Aesthetic Considerations for Automated Platformer Design

Michael Cook, Simon Colton and Alison Pease
Computational Creativity Group
Department of Computing
Imperial College, London
ccg.doc.ic.ac.uk

Abstract

We describe ANGELINA3, a system that can automatically develop games along a defined theme, by selecting appropriate multimedia content from a variety of sources and incorporating it into a game’s design. We discuss these capabilities in the context of the FACE model for assessing progress in the building of creative systems, and discuss how ANGELINA3 can be improved through further work.

The design of videogames is both a technical and an aesthetic task, and a holistic approach is necessary when constructing systems which aim to automate the process. Systems previously demonstrated as automated game designers have been shown to tackle, in a basic way, many of the technical tasks associated with game design including level creation and ruleset design, for both simple arcade-style games (Cook and Colton 2011a) and platform games (Cook and Colton 2012). However, in such systems the art, sound and theme are chosen by a human. This weakens the claim that these systems automate the process of game design.

Today, people play videogames for many reasons beyond simply the challenge they offer. Dan Pinchbeck’s experiment in narrative technique Dear Esther1 enjoyed 50,000 sales in its first week2, while Jenova Chen’s Flower3 has been used in a church in the UK as part of a service of worship, with one attendee describing the game as ‘spiritual’4. Automating the design of games that carry emotional weight or attempt to convey a complex meaning is a compelling research problem that lies at the intersection of game design theory and Computational Creativity, and is almost entirely unexplored.

ANGELINA, A Novel Game-Evolving Labrat I’ve Named ANGELINA, is a system for investigating the automation of simple videogame design. We describe here a first step for the latest version of the software, ANGELINA3, towards producing a system that not only takes on the technical task of game and level design, but also independently selects and arranges visual and aural media as part of the design process, to achieve a creative and artistic goal in the finished game. Our long term goal is to develop a fully automated creative videogame design system. This paper reports our progress towards this goal, in which we describe the third iteration of the ANGELINA3 system and employ the FACE model (Colton, Charnley, and Pease 2011) of evaluation from Computational Creativity to argue that ANGELINA3 is more creative than an earlier version of the software. We make the following contributions:

1. We describe an automated videogame design system, ANGELINA3, which is able to generate conceptual information gleaned from news articles, form aesthetic evaluations of a particular concept, invent example videogames which express these concepts, and generate its own framing information about its products and processes.

2. We demonstrate the use of evaluation criteria from Computational Creativity to game design systems, and use it to argue that our system has progressed in terms of creativity since a previously described version of the software.

The remainder of this paper is organised as follows: in the section titled Background we describe the structure of ANGELINA2 and extensions made in ANGELINA3; we then describe the modules that provide the system’s creative abilities; in the Example Games section we give examples of games produced by the system; we then evaluate ANGELINA3 as the system currently stands; in Related Work we outline some existing work in the area and its relation to ANGELINA3; finally we discuss future directions for the project to improve ANGELINA3’s creative abilities and independence as a designer.

Background

ANGELINA

First proposed in (Cook and Colton 2011a), ANGELINA1 is a co-operative co-evolutionary system that designs games iteratively by decomposing the design process into separate but interrelated design tasks. In (Cook and Colton 2012) we refer to these processes as species. In a co-operative co-evolutionary system, these species operate in a similar manner to standard evolutionary processes; they have a population, a fitness function, a procedure for crossover and so on. The primary difference comes in the evaluation of fitness for a candidate solution. In co-operative co-evolution, a
candidate solution is evaluated alongside the highest-fit ness members of all other species’ populations, and the fitness of the overall resulting system is measured as well as fitness of the single species on its own. This allows fitness func tions to take into account both individual performance as well as how well a solution co-operates with the solutions being produced by other species. Better co-operating indi viduals are preferentially selected, and over time solutions improve both on a species level and the inter-species level.

In (Cook and Colton 2011a), ANGELINA2 used three species - Maps, Layouts and Powersets. Maps defined pass able and impassable areas in the game world; Layouts de fined the placement and design of enemies, as well as the player start and game exit; Powersets defined a set of powerup items which enhanced the player’s abilities and en abled them to complete levels. For further details on these species, their representation within the system and their evalu ation via fitness functions, see (Cook and Colton 2012).

As with the previous version, games produced by ANGELINA3 are 2D platform games based on the Metroid vania genre: players are tasked with finding a goal some where in the game; initially the player’s access to regions of the game are restricted by their abilities, such as the height they jump to. By collecting powerups, the player’s abilities change and new areas of the world become accessible. Some simple combat is included, although these only serve as temporary hindrances as the player cannot die. Further description of the games can be found in (Cook and Colton 2012).

The FACE Model

The evaluation of creativity in systems is an active area of research (see (Jordanous 2012, Chapter 2) for an overview), and only two frameworks have achieved take-up by the community: Ritchie’s artefact-based criteria (Ritchie 2001) and Colton’s creative tripod (Colton 2008). In the creative tripod, Colton argues that if the system exhibits skill, appreciation and imagination then it will be perceived as creative.

The recent FACE model (Colton, Charnley, and Pease 2011) extends the creative tripod by breaking down the creative act into constituent parts and providing computational interpretations of each aspect, inspired by the psychology of human creativity and analyses of acts of human creativity ((Pease and Colton 2011) for details). It breaks down creative acts into 8 types of generative acts producing:

\[ F^g: \text{a method for generating framing information} \]
\[ F^p: \text{an item of framing information for } A/C/E^{p/g} \]
\[ A^g: \text{a method for generating aesthetic measures} \]
\[ A^p: \text{an aesthetic measure for process or product} \]
\[ C^g: \text{a concept} \]
\[ C^p: \text{a method for generating concepts} \]
\[ E^g: \text{an expression of a concept} \]
\[ E^p: \text{a method for generating expressions of a concept} \]

Creative episodes are then expressed in terms of tuples of at least one of these types of generative acts (not necessarily all). For instance, the creation of the notion of prime numbers involved the invention of the concept (prime number) (C^g); finding examples of the concept (E^g), and inventing ways of generating further primes (E^p).

The FACE model affords the comparison of two creative systems, which may be versions of the same software. In particular, under a straightforward cumulative approach, a system which performs the creative act comprising three generative acts: \(<A^g; C^g; E^g>\) might be seen as more creative than one which only performs creative acts \(<C^g; E^p>\).

Note that the FACE model does not take into account the quality of the artefacts produced. It is designed to gauge the progress of the system itself, and the authors acknowledge in (Colton and Wiggins 2012) that the quality of any generated artefacts may drop in line with initial increases in the creativity of the system. They liken this to the phenomena of latent heat in thermodynamics: “as the creative responsibility... increases, the value of its output does not (initially) increase, much as heat input to a substance on the boundary of state change does not increase temperature”.

Towards a fully automated creative videogame design system

In this section we briefly describe the additions to ANGELINAS’s co-operative co-evolutionary system that facilitate increased creativity. We then go on to describe the processes that allow the system to make creative decisions, obtain media resources, and create a themed game.

Creativity and Evolution

In order to integrate downloaded resources into the design of a game, we have added a fourth species to the co-operative co-evolutionary system, which evolves Artistic Direction objects. A single Artistic Direction (AD) is a set of ImagePlacement and SoundPlacement objects that define the positioning of media content within a game. ImagePlacements define co-ordinates for an image’s position in the game, as well as width and height parameters that define how the image is scaled. Images are invisible by default and fade into view when the player passes over them in the game. SoundPlacements define co-ordinates for a sound effect’s position in the game, as well as a range parameter that defines a region around the sound effect’s position. When a player enters this region, the sound effect begins to play.

Crossover of two AD solutions employs uniform crossover across the concatenated lists of Image- and SoundPlacements, while mutation randomly selects one or more Placement objects and randomly adjusts their co-ordinate values or other parameters. In order to evaluate a Placement object, we first ensure it is not overlapping with any other Placement objects, or overlapping with the edge of the game world. We also use data from the Map species to penalise ImagePlacement objects which overlap with game tiles (as this would obscure their view). ANGELINA3 generates reachability maps by simulating the player’s path around the game world, and this data is also used in the evaluation function to ensure that all Placement objects can be triggered by the player in a standard playthrough.

Media Acquisition and Use

Currently, the starting point for any execution of ANGELINA3 is the website of The Guardian newspaper,
inspired by a collage-generator described in (Krzeczkowska et al. 2010) that created image mashups using current news stories. ANGELINA3 reads the current top five news headlines, and ranks them as follows. Articles which feature tags ANGELINA3 has no record of seeing before are considered more interesting, but the system will also use a sentiment analysis technique to query Twitter about people whose names it has heard of before. If ANGELINA3 detects a shift in opinion about a person, that raises how interesting an article is, as described below.

Once ANGELINA3 has selected an article to use, it extracts the headline, the body text, and the set of tags which the Guardian has assigned to the article. Because tags summarise the article’s contents, they provide a useful shorthand for the topics the article covers. Once the system has collected this data from the article, it then proceeds through several media acquisition phases to obtain resources for use in the game’s design. These are outlined below.

**Country Detection**

ANGELINA3 identifies a word as a country by using the Wikipedia list of sovereign states. Once a country has been identified, ANGELINA3 uses another Wikipedia page to convert a country’s name into its adjectival form, C, which it uses to search Flickr using the term “C landscape”. It selects a result to be used as a background image for the game.

**Person Detection & Sentiment Analysis**

We consider a person notable if they have a Wikipedia page. Using this as a metric, ANGELINA3 can detect if a tag refers to a person by checking Wikipedia for the existence of a page about them. The system then attempts to gauge whether a person is liked or not by the general populace. This is done via a basic sentiment analysis of Twitter. For a person P, ANGELINA3 searches Twitter for popular tweets matching the search term “P is”. For each tweet, it collects the word directly following the search term into a set of words, Q, and then calculates an average emotional weight for the set Q using the AFINN database (a collection of 2477 words with hand-assigned valences) (Nielsen 2011). This average is then used to update a database of prevailing opinion that is persistent across all executions of ANGELINA3.

The sentiment and the collected data about a person is used in two ways. Firstly, in the event that no country is found in the story tags, ANGELINA3 can use a person’s nationality as a basis for a background image search. Secondly, ANGELINA3 will select images of this person for integration into the game. We employ an augmented search technique as described in (Cook and Colton 2011b) to emphasise an emotion based on the sentiments recorded. If negative sentiments were recorded, the search was augmented with ‘angry’; if positive, the search was augmented with ‘happy’.

The intention here is to present an image of the person likely to elicit the sentiment popularly held about them - seeing an angry face is likely to present the person negatively. Figure 1 shows a sample augmented search.

**General Tag Use**

If a tag refers to neither a country nor a person, ANGELINA3 uses it as a basis for searching online image and sound databases for relevant media to use in the game. Image searches were performed using Google Images, while the FreeSound database was used for sound effects.

ANGELINA3 can preferentially select tags as being the focus of a game, which leads to the inclusion of more image and sound resources bearing those tags. The software has different methods for choosing a focus - it can prioritise the inclusion of tags which appear in the headline, tags which appear frequently in the body text, or tags which are less common words in general English. This emphasis on certain tags changes the balance of a game’s aesthetic by exposing the player to far more images or sounds of a certain kind.

**Title Generation**

ANGELINA3 has two methods it can use to generate a title for a game. The first approach is to attempt to generate a pun based on one of the tags attached to the article. For each tag, the system queries the RhymeZone and WikiRhymer databases for a list of perfect rhymes for the tag. It then uses the list to search four corpora of pop culture phrases: the Guardian’s Top100 Songs To Heart Before You Die; the NY Times’ Top 1000 Films; Tony Mott’s book 1001 Videogames You Must Play Before You Die; and WikiRhymer’s own database of common phrases or proverbs. If ANGELINA3 finds any matches, it substitutes the original tag for the rhyming word, which it adds to a list of possibilities and randomly selects one after completing its search.

If no rhyme matches are found, it uses an alternative approach that employs the TextRank algorithm outlined in (Mihalcea and Tarau 2004). By concatenating the headline and body text and performing a TextRank search on it, ANGELINA3 receives a set of phrases and words ordered by importance as assessed by TextRank. Using a method similar to that described in (Colton, Goodwin, and Veale 2012), we analyse the results of TextRank using the Kilgarriff database of word frequencies to assess how common each word is in the English language. Through initial experimentation, we found that by ordering the TextRank results based

---

5http://www.freesound.org
6http://www.rhymezone.com
7http://wikirhymer.com/
8http://www.kilgarriff.co.uk/
I was reading the Guardian website today when I came across a story titled “Obama to urge Afghan president Karzai to push for Taliban settlement”. It interested me because I’d read the other articles that day, and I prefer reading new things for inspiration. I looked for images of United States landscape for the background because it was mentioned in the article. I also wanted to include some of the important people from the article. For example, I looked for photographs of Barack Obama. I searched for happy photos of the person because I like them. I also focused on Afghanistan because it was mentioned in the article a lot.

Figure 2: An excerpt from the commentary for the game Hot NATO

Music Selection
ANGELINA_3 uses a collection of Creative Commons music by Kevin Macleod⁹. By running the body text of the Guardian article through the AFINN database, the system can gauge an average tone of the article. If the tone is positive, it selects a piece of music that is upbeat or bright. If the tone is negative, it chooses a suspenseful piece. Selections are made at random from tracks tagged with that emotion.

Commentary Generation
After the generation of a game ANGELINA_3 is able to create a commentary describing the creative process, inspired by the commentaries generated in (Colton, Goodwin, and Veale 2012). During the production of the game ANGELINA_3 records decisions as well as the justifications for them, logging them for synthesis into a templated commentary. The system then replaces segments of the commentary template with the appropriate contextual information.

The commentaries mention both static features of the creative process, such as the headline of the story, as well as decisions made by the system such as the reasons for selecting an article or tags which were focused on in depth by ANGELINA_3. Figure 2 shows an example commentary.

Example Games
In this section we give two examples of games produced by the system, selected by hand from a week of daily executions of ANGELINA_3.

Sex, Lies and Rape
On May 8th 2012 nine men were convicted in the UK of sexually exploiting young girls in Greater Manchester. The Guardian reported on the story under the headline Rochdale gang found guilty of sexually exploiting girls. ANGELINA_3 retrieved the article, along with the tags Crime, Police, Child Protection, Children and Social Care, and created a game called Sex, Lies and Rape. It can be played online at http://www.gamesbyangelina.org/aiide/slar.

Because no country is explicitly mentioned in the headline or tags, and no people are named, ANGELINA_3 retrieves a generic landscape image for the background of the game. A suspenseful piece of music was selected because the overall tone of the article is judged to be negative. Images were selected based on the tags, including a cartoon of a criminal; a drawing of two parents protecting a child; a photograph of a young girl with the text ‘Because nothing matters more’ underneath it; and a painting by Titian depicting the rape of Lucretia. There is one sound effect that can be triggered by the player - a children’s song being sung in Greenlandic. Figure 3 shows a screenshot from this game.

The Conservation of Emily
On May 10th 2012 Lord Mandelson, a peer in the UK’s House of Lords, admitted that he was working for a multinational firm accused of illegal logging activities. The Guardian reported on the story under the headline Lord Mandelson confirms he is advising company accused of illegal logging. ANGELINA_3 retrieved the article, along with the tags Peter Mandelson, Greenpeace, Activism, Deforestation, Endangered Habitats, Endangered Species, Conservation, Forests and Animals, and created a game called The Conservation of Emily. It can be played online at

---

⁹http://www.incompetech.org

No country is mentioned in the tags, however Peter Mandelson is identified as being English, so a picture of the English countryside is retrieved and used as background. The article is assessed as being negative in tone, so a suspenseful piece of music is selected. Ambient rainforest sound effects and birds singing can be found throughout the level, as well as a man screaming. Inset images retrieved for the story’s tags include some small animals; a photograph of Peter Mandelson; a collage of animals with the words ‘Help Us’ in the centre; and a placard reading ‘Oil Fuels War’. Figure 4 shows a screenshot from this game.

**Evaluation - The FACE Model**

We have evaluated our system with respect to the FACE model introduced in the background section. We break this down into four parts and discuss ANGELINA’s functionality with respect to generating particular types of product. Evaluating ANGELINA under such a model provides a formalised manner in which to compare different approaches to automated game design and allows us to better chart future directions for the system’s development.

**Concept** ANGELINA produces videogames which attempt to convey a sentiment about a particular person, which we represent as a concept under the FACE model, of which a videogame is an expression. ANGELINA contributes to this concept by acquiring information about particular people, evaluating sentiments and using them to inform the design process, which represents a \( C^3 \) act.

**Examples** By designing games that follow basic tenets of Metroidvania design, as described in (Cook and Colton 2012), as well as producing games that feature content directly inspired by a current news event, ANGELINA demonstrates an ability to produce basic expressions of concepts \( E^3 \) such as platformer videogames with a consistent theme and mood.

**Aesthetic** The aesthetic judgement \( A^3 \) of whether a game or a set of media convey a sentiment about a person is used in the media acquisition stage of ANGELINA. Although it is not integrated fully into the evolutionary design process, it contributes to the production of the final game by helping evaluate the media that are selected for inclusion in the game’s design. We discuss the expansion of aesthetic judgements in ANGELINA as part of future work.

**Framing information** ANGELINA can generate framing information \( F^3 \) in the form of commentaries and titles that reference both popular culture and the news articles that served as inspiration, as well as justifying decisions that affected the outcome of the generative process. In Figure 2 the commentary states that ‘I searched for happy photos of the person because I like them,’ which shows the system can justify design decisions with reference to a particular concept.

**Discussion** ANGELINA is the result of our attempt to build a system that can make decisions, implement them in an artefact, and justify them after the fact. In terms of the FACE model, ANGELINA has functionality in some aspect of four generative acts on the product level: \( F^3; A^3; C^3; E^3 \). In terms of the cumulative approach described in the Background section, we can compare ANGELINA to the version of the software described in (Cook and Colton 2012), ANGELINA, which is only capable of generating expressions of the Metroidvania genre in the form of playable games \( E^3 \). ANGELINA is unable to make decisions or alterations to its design process, nor is it able to produce information framing the process after the fact, meaning the system neither employs the use of aesthetic values nor generates framing information whilst designing a game. Thus, the creative act undertaken by ANGELINA can be expressed solely by the tuple \( E^3 \) and ANGELINA is therefore an advance on this work.

Note that ANGELINA does not invent any of its own processes (these are human-developed), suggesting areas for further work.

**Related Work**

In (Treanor et al. 2012) the authors describe Game-O-Matic, a system for assisting in the production of news games; games which are designed to highlight a current news story, often created in conjunction with journalists to complement traditional journalism. A human describes relationships between two or more real-world concepts (such as protesters and police) and the tool attempts to design a game in which the mechanics of gameplay reflect these relationships. Although both (Treanor et al. 2012) and ANGELINA share news stories as their subject matter, the aims of the research projects are quite divergent. Game-O-Matic’s intention is to provide a tool for assistive game design, whereas our aim with ANGELINA is to create software that can design independently about general themes or topic areas. We chose news stories as source material here due to the richness of the data associated with them in the form of current social discourse and available multimedia.

Game-O-Matic uses a human-defined set of verbs and mechanics in order to construct possible gameplay scenarios, but in doing so designs games which convey meaning through their mechanics. In one example in (Treanor et al. 2012), the player plays as a protester and must avoid the police entities that are on-screen. ANGELINA’s theming is far more aesthetic at this point, and does not affect the way the player interacts with the game on a mechanical level. This is discussed in further work as an area for development.

In (Nelson and Mateas 2007), the authors describe a system that generates simple games based on keyword nouns and verbs, such as shoot and pheasant. The system employs the WordNet corpus to link nouns and verbs to a set of pre-known game mechanics and nouns, from which it produces a small game. This gives the system a lot of flexibility, and also allows the games produced to have some visual components, such as a picture of a bird for the keyword ‘pheasant’. However, the games never increase in complexity beyond a simple minigame, and the creative variation in the games is restricted to visual and mechanical components only.
Further Work

There are many areas of expansion for ANGELINA, both in technical terms as well as the creative skills used in design. Mechanically, the ability to understand directionality and flow would enable higher-level planning of game designs, where the evolutionary system can take into account the order in which the player is exposed to information, and what directions they are likely to move in. This opens up the possibility that ANGELINA would be able to convey a narrative of events through the ordered presentation of content.

Considering the FACE evaluation above, a key area of growth for ANGELINA is into the process space, rather than the generative steps of product construction. One example of such growth could be to give ANGELINA metalevel control over its own design process, by allowing it to alter its fitness functions prior to evolution. This would allow the system to develop its own aesthetic measures for the design process, strengthening its performance creatively.

Conclusion

We have introduced a new set of capabilities for the automated game design system ANGELINA, which demonstrate an ability to creatively design games around a theme, using a variety of multimedia resources. We have evaluated ANGELINA’s current ability under the FACE model, and used it to point towards future developments for the system, as well as showing progress from the previous version of the software. In addition, we claim that ANGELINA is the first game design system that performs under all four aspects of the FACE model in a generative capacity.

The production of framing information and the application of aesthetics to creative processes is integral to the creative autonomy of a system (Pease, Charnley, and Colton 2012), as well as contributing towards the perception of the system as being creative (Colton 2008). While the games produced by the system may not be remarkable, the underlying systems are creatively broader than any previous version, and we hope to continue this improvement in future.

Applying ideas from Computational Creativity to game design opens up new avenues for development and evaluation of automated systems, as well as providing a new perspective on the creative processes involved. By giving more creative responsibility to our systems we hope to assist them in developing a new wave of meaningful videogames.

Acknowledgements

The authors would like to thank the reviewers for their comments which improved the quality of many aspects of the paper. Thanks also to Phillipe Pasquier and Antonios Liapis for useful discussions and suggestions.

References
