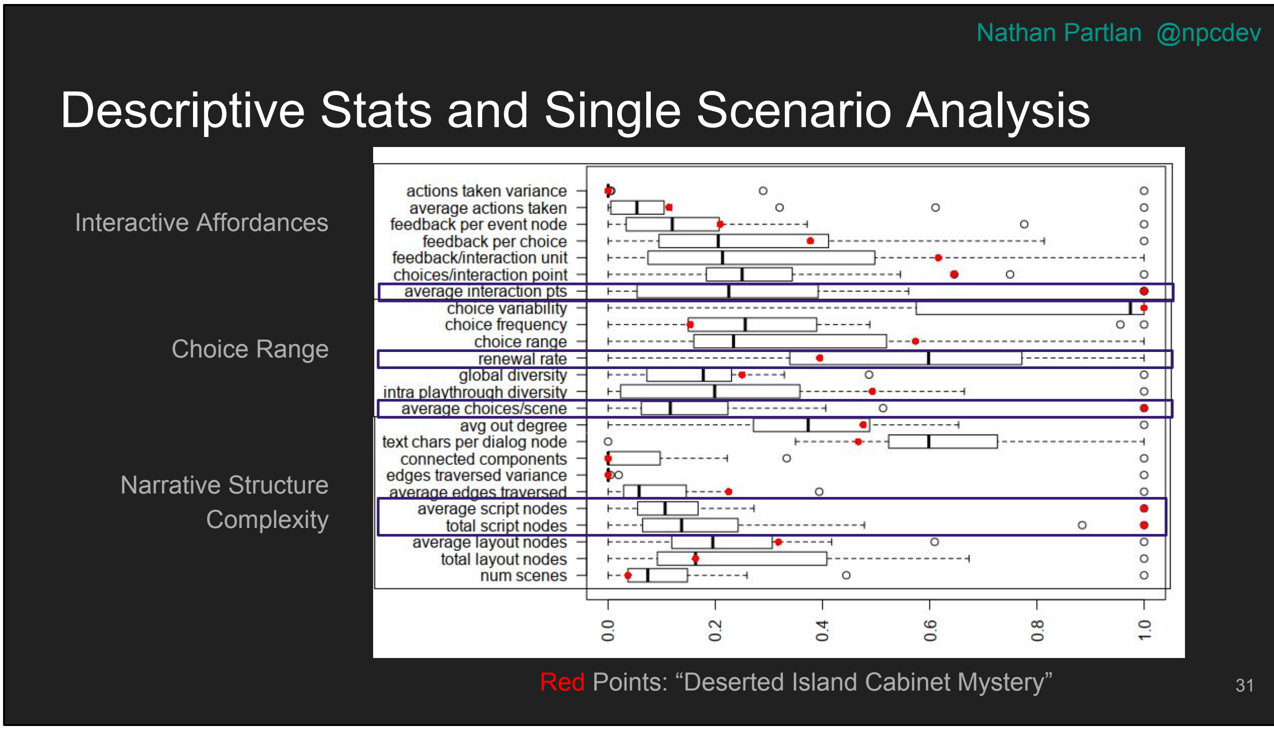


Image: Gillian M. Smith, “Quilt Design a Day.” http://sokath.com/main/blog/2015/05/23/quilt-design-a-day/

We applied these metrics to the 20 scenarios in our data set.



We used several different forms of analysis to glean insights about their design.



One way to look at the metrics is through descriptive statistics. In this plot, you can see box-and-whisker plots of each metric (normalized). In this view, we can see an overview of how a single scenario (specifically, Deserted Island Cabinet Mystery, our earlier example) compares to others.

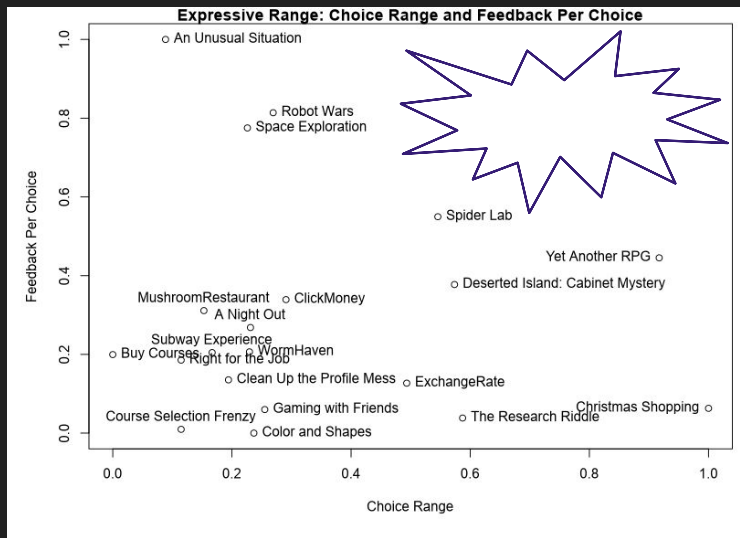
For instance, this scenario has an extremely high number of interaction points, and a correspondingly high number of choices.

It also appears to be high in other complexity metrics related to the number of script nodes. This scenario must have quite a bit of code to run, but it provides a lot of choices to the player.

However, it has a low renewal rate, which means that most of the content is similar during random playthroughs. This may be a sign that the choices do not lead to highly-varied paths through the scenario, and that each playthrough may feel somewhat similar despite the large number of choices available.

Based on this low renewal rate, we might ask the designer whether they could reduce the choice complexity of their scenario, but provide more meaningful changes in the story based on the player's choices, focusing their design and enabling more meaningful choice.

Comparisons



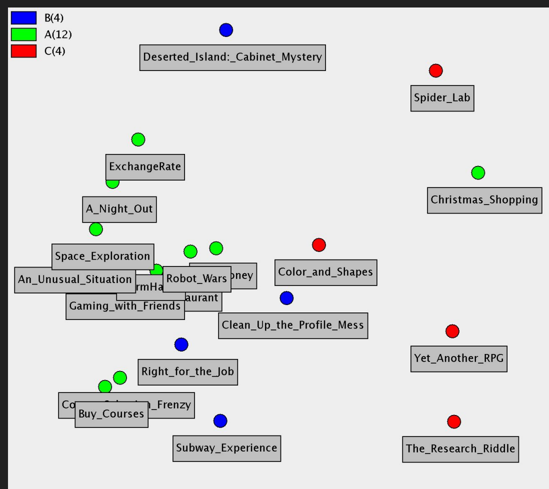
Another useful method for comparing scenarios is "Expressive Range," based on work by Gillian Smith et. al. Usually used for analyzing the range of output for a generative method, here we use it instead to understand the range of human-designed scenarios.

We select two metrics, and visualize the expressive range of our collection of scenarios in those.

In this case, we see that there is a tradeoff between choice range and feedback per choice; no scenario is high in both!

This makes sense, because with high choice range, there would be many options to write specific content for at each choice point, leading to a higher authoring burden for providing meaningful feedback for each.

Clustering



Action Space

In addition to individual scenario analysis and expressive range, we used clustering to understand groupings of scenarios. Using K-Medoids, we clustered scenarios on each of our metric factors.

Here we see the clusters for “action space.” In this case, the three clusters fairly effectively separate the scenarios by how they provide the player with choices.

Deserted Island: Cabinet Mystery is in the blue cluster with several other scenarios that repeatedly provide choices to the player, each with multiple options, but that do not significantly change the rest of the scenario based on those choices - the options are mostly superficial. The red scenarios, on the other hand, are those with significant puzzles or combat gameplay, where choices can significantly change how the game progresses and how quickly the player reaches the end.

Future Work

Image: https://twitter.com/powerofcats/status/290392045162471424

We have several future avenues for improvement and further research.

Image: https://twitter.com/powerofcats/status/290392045162471424

Revising the Representation and Metrics

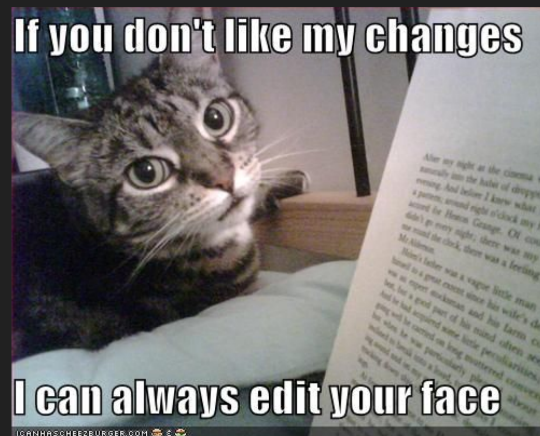


Image: https://www.youtube.com/watch?v=gZ6J2ioMM0A

We can improve and revise our representation and metrics - they do not yet fully capture everything we can about the scenarios. We can also use more motivated AI playthroughs based on automated playtesting methods to improve our metric calculations.

Image: https://www.youtube.com/watch?v=gZ6J2ioMM0A

Nathan Partlan @npcdev

Creativity Support:
Analysis and Prompting

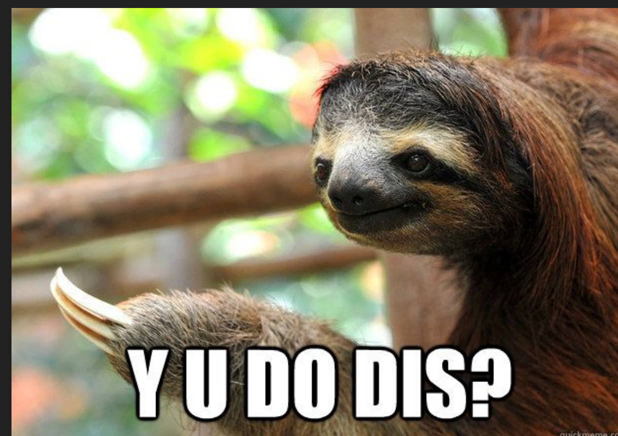
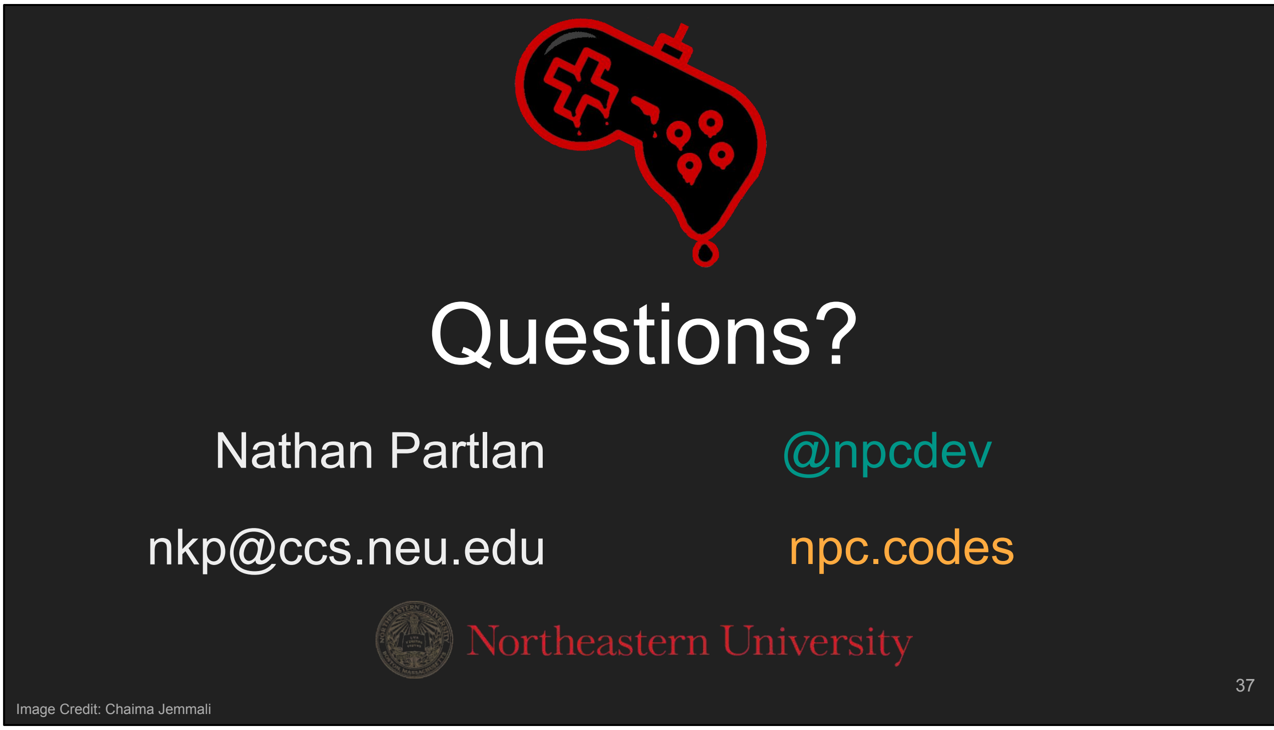


Image: http://www.youtube.com/watch?v=q6Kffa4ofG4

Our next major step is to determine how we can use this representation and metrics to help inform a full creativity support system, achieving our original goals of mixed-initiative AI. We will be developing questions, visualizations, and other support tools based on this work.

Image: http://www.youtube.com/watch?v=q6Kffa4ofG4



Thank you! What questions do you have?

Image Credit: Chaima Jemmali

Nathan Partlan @npcdev

References

Projects and metrics archived at:
http://hdl.handle.net/2047/D20291320

Ashwell, Sam Kabo. 2015. "Standard Patterns in Choice-Based Games." These Heterogenous Tasks (blog). January 27, 2015. https://heterogenoustasks.wordpress.com/2015/01/26/standard-patterns-in-choice-based-games/.

Baldwin, Alexander, Steve Dahlskog, Jose M. Font, and Johan Holmberg. 2017. "Towards Pattern-Based Mixed-Initiative Dungeon Generation." In Proceedings of the 12th International Conference on the Foundations of Digital Games, 74:1–74:10. FDG '17. New York, NY, USA: ACM.

Bernstein, Mark. 1998. "Patterns of Hypertext." In Proceedings of the Ninth ACM Conference on Hypertext and Hypermedia: Links, Objects, Time and Space, 21–29. HYPERTEXT '98. New York, NY, USA: ACM.

Carstensdóttir, Elín, Erica Kleinman, and Magy Seif El-Nasr. 2017. "Towards an Interaction Model for Interactive Narratives." ICIDS 2017.

Csikszentmihalyi, Mihaly. 1990. Flow: The Psychology of Optimal Performance. New York: Harper and Row.

Guzdial, Matthew, and Mark Riedl. 2016. "Toward Game Level Generation from Gameplay Videos." ArXiv:1602.07721 [Cs], Procedural Content Generation Workshop 2015, February.

Liapis, Antonios, Georgios N. Yannakakis, and Julian Togelius. n.d. "Sentient Sketchbook: Computer-Aided Game Level Authoring." In In Proceedings of ACM Conference on Foundations of Digital Games, 2013.

Millard, David E., Charlie Hargood, Michael O. Jewell, and Mark J. Weal. 2013. "Canyons, Deltas and Plains: Towards a Unified Sculptural Model of Location-Based Hypertext." In Proceedings of the 24th ACM Conference on Hypertext and Social Media, 109–118. ACM.

Ryan, James. 2017. "Popular Visual Depictions of Early Generative Systems." In Seeds vol. 2.

Short, Emily. 2016. "Small-Scale Structures in CYOA." Emily Short's Interactive Storytelling (blog). November 5, 2016. https://emshort.blog/2016/11/05/small-scale-structures-in-cyoa/.

Smith, Gillian, and Jim Whitehead. 2010. "Analyzing the Expressive Range of a Level Generator." In Proceedings of the 2010 Workshop on Procedural Content Generation in Games, 4:1–4:7. PCGames '10. New York, NY, USA: ACM.

Smith, Gillian, Jim Whitehead, and Michael Mateas. 2010. "Tanagra: A Mixed-Initiative Level Design Tool." In *Proceedings of the Fifth International Conference on the Foundations of Digital Games*, 209–216. FDG '10. New York, NY, USA: ACM.

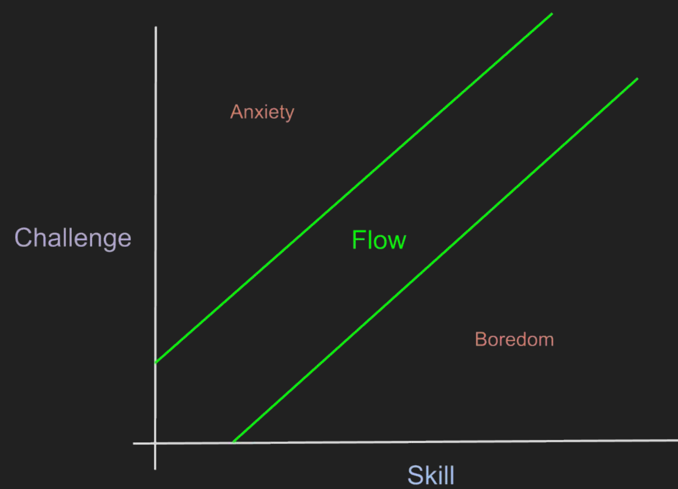
Szilas, Nicolas, and Ioana Ilea. 2014. "Objective Metrics for Interactive Narrative." In Interactive Storytelling, 91–102. Lecture Notes in Computer Science. Springer, Cham.

Thanks To:
The StudyCrafter team, Dr. Camillia Matuk and Dr. Steven Sutherland
The StudyCrafter project has been supported by:
Northeastern University, NSF (IIS-1736185), and DARPA (D16AP0011)

Correlations with **IRIS** Questionnaire

Choice Variability: correlated with **Flow** and **Positive Emotional State**

Choice Range: correlated with **Negative Emotional State** (over-saturation?)



Szilas, Nicolas, and Ioana Ilea. 2014. "Objective Metrics for Interactive Narrative." In *Interactive Storytelling*, 91–102. Lecture Notes in Computer Science. Springer, Cham.

(adapted from) Csikszentmihalyi, Mihaly. 1990. *Flow: The Psychology of Optimal Performance*. New York: Harper and Row.

They tested their metrics against the IRIS survey, which measures player responses to interactive narrative.

They found some correlations with two of these metrics, specifically in terms of "flow" and "positive emotional response."

We incorporate and expand on these metrics in our work.