

CESAR RUBEN FRANCISCO GENNARO CAMPOS

**MIXED-INITIATIVE TOOL TO SPEED UP
CONTENT CREATION IN PHYSICS-BASED
GAMES**

Dissertação apresentada a Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Ciência da Computação, para obtenção do título de *Magister Scientiae*.

VIÇOSA
MINAS GERAIS - BRASIL
2017

CESAR RUBEN FRANCISCO GENNARO CAMPOS

**MIXED-INITIATIVE TOOL TO SPEED UP
CONTENT CREATION IN PHYSICS-BASED
GAMES**

Dissertation presented to the Graduate Program in Ciência da Computação of the Universidade Federal de Viçosa in partial fulfillment of the requirements for the degree of Master in Ciência da Computação.

VIÇOSA
MINAS GERAIS - BRASIL
July 2017

© 2017, Cesar Ruben Francisco Gennaro Campos.
Todos os direitos reservados.

Gennaro Campos, Cesar Ruben Francisco
d0000 Mixed-Initiative Tool to Speed Up Content
Creation in Physics-Based Games / Cesar Ruben
Francisco Gennaro Campos. — VIÇOSA

MINAS

GERAIS

-

BRASIL,
2017

ix, 39 f. : il. ; 29cm

Dissertação (mestrado) — Universidade Federal de
Viçosa

Orientador: Levi Henrique Santana de Lelis

1. Mixed-Initiative tool. 2. procedural content
generation. 3. physics-based games.. I. Title.

CDU 0000

[Folha de Aprovação]

I dedicate this work to my beloved parents César and Carmen (In memoriam).

Acknowledgments

First of all, I thank to God because  is the light, strength, fortress and wisdom that gives sense to my life. For all that He has accomplished for me.

To my adviser, Prof. Levi, for his dedication, patience and teachings. For contributing to my professional growth and for being also an example to be followed.

To Professors André, Alcione, and Jugurta for reaching out to me when I needed it, and I could not have done it is wasn't for their help.

To Walter and João for their invaluable contributions to this work, for the dedication and willingness to help.

To CNPq for funding my studies. To all the professors and employees of the Department of Informatics of UFV, who somehow contributed to my professional growth, especially Altino.

To my parents, Cesar and Carmen, my infinite thank for the education and for all the support and encouragement in my studies. You are my examples of life, generosity, hospitality, honesty, responsibility and determination.

To my sisters, Karem and Ingrid, for always being ready to help me and also to my nephews and nieces.

To the Rosado Corrêa family, for welcoming me to their homes and becoming my family in Ponte Nova, for the delicious lunches on Sundays and for giving me so many leisure moments.

To Guidson, Fernando and Camilo for the friendship and  welcome me.

To all the friends of the master's degree by the companionship and support in special to Paulo, Rubens, Dâmaris, Barbara, Cleyton, Gilson and Jerônimo.

Finally, and very importantly, I thank my partner, José Roberto (Betinho), for helping me whenever I needed to, often leaving his obligations to worry about mine, accompanying me in my activities. His participation was very essential for the accomplishment of this work. Without him by my side I would not get here. Thank you so much for everything, especially for helping me be a better person.

Contents

List of Figures	vi
List of Tables	vii
Resumo	viii
Abstract	ix
1 Introduction	1
1.1 Contributions	2
1.2 Dissertation Structure	3
2 Related Work	4
3 Angry Birds	7
4 Manual Module for Content Creation	9
5 Mixed-Initiative Module for Content Creation	11
5.1 Sketch Representation	11
5.2 Handling Lines Intersections	12
5.2.1 Brute-Force Search for Length Discovery	13
5.2.2 Extending Horizontal Structures	15
6 User Study	17
6.1 Empirical Methodology	17
6.1.1 Timing Results	19
6.1.2 Similarity Results	19
6.1.3 Questionnaire Results	21
7 Future Works	23

8 Conclusions	25
Bibliography	26
Appendix A Manual Module Scenarios	28
A.1 Scenario 1	28
A.2 Scenario 2	31
Appendix B Mixed-Initiative Module Scenarios	34
B.1 Scenario 1	34
B.2 Scenario 2	37

List of Figures

3.1	Example of an AB level.	7
3.2	Elements used to build levels.	8
4.1	Example of an AB level being created with the manual module.	10
5.1	Example of a structure created with the mixed-initiative module	12
6.1	Example of a structure created with the mixed-initiative module	18
6.2	User study results for the mixed-initiative module.	21
6.3	User study results for the manual module.	21
A.1		28
A.2	29
A.3	30
A.4	31
A.5	32
A.6	33
B.1	34
B.2	35
B.3	36
B.4	37
B.5	38
B.6	39

List of Tables

6.1	Scenario creation time in seconds.	19
6.2	Similarity results in Scenario 1	19
6.3	Similarity results in Scenario 2	19

Resumo

Gennaro Campos, Cesar Ruben Francisco, M.Sc., Universidade Federal de Viçosa, julho de 2017. **Ferramenta de Iniciativa Mista para acelerar a Criação de Conteúdo em Jogos Baseados em Física** Orientador: Levi Henrique Santana de Lelis .

Neste trabalho apresentamos uma ferramenta de iniciativa mista para acelerar o processo de geracão de conteúdo em jogos baseados em ca. Nosso sistema usa busca exaustiva e álgebra linear para permitir que o designer do jogo se concentre no processo criativo de geração de conteúdo, aliviando o ônus da construção manual de estruturas de jogos. Usamos um clone de Angry Birds chamado Science Birds como objeto de nossa pesquisa. Um estudo com usuários mostra as vantagens de empregar nossa ferramenta de iniciativa mista para criar níveis de Angry Birds. Nosso estudo mostrou que as pessoas são capazes de criar níveis de Angry Birds muito mais rapidamente usando nossa ferramenta de iniciativa mista do que um sistema mais tradicional que desenvolvemos como base de comparação. Além disso, os níveis criados com nossa ferramenta de iniciativa mista são comparáveis, em termos de qualidade, com aqueles criados com o sistema mais tradicional. Finalmente, os participantes do nosso estudo relataram ser fácil utilizar nossa ferramenta e que ficaram mais satisfeitos com suas experiências do que os participantes que utilizaram o sistema tradicional.

Abstract

Gennaro Campos, Cesar Ruben Francisco, M.Sc., Universidade Federal de Viçosa, July of 2017. **Mixed-Initiative Tool to Speed Up Content Creation in Physics-Based Games** Adviser: Levi Henrique Santana de Lelis

In this work we introduce a mixed-initiative tool to speed up the process of content generation in physics-based games. Our system uses exhaustive search and linear algebra to allow the game designer to focus on the creative process of content creation by alleviating the burden of manually constructing game structures. We use a clone of Angry Birds called Science Birds as the object of our research. A user study shows the advantages of employing our mixed-initiative tool for creating Angry Birds levels. Namely, our study showed that people are able to create Angry Birds levels much more quickly using our mixed-initiative tool than a baseline system we developed. Moreover, the levels created with our mixed-initiative tool are comparable in terms of quality with those created with the baseline system. Finally, the participants of our study reported to be easy to use our tool and that they were more satisfied with their experience than the participants who used the baseline.

Chapter 1

Introduction

The creative process of computer game level design is often time consuming as the designer might have to perform multiple iterations between prototyping and play testing before achieving the desired goals. The level design process can be specially time consuming in physics-based games such as Angry Birds (AB). This is because the ~~elements of the game~~ are subject to real-world effects such as gravity and friction, and it is often hard to predict the result of the interaction of several ~~game~~ elements without careful testing.

In this work we present a mixed-initiative system to speed up the process of level design in physics-based games. Our goal is to allow the game designer to focus on the creative part of the process by having the computer doing most of the tedious and time consuming tasks for the designer. Although the object of our research is a clone of AB called Science Birds,¹ the system we introduce in this work hinges on ideas that are applicable to other real-world problems that involve the collaboration between humans and algorithms, such as coordination of patient care in hospitals (BLOSS, 2011; GOMBOLAY ET AL., 2016), autonomous driving (MOK ET AL., 2015), dialog systems (ALLEN ET AL., 2001), and general collaborative planning problems (KAMAR ET AL., 2009). In this work we deal with general and important questions that are likely to appear in other human-algorithm interaction scenarios. For example, our system is concerned with balancing how much it is able to speed up the process of level creation while not interfering with the quality of the designer's work. Similar balancing problems arise in automated driving, where the system tries to reduce as much as possible the driver's workload while not interfering with the passengers' safety.

The traditional process of creating AB levels consists of manually placing game

¹<https://github.com/lucasnife/Science-Birds>

elements on the screen so that they form stable structures (i.e., structures that do not move due to the effects of gravity). Our mixed-initiative tool partially replaces this process of manually placing elements on the screen by a drawing tool. The drawing tool allows the designer to quickly sketch the structures that will compose the level. The sketches are then automatically transformed into structures that use elements of the game. Once the basic structure of the level is created through the sketching tool, the designer is then able to perform fine adjustments to the structures thus created by manually adding and removing game elements. Our goal is to reduce the amount of work the designer has to perform in order to test an idea. We achieve that by reducing the number of elements the designer has to place on the screen. We believe our mixed-initiative tool might be able to enhance the designer’s creativity by allowing them to quickly prototype different level ideas.

We performed a systematic user study to evaluate the mixed-initiative tool we introduce. Our primary goal with this study was to measure whether people are able to create levels more quickly with our mixed-initiative tool than with a baseline system that requires the manual placement of game elements. The results of our study showed that our mixed-initiative tool is indeed able to significantly and substantially speed up the process of level creation. However, this speed up is meaningful only if the quality of the levels created with both approaches (mixed-initiative and baseline) are comparable. We derived a similarity metric for assessing how similar the levels created by the participants were from the levels they were told to create. The results suggest that people create levels with similar quality while using either system. The timing results taken together with the similarity results support the hypothesis that our system is able to speed up the level creation process while not interfering with the quality of the levels created. Finally, the user study suggests that people tend to find it easier and more satisfying to use the mixed-initiative tool than the baseline system.

1.1 Contributions

The mixed-initiative tool we introduce in this work opens several research directions. Our system might be used in future works to evaluate the generation of difficulty progression plans as well as to evaluate Artificial Intelligence systems that collaborate actively with the designer by predicting the designer’s actions. The code of our system will be made available so that other researchers may also investigate these, and possibly other, research directions.

1.2 Dissertation Structure

This work is organized as follows. In Chapter 2 we review some relevant related work. In Chapter 3 we describe the AB game. In Chapter 4 we describe what we call the manual module of our system, which is used as part of our mixed-initiative system and also as the baseline system employed in our user study. In Chapter 5 we detail our mixed-initiative tool. In Chapter 6 we describe the empirical methodology used in our user study as well as our main results. In Chapter 7 we describe possible future works, and finally, in Chapter 8 we state the conclusions of this work.

Chapter 2

Related Work

The definition of mixed-initiative systems is arguably fuzzy (NOVICK AND SUTTON, 1997). In this work, we consider a system as a mixed-initiative system if it makes decisions on its own while jointly solving a task, either with a human user or another system. Allen 1999 defines the following levels of mixed initiative.

- **Unsolicited reporting.** The system reports to the user as critical situations emerges.
- **Subdialogue initiative.** The system asks for clarifications and corrections automatically.
- **Fixed subtask initiative.** The system solves a fixed task on its own.
- **Negotiated mixed initiative.** The system negotiates with other agents (human or computational) who takes the initiative.

The mixed-initiative tool we introduce in this work falls into the fixed subtask initiative, as the system solves on its own the task of creating game structures from the sketches drawn by the user. As we explain in Chapter 5, our mixed-initiative tool makes some decisions on its own while creating the level structures. These decisions might slightly change the designer's sketch, but it might substantially reduce the amount of work the designer will have. Once the system finishes creating the game structures from the sketches it returns the control to the user. Most of the mixed-initiative systems we review in this Chapter falls into the fixed subtask initiative category. We focus our review in games-related systems.

Tanagra is a mixed-initiative tool to design 2D platform levels (SMITH ET AL., 2011). Tanagra ensures that all levels created are playable and provides the ability

to look

for a human designer ~~look~~ for many different levels that meet with their goals. The human designer can iteratively refine the level by placing and moving game elements level as well as through direct manipulation of what they called the level's rhythm. Tanagra also falls into Allen's fixed subtask initiative category as it solves specific tasks when invoked by the user.

 Sketchbook Sentient is a computer-aided design tool for creating game maps that operates on high-level sketches instead of detailed maps (LIAPIS ET AL., 2014). Sketchbook Sentient enables quick prototyping of level concepts and provides quality assessment schemes through graphical interfaces. The tool uses genetic algorithms to create playable maps of games such as Blizzard's StarCraft. The game maps created by the tool are presented in real time, and while the human user uses the tool, suggestions appear on the screen allowing the user to quickly replace parts of the map being created. A computer system that contributes actively to the design process allows not only the co-creation of playable levels with the human designer, but it also allows one agent (human or computer) to "inspire" the other. Sketchbook Sentient falls into multiple mixed initiative categories. Namely, it provides design suggestions to the user (unsolicited reporting) and solves specific tasks (fixed subtask initiative).

Ropossum is a tool to design levels for the physics-based game Cut The Rope (SHAKER ET AL., 2013A,C,B, 2015). Ropossum is composed of two modules, one for automatic content generation and another to ensure level playability. Ropossum creates complete, playable Cut the Rope levels, and the designer is able to modify the levels created with an editing tool. Similarly to our system, Ropossum also falls into Allen's fixed subtask initiative category, as it is able to solve specific tasks such as verifying playability. For also dealing with a physics-based game, Ropossum is perhaps the most similar system to the one we introduce in this work. Nonetheless, our system substantially differs from Ropossum due to the number of game elements used in Cut the Rope and in AB. While levels of the former usually have a **handful of elements**, levels of the latter often have dozens of elements. The challenge in designing levels of Cut the Rope is to carefully choose where to place a ~~handful of the elements~~. By contrast, in AB the designer has to carefully choose where to place dozens of elements. Thus, the level design process is expected to be more time consuming IN AB. Moreover, while Ropossum takes the initiative to fully create levels, the user of our system does not relinquish the design control to the computer, ~~they just allow~~ the user just allows the computer to help with the level creation. Butler et al. introduced a mixed-initiative system to generate levels for the educational game of Refraction (BUTLER ET AL., 2013). The users of Refraction learn how to

perform simple math operations by playing the game. Butler et al.’s system allows the designer to define a progression plan. That is, which concepts (i.e., math concepts) must be included in a sequence of levels to allow the user to progressively learn math operations as they advance in the game. Once the progression plan is defined, the system is able to generate levels satisfying such a plan. This makes Butler et al.’s system fall into Allen’s fixed subtask initiative system. However, in addition to solving specific tasks such as “create a level satisfying a given set of progression constraints”, Butler et al.’s system alerts the user in case a modification made by the designer makes a constraint that was being satisfied as not satisfied. For that their system can also be classified into the unsolicited reporting category too. All systems mentioned above differ from our mixed-initiative system in their primary goal. That is, none of the systems reviewed explicitly aim at speeding up the level design process. Moreover, AB is perhaps the most complex game amongst all games previous mixed-initiative works have dealt with. The contributions we make in this work are certainly orthogonal to the contributions of previous works. For example, how to design good level progressions for the game of AB remains an open question. Thus, future works might investigate how the contributions of the other mixed-initiative systems can be applied to AB, and the system we introduce in this work will be a good starting point to these investigations.

for

Chapter 3

Angry Birds

Developed by Roxio, AB is a puzzle game based on physics. The player's objective is to destroy green pigs which can be guarded by block structures. The player is able to destroy blocks and pigs by throwing birds with a slingshot placed on the lefthand side of level. Figure 3.1 shows an example of a level generated using our mixed-initiative tool. AB levels can be constructed by using the elements shown in Figure 3.2. Each element can be made of rock, wood, or ice, as demonstrated by rounded object in the first row of Figure 3.2. Different materials offer different resistance, where objects made of rock are the hardest to break, while objects made of ice are the easiest. The second and third row of Figure 3.2 show all elements available to the designer in our mixed-initiative tool. Although all elements are shown made of wood, they are also available to be used as made of rock and ice ~~in our tool~~. Our mixed-initiative tool is built on top of an AB's clone called Science Birds (FERREIRA AND TOLEDO, 2014).



Figure 3.1: Example of an AB level.

In addition to its material, each game element has properties such length l and width w . In our tool the objects can be rotated and we assume that the length l of an element denotes its longest side, while its width w its shortest side, independently

orientation

of its ~~position~~. In the current version of our tool we automatically manipulate only the ~~elements showed~~ in the second row of Figure 3.2. Note, however, that the designer has access to all other elements, but they are not used in the automatic process of ~~described further in Chapter 5.~~ described in Chapter 5 below. We intend  use other elements in the process automatic element placement of future versions of our tool.

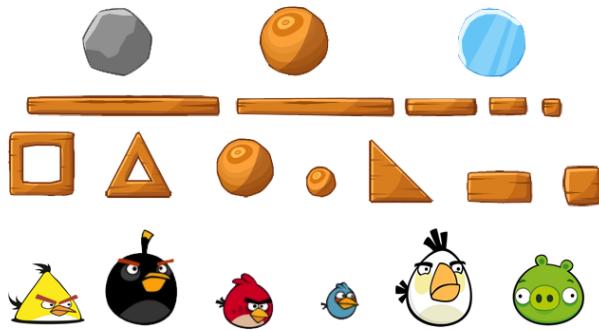


Figure 3.2: Elements used to build levels.



Chapter 4

Manual Module for Content Creation

Our mixed-initiative tool for constructing AB levels is divided into two modules. The first module, which we call the manual module, allows the designer to manually construct AB levels. We describe the manual module in this chapter. The second module, which we call the mixed-initiative module, is described in Chapter 5. A user testing user using the mixed-initiative features of our system uses both the manual module and the mixed-initiative module as the latter allows the designer to edit the content generated by the former. We also use the manual module as a baseline in our user study.

Figure 4.1 shows a screenshot of the manual module. On the upper part of the screen the designer is able to select which element they will place on the level. In this module the objects are placed manually by the designer who uses the computer's mouse to choose the element's location on the screen. On the top left corner of the screen there is a button in which the designer can choose the material of the elements (rock, wood, or ice). After placing an element on the screen the designer is able to rotate it. For example, the wooden blocks at the bottom of the structure were first placed in its default horizontal orientation, as depicted at the top of the figure, and then rotated vertically. At the bottom of the figure the designer can choose how many of each bin type the level will contain.

The designer is allowed to playtest the level being created by clicking on the green button button at the bottom-right corner of the screen. The playtest is important because the elements being placed in the level do not suffer the effects of gravity and friction while being placed in the level. By playtesting, By playtesting the designer is able to verify whether the structures are stable. Should the designer discover that a given structure is unstable, they can return to the module's editing mode to fix the source of instability. Another level property which is as important as stability is level's level's feasibility, i.e., can the level be solved? The

designer also ensures feasibility with playtesting. Finally, aiming at speeding up the process of level creation the manual module contains hotkeys for quickly selecting elements and changing the elements' materials.



Figure 4.1: Example of an AB level being created with the manual module.

Chapter 5

Mixed-Initiative Module for Content Creation

In our mixed-initiative module the designer is able to quickly sketch AB structures with our drawing tool, and our system generates a structure according to the designer's sketch. An example of this process can be seen in Figure 5.1. The drawing tool is shown in Figure 5.1a, which initially displays an empty gridded canvas. The grid contains 35×45 cells. The designer is able to draw horizontal and vertical lines that can be edited with respect to their material—we intend to implement orientations other than vertical and horizontal for the lines drawn by the designer in future versions of our system. Purple lines represent elements made of rock, blue lines elements made of ice, and brown lines elements made of wood. Once the designer finishes sketching the lines representing the level's initial structure, the system generates the structure from the sketch. Figure 5.1b shows the structure generated from the lines shown in Figure 5.1a. In the following subsections we detail how our system creates the AB structures from the ~~designer~~ designer's sketch.

5.1 Sketch Representation

The sketching tool creates a collection L of line segments (henceforth line segments will be referred as lines) where each line in L is represented by a pair of points $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$. For each line i we compute its length  follows.

$$c_i = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2},$$

and the inclination d_i ,

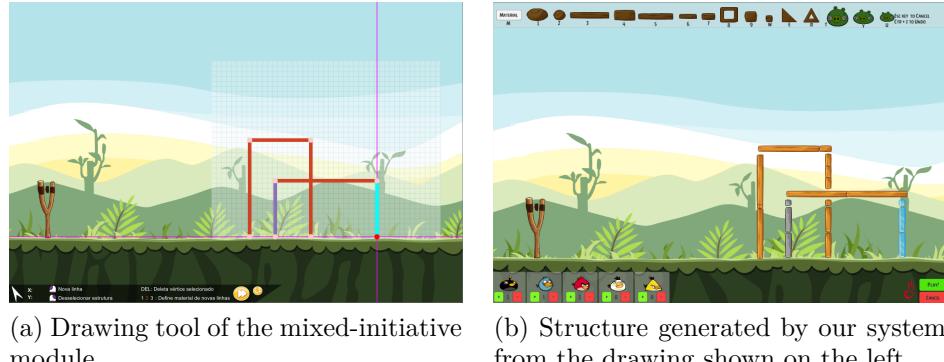


Figure 5.1: Example of a structure created with the mixed-initiative module

$$d_i = \tan^{-1} \left(\frac{y_2 - y_1}{x_2 - x_1} \right).$$

Note that in this version of the system we present in this work the inclination will be either 0 or 90 degrees (horizontal or vertical), due to the constraint we impose in our drawing tool. The lines inclination is important because we treat vertical and horizontal lines differently while searching for a combination of game elements that will form a structure representing each line, and while handling lines intersections.

5.2 Handling Lines Intersections

Our drawing tool allows the designer to create lines that intersect with each other; as an example, see the vertical and horizontal brown lines shown in Figure 5.1a. For every intersection of lines we alter the vertical line, by replacing it by two shorter lines, in a way that the intersection disappears. We replace the vertical instead of the horizontal lines because this way we are likely to reduce the amount of work the designer has to perform to make the resulting structure stable. Let us consider again the example of the crossing lines in Figure 5.1a. Figure 5.1b shows the final result, where the vertical line was replaced and the horizontal line “cuts across” the original vertical line. If we did the opposite (i.e., replaced the horizontal line by two shorter lines), the designer would have to place supporters underneath the horizontal structures to make the whole structure stable. By contrast, since we replace the vertical line, the two shorter vertical lines already serve as a support to the horizontal structure, thus likely reducing the designer’s amount of work.

We use Bourke’s algorithm (BOURKE, 1989) to detect line intersections in the collection L . Let a line i be determined by points $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$,

and line j be determined by points $p_3 = (x_3, y_3)$ and $p_4 = (x_4, y_4)$. Lines i and j intersect if the values of u_a and u_b are both between 0 and 1, otherwise i and j do not intersect. The values of u_a and u_b are computed as follows.

$$u_a = \frac{(x_4 - x_3)(y_1 - y_3) - (y_4 - y_3)(x_1 - x_3)}{(y_4 - y_3)(x_2 - x_1) - (x_4 - x_3)(y_2 - y_1)},$$

$$u_b = \frac{(x_2 - x_1)(y_1 - y_3) - (y_2 - y_1)(x_1 - x_3)}{(y_4 - y_3)(x_2 - x_1) - (x_4 - x_3)(y_2 - y_1)}.$$

Given that i and j intersect, we compute the coordinate (x, y) of such intersection as follows:

coordinates (x, y) of such intersection using the Linear Equation, as follows.

$$x = x_1 + u_a(x_2 - x_1),$$

$$y = y_1 + u_b(y_2 - y_1).$$

Formally, if vertical line i defined by points $p_1 = (x, y_1)$ and $p_2 = (x, y_2)$ intersects a horizontal line at coordinate (x, y) , assuming $y_1 < y_2$ (i.e., y_2 is above y_1) i is replaced in L by line i' defined by points $p'_1 = (x, y_1)$ and $p'_2 = (x, y - 1)$ and by line i'' defined by points $p''_1 = (x, y + 1)$ and $p''_2 = (x, y_2)$. The resulting collection L has no lines intersecting each other, and this is the collection used in the next step of our system, described below.



Example 1 The second vertical brown line (from left to right) in Figure 5.1a intersects with two horizontal brown lines. The vertical line originally occupies 20 grid cells and is replaced by two vertical lines, one that occupies 11 grid cells, from the 1st to the 11th cell in the column occupied by the line (assuming the 1st grid cell to be at the bottom of the figure), and another that occupies 7 grid cells, from the 13th to the 19th grid cells; the 12th and the 20th cells are then free to be occupied by the horizontal lines.

5.2.1 Brute-Force Search for Length Discovery

We use the elements shown on the second row of Figure 3.2 to create structures according to the lines in L provided by the drawing tool. The first element has dimensions 10.5×1 , the second 8.5×1 , the third 4.5×1 , the fourth 2×1 , and the last 1×1 where the first number represents l and the second w (we assume the elements to be placed horizontally, as shown in Figure 3.2). The dimension values of the elements are approximate as they depend on the bounding boxes used

in the game. We discovered empirically that by using these values and removing one unit from the length of each line in L result in visually pleasing structures. For instance, in Example 1, the brown line that occupies 20 grid cells is replaced by two lines, one occupying 10 grid cells ($11 - 1 = 10$) and another occupying 6 grid cells ($7 - 1 = 6$). This reduction of one unit is performed for all lines in L , not only those that intersect with other lines.

Our mixed-initiative tools chooses a sequence of elements to form each line drawn by the user. For example, should the user draw a vertical line of length 15, our system will combine smallest number of elements that together will form a line of length 15. An optimal solution is the combination of elements of length 10.5 and 4.5. Another solution is the combination of two elements of length 4.5 and three elements of length 2. We prefer the former than the latter because the former uses only two elements, while the latter uses five elements. We minimize the number of elements composing the structures generated by our system because structures with fewer elements tend to be stronger and easier to be made stable, as discussed in Section 5.2. We call the problem of finding the set of AB elements that combined form a line of specific length l as the *AB length-discovery problem*.

The main issue of minimizing the number of elements composing our structures is that the AB elements' length form a *non-canonical* problem instance. Similarly to the coin change problem, we say that a problem instance is non-canonical if a greedy algorithm is unable to produce optimal solutions (i.e., solutions with the smallest number of elements). The following proposition shows that the AB length-discovery problem is non-canonical.

Proposition

Proposition 1 *The AB length-discovery problem is non-canonical.*

Proof. Consider a set of elements  length 10.5, 8.5, 4.5, 2 and 1, and a line of length 17. An iterative greedy algorithm that chooses the largest possible element in every iteration would choose the elements of length 10.5, 4.5, and 2. By contrast, the optimal solution is to choose two elements of length 8.5. \square

Since the lines drawn by the designer are defined by integer values, there will always be a solution to the AB length-discovery problem. That is, the trivial solution of using a collection of 1×1 elements will always be applicable, independently of the line length.

In order to minimize the number of elements composing a given structure and still obtain answers in real-time, we devised a scheme in which we pre-compute the optimal combinations of elements for all sizes from 1 to 45 (recall that 45 is

the longest line possible). The pre-computed combinations are stored in a lookup table which is used by the system while constructing structures from the designer's sketches. We use a simple depth-first search algorithm to enumerate all possible combinations of elements up to the maximal length of 45. Note that other schemes such as dynamic programming are possible, but since our problem is relatively small and the lookup table is computed only once, as a preprocessing step, we chose to use the simpler depth-first search approach. Our depth-first search approach finishes in approximately 3 minutes on a 2.7 GHz CPU. Once the lookup table is computed our system is able to efficiently access the combinations through the length of the desired structure—our lookup table is implemented as a hash table, which allows access to the element combinations in constant time.

Note that the structure generated by our system is not necessarily stable. Consider for example the structure shown in Figure 5.1b—the structure shown in the figure is in editing mode, which means that the structure does not suffer the effects of gravity and friction. All vertical elements in the example will be stable once when the gravity is "turned on", as they are supported by other vertical elements once we "turn gravity on", as they are supported by other elements. Nevertheless, some of the horizontal elements will be unstable and the designer has to use the manual module of our system to edit the structure to make them stable. Intuitively, it is easy to see that vertical elements tend to be stable in the structure generated, while the horizontal elements are often unstable as they might not have supports. We experimented with a system that automatically creates supports to the structure generated, but decided not to use it as it often interfered with the design choices made by the user. For that, we decided that making the structure stable was to be done manually by the designer.

5.2.2 Extending Horizontal Structures

In order to reduce the amount of work required to make the horizontal elements stable, we implemented the following enhancement in our algorithm. Whenever creating a structure representing a horizontal line i of length c_i , we create a structure with length equal to the structure with the smallest number of elements amongst the following lengths $\{c_i, c_i + 1, \dots, c_i + \epsilon\}$, where ϵ is a positive integer. For example, for horizontal line i with length $c_i = 16$ and $\epsilon = 2$, we create a structure of size 17 instead of 16. This is because amongst the structures of lengths $\{16, 17, 18\}$, 17 is the length that requires the smallest number of elements to be formed—lines with lengths of 16 and 18 require three elements, while a line with 17 grid cells requires only two elements.

The disadvantage of the ϵ -enhancement is that our system generates structures that might be slightly different from those sketched by the designer. The advantage is that the enhancement might reduce the number of elements composing the horizontal lines, which can reduce the amount of work the designer has to perform to ensure structure stability. As an example, consider the lowest horizontal structure shown in Figure 5.1b. This structure is composed by two elements of length 10.5, resulting in a structure of length 21, while the length of the line drawn by the designer is 20 (after all adjustments described in the previous sections). A line of length 20 requires 3 elements (elements with length 10.5, 8.5, and 1), while a structure with length 21 requires only two elements (two elements with length 10.5). The designer will have much less work ensuring structure stability should the system creates a structure with 21 grid cells instead of 20 cells. This is because the structure with 21 cells contains only two elements. It is intuitively easier to add supporting structures to two large horizontal elements than to three smaller elements. We tested several ϵ -values in preliminary experiments and noticed that $\epsilon = 3$ offered a good balance between not affecting the sketch created by the designer while still reducing the amount of work required to ensure structure stability.

Note that there is no need to use the ϵ enhancement with vertical structures. This is because vertical elements tend to be stable, and they do not require editing by the designer to ensure stability.



Chapter 6

User Study

We performed a user study to evaluate our mixed-initiative approach. The baseline we use in our study is the manual module for level creation. Since our goal is to allow rapid prototyping of AB levels, we are primarily interested in measuring the time ~~the participants~~ volunteers take to reproduce AB levels that are shown to them. We also introduce a measure of level similarity to verify how accurate the ~~participants~~ level were from the original levels. We also evaluate the ~~participants~~ experience through a questionnaire.

6.1 Empirical Methodology

Our user study follows a between-subject design, i.e., each participant evaluates either the mixed-initiative module or the manual module. We decided to use such a design to avoid carry-over effects. Since the mixed-initiative module also uses the manual module for the designer's final adjustments, the ~~participants~~ would likely construct AB levels faster when using the second system (either the baseline or the mixed-initiative system), as the usage of the first system can potentially train the ~~participants~~ to create levels more quickly. The between-subject design solves this issue. *Os termos participant e participants estão se repetindo muito e com muita proximidade.*

Each participant created three AB levels. For the first level the participants were free to create whichever structures and challenges they liked. The goal of this first level was to get the participant acquainted with the tool. Once the participant finished creating the practice level we showed on a separate screen the picture of an AB level and asked the participant to reproduce that level as quickly and as accurate as possible; we call Scenario 1 (S1) this level. Once the participant finished reproducing S1, we showed the picture of another level to be reproduced, which we

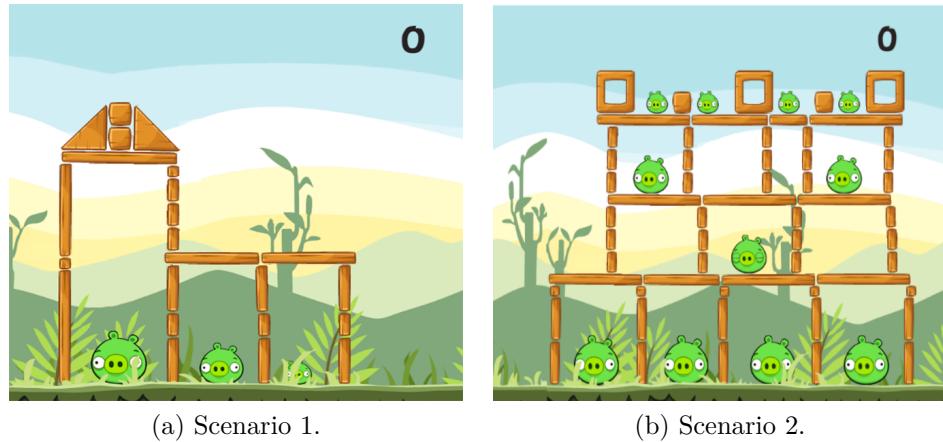


Figure 6.1: Example of a structure created with the mixed-initiative module

call Scenario 2 (S2). Figure 6.1 shows Scenarios 1 and 2, which differ considerably in size: the former has 24 structural (wooden) elements, while the latter has 58. The size difference allows us to test the efficiency of our tool in scenarios of different sizes. We measured the time it took for each participant to complete their reproduction of S1 and S2. **Para tornar a leitura mais agradável, alterne entre sinônimos**

After reproducing Scenarios 1 and 2 the participants answered a demographics questionnaire. In total the experiment had 24 participants, 12 for each system. The participants who evaluated the baseline were 26 years old on average, 9 were male and 3 were female. The participants who evaluated the mixed-initiative system were 25.75 years old on average, 8 were male and 4 were female. In addition to the demographics questionnaire, the participants answered the following three questions in a 5-Likert scale.

1. How difficult was it to use the tool? (a score of 1 means not difficult at all and a score of 5 means very difficult)
 2. How satisfied were you with the tool while reproducing S1? (a score of 1 means not satisfied at all and a score of 5 means very satisfied)
 3. How satisfied were you with the tool while reproducing S2? (a score of 1 means not satisfied at all and a score of 5 means very satisfied)

Next, we present the results of our user study in terms of timing, similarity of the levels created by the participants with the levels they were instructed to recreate, and in terms of overall user experience while using the tools.

Table 6.1: Scenario creation time in seconds.

	Scenario 1	Scenario 2
Mixed Initiative	205.55 ± 80.21	441.31 ± 140.72
Manual Module	290.34 ± 89.80	720.36 ± 409.28

6.1.1 Timing Results

The average timing results as well as the standard deviations of our user study are presented in Table 6.1. The values are presented in seconds. Welch's t tests indicate that the differences between the baseline and the mixed-initiative systems are significant with $p < .05$ in both scenarios. The timing difference between the two approaches in Scenario 1 has a Cohen's d -value of 0.98, which indicates a large effect size. The effect size is smaller for Scenario 2, as the Cohen's d -value is 0.71, which indicates an effect size between medium and large. These results support our hypothesis that our mixed-initiative system is able to significantly ($p < .05$) and substantially (effect sizes in the medium and large marks) speed up the process of prototyping AB levels.

6.1.2 Similarity Results

Table 6.2: Similarity results in Scenario 1

	Mixed Initiative	Manual Module
Structural Elements	4.08 ± 2.31	2.08 ± 3.34
Vertical Elements	21.25 ± 11.74	17.00 ± 38.87
Horizontal Elements	11.58 ± 13.63	11.5 ± 14.64
Pigs	0.17 ± 0.39	0.17 ± 14.64

Table 6.3: Similarity results in Scenario 2

	Mixed Initiative	Manual Module
Structural Elements	8.25 ± 4.94	7.00 ± 8.80
Vertical Elements	34.17 ± 32.90	44.18 ± 81.04
Horizontal Elements	76.75 ± 70.56	44.45 ± 56.36
Pigs	0.67 ± 0.78	0.18 ± 0.40

The significant and substantial positive timing results are meaningful only if the levels created by the participants are “similar enough” to the original levels. While we understand that the levels created by the participants do not have to match perfectly the levels they were told to reproduce, they have to have some

structural similarity to allow for rapid tests of level prototypes (i.e., is the initial idea for the level going to deliver the designer's intended goals)? We measure if the levels created by the participants were “similar enough” to the originals through a similarity metric.

We compute the similarity of the participants' levels and the original levels by computing the average absolute difference between the following level properties: total number of structural elements (wooden elements), total number of grid cells occupied by vertical elements, total number of grid cells occupied by horizontal elements, and total number of pigs. Intuitively, if a participant creates a level that is identical to the level they were told to recreate, then all absolute differences will be zero for that level. The average absolute differences for S1 and S2 are shown in Tables 6.2 and 6.3, respectively. As an example of how to read the table, a value of 4.08 for the mixed-initiative system in terms of structural elements means that the levels created by the participants either had 4.08 more structural elements than the original or had 4.08 fewer structural elements than the original on average.

The results shown in Table 6.2 suggest that the levels created by the participants using either system are equally similar to the original S1 level. Recall that S1 has 24 structural elements, which suggests that a difference of 2 elements (mixed-initiative absolute difference is 4.08 and the baseline absolute difference is 2.08 for structural elements) is not substantial. By analyzing the values of vertical and horizontal elements one can notice that the levels created with the mixed initiative tool were either slightly shorter or slightly taller than those created with the manual module. This is because there is almost no difference between the horizontal elements values of the two approaches, but a difference is noticed in terms of vertical elements. This result is somewhat expected, as while sketching the lines in our drawing tool the designer might miss the exact size of the structure to be reproduced. Although slightly shorter or taller on average, the levels created with the mixed-initiative tool retained the level structure shown in the original image of S1.

As can be observed in Table 6.3, the S2 levels created by the participants were less similar to the original than what was observed for S1. This is because S2 is more complex than S1. However, the results of both systems are still comparable. The mixed-initiative system presented better results in terms of vertical elements, but worse results in terms of horizontal elements. The mixed-initiative participants had problems estimating the size of the horizontal lines in the level. The poor horizontal estimations led to narrower structures in which the pigs could not fit. As a result, the absolute difference of pigs was also larger for the mixed-initiative

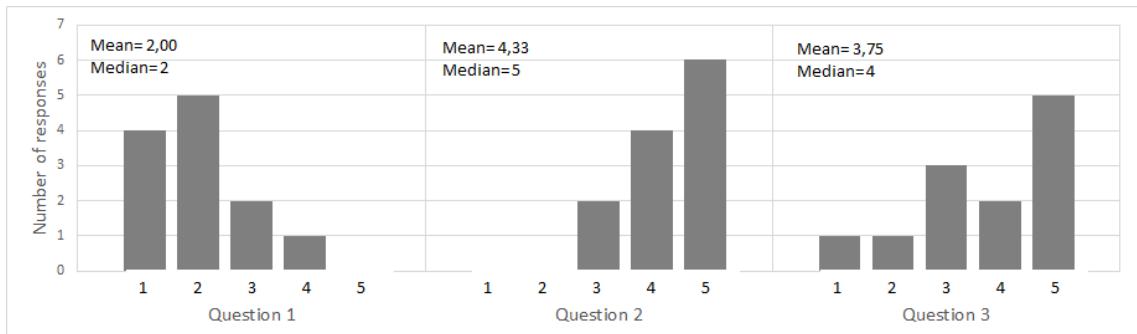


Figure 6.2: User study results for the mixed-initiative module.

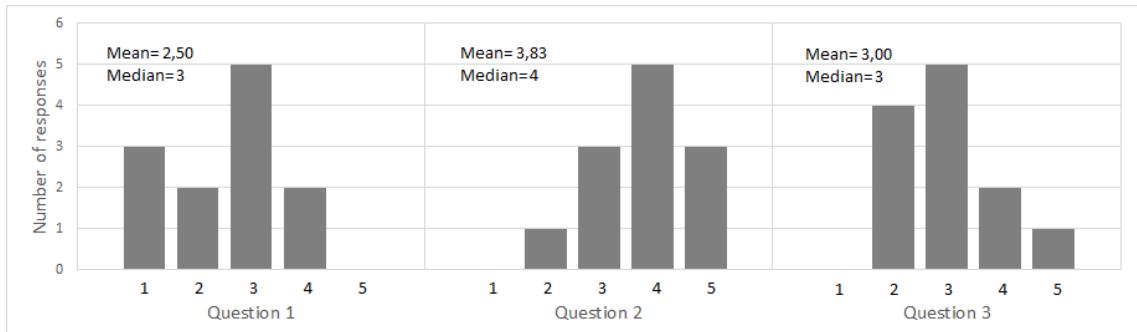


Figure 6.3: User study results for the manual module.

system. Despite these nuanced differences, the levels created by the participants using either system retained the main concepts of the original S2 level.

The similarity results taken together with the timing results strongly support our hypothesis that our mixed-initiative tool might help to speed up the process of prototyping and testing levels of physics-based games. This is because the participants who used our tool were able to recreate levels much more quickly than the participants who used a baseline system, and the quality of the levels created (in terms of similarity with the original scenarios) using either system is comparable.



6.1.3 Questionnaire Results

In addition to verifying if our tool is able to speed up the prototyping process of AB levels, we are also interested in verifying the user's experience. We measure the user experience in terms of difficulty of using the tool to recreate the scenarios (Question 1) and the users' satisfaction while recreating the scenarios (Questions 2 and 3). Figures 6.2 and 6.3 show the distribution of answers provided by the participants in a 5-Likert scale for the mixed-initiative system and the baseline, respectively.

By analyzing the leftmost histogram of the two figures one can notice that the

participants tend to find it easier to use the mixed-initiative tool than the baseline. This can be observed in terms of mean and median values, as well as in terms of the distribution of answers. For the mixed-initiative the answers are biased towards smaller numbers (1 and 2), indicating that people found it easy to use the tool. By contrast, the trend for the baseline is not as clear.

While most of the participants found it satisfying the experience of using both of our systems to reproduce S1 (see the histogram in the middle of both figures), we still observe a slight advantage for the mixed-initiative module. This is because the mixed-initiative tool has twice as many 5's than the baseline, indicating that more people found it very satisfying to use the mixed-initiative tool. Moreover, one participant reported their experience as not satisfying for the baseline. By contrast, all participants who used the mixed-initiative system said that it was either indifferent (2 responses), satisfying (4 responses), or very pleasing (6 responses).

The rightmost histogram shown in Figures 6.2 and 6.3 depict the distribution of answers of how satisfied the participants were while recreating S2. Again we observe that people tend to find more satisfying to use the mixed-initiative tool than the baseline. In fact, most of the people who used the baseline were either indifferent or not satisfied with the system (9 responses out of 12). This is because S2 is a relatively large level that requires the placement of many structural elements, what makes the task tedious and time consuming. By contrast, most of the participants who used the mixed-initiative tool were either indifferent (3 responses), satisfied (2 responses), or very satisfied (5 responses). Two participants answered that they were not satisfied with the mixed-initiative tool while recreating S2. These two participants had problems making the structure stable. We have experimented in preliminary tests with a system that automatically ensures structure stability. However, such a system was often substantially ~~changing~~ changing the user's design, and for that we decided to not use the stability system.



Chapter 7

Future Works

There are several ways of improving and continuing the mixed-initiative tool proposed in this work. First, it will be interesting to investigate how to ensure stability while not seriously affecting the user's design. Such a "stability system" could have improved the experience of the two participants who were not satisfied with the task of recreating S2 with the mixed initiative tool.

Another direction of future research is to develop a system to automatically adjust the difficulty of the levels created. There are several ways the difficulty of the level being created can be adjusted. For example, one can make the structures more resistant by duplicating their width. That is, instead of using a single line of wooden elements, the system could use two lines of wooden elements. More resistant structures usually offer more challenges to the player, as the pigs are more likely to be better protected from the shooting birds. Another way of changing the difficulty of the level is by changing the number of birds the player can shoot, where fewer birds usually imply in more challenging levels. Ideally, the mixed-initiative system would have an Artificial Intelligence (AI) system to automatically play the levels and verify their feasibility. One could even try to infer the difficulty of the level through the AI system, and then choose the set of birds accordingly.

We believe that our system can be used in the future to study difficulty progression in learning tasks, similarly to how the game of Refraction (BUTLER ET AL., 2013) was used. The chief advantage of our system over Refraction is the complexity of the underlying task. While Refraction deals with simple math problems, AB deals with complex tasks in which the definition of difficulty progression is an open question.

Our system can potentially be used to study human-AI collaboration. Currently our system reproduces almost exactly what the designer expects, but one

could imagine having an AI system collaborating actively with the designer. The AI could try to predict what the designer is trying to accomplish and then automatically build structures that will lead the designer to their goals.

An application of our system that requires further research is the use of lines and forms extracted from images, instead of our drawing tool. That is, instead of having the designer drawing the lines, one could provide images that are processed and the main forms extracted. These forms would then be provided to our system that would build the level based on such forms. For example, the user could provide the image of a house, and our system would automatically build an AB level that looks like the house in the image.

Chapter 8

Conclusions

In this work we presented a mixed-initiative system for rapid prototyping of levels of a physics-based game. The main goal with our system is to allow a quick turnaround on the designer's ideas. That is, is a given level going to deliver the experience the designer expects? Our mixed-initiative tool provides a way of quickly implementing different level ideas in domains where the process of level construction are typically tedious and time consuming.

Our mixed-initiative system uses a drawing tool to allow the designer to sketch the structure of the level to be constructed. Then, it uses linear~~zebra~~ and brute-force search to build a structure that matches the designer's sketches. A detailed user study showed the advantages of our system over a baseline for level construction. Namely, our study showed that people using our mixed-initiative tool is able to create level much more quickly than people using the baseline tool. Moreover, we showed that the levels created with the mixed-initiative tool are comparable in terms of quality with those created with the more laborious baseline. We also discovered through a questionnaire that people found it much easier to use our mixed-initiative tool than the baseline and in general were more satisfied with their experience than the other participants.

This work opens several new exciting research directions. The code of the system proposed and studied in this work will be made available online so that other researchers will be able to explore some of these ~~exciting~~ research possibilities.



Bibliography

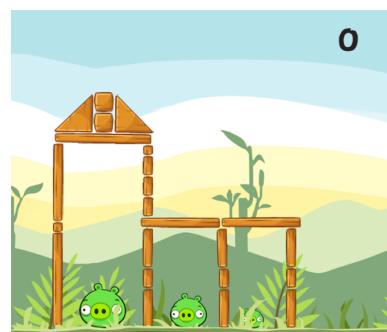
- Allen, J. F. (1999). Mixed initiative interaction. In *Proceedings of the IEEE Intelligent Systems*. IEEE Computer Society.
- Allen, J. F., Byron, D. K., Dzikowska, M., Ferguson, G., Galescu, L., and Stent, A. (2001). Toward Conversational Human-Computer Interaction. *AI Magazine*, 22(4):27--38.
- Bloss, R. (2011). Mobile hospital robots cure numerous logistic needs. *Industrial Robot*, 38(6):567--571.
- Bourke, P. (1989). Intersection point of two lines (2 dimensions). *Retrieved July, 6:2010.*
- Butler, E., Smith, A. M., Liu, Y.-E., and Popovic, Z. (2013). A mixed-initiative tool for designing level progressions in games. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, pages 377--386. ACM.
- Ferreira, L. and Toledo, C. (2014). A search-based approach for generating angry birds levels. In *Proceedings of the IEEE International Conference on Computational Intelligence in Games*.
- Gombolay, M., Yang, X. J., Hayes, B., Seo, N., Liu, Z., Wadhwania, S., Yu, T., Shah, N., Golen, T., and Shah, J. (2016). Robotic assistance in coordination of patient care. In *Proceedings of Robotics: Science and Systems*, Ann Arbor, Michigan.
- Kamar, E., Gal, Y., and Grosz, B. J. (2009). Incorporating helpful behavior into collaborative planning. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems- Volume 2*, pages 875--882. International Foundation for Autonomous Agents and Multiagent Systems.

- Liapis, A., Yannakakis, G. N., and Togelius, J. (2014). Designer modeling for sentient sketchbook. In *2014 IEEE Conference on Computational Intelligence and Games*, pages 1--8. IEEE.
- Mok, B. K.-J., Johns, M., Lee, K. J., Ive, H. P., Miller, D., and Ju, W. (2015). Timing of unstructured transitions of control in automated driving. In *2015 IEEE Intelligent Vehicles Symposium (IV)*, pages 1167--1172. IEEE.
- Novick, D. G. and Sutton, S. (1997). What is mixed-initiative interaction. In *In Proceedings of the AAAI Spring Symposium on Computational Models for Mixed Initiative Interaction*, pages 114--116.
- Shaker, M., Sarhan, M. H., Al Naameh, O., Shaker, N., and Togelius, J. (2013a). Automatic generation and analysis of physics-based puzzle games. In *Computational Intelligence in Games (CIG), 2013 IEEE Conference on*, pages 1--8. IEEE.
- Shaker, M., Shaker, N., Abou-Zleikha, M., and Togelius, J. (2015). A projection-based approach for real-time assessment and playability check for physics-based games. In *European Conference on the Applications of Evolutionary Computation*, pages 430--442. Springer.
- Shaker, N., Shaker, M., and Togelius, J. (2013b). Evolving playable content for cut the rope through a simulation-based approach. In *AIIDE*.
- Shaker, N., Shaker, M., and Togelius, J. (2013c). Ropossum: An authoring tool for designing, optimizing and solving cut the rope levels. In *AIIDE*.
- Smith, G., Whitehead, J., and Mateas, M. (2011). Tanagra: Reactive planning and constraint solving for mixed-initiative level design. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(3):201--215.

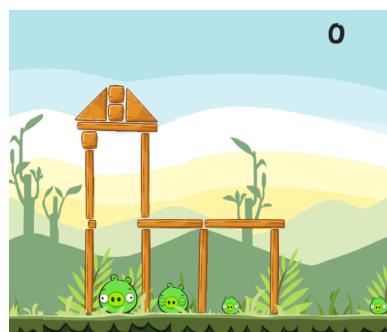
Appendix A

Manual Module Scenarios

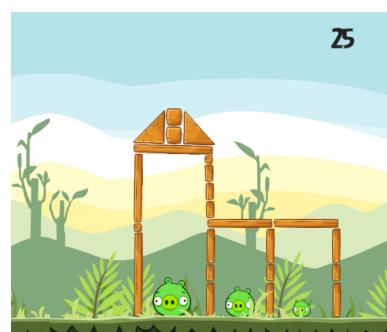
A.1 Scenario 1



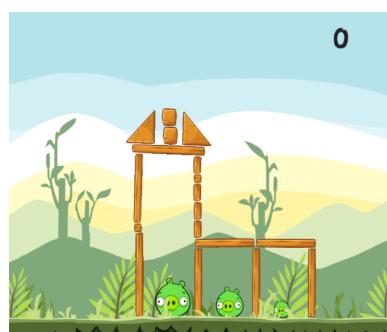
(a) Original



(b) Participant 1



(c) Participant 2



(d) Participant 3

Figure A.1

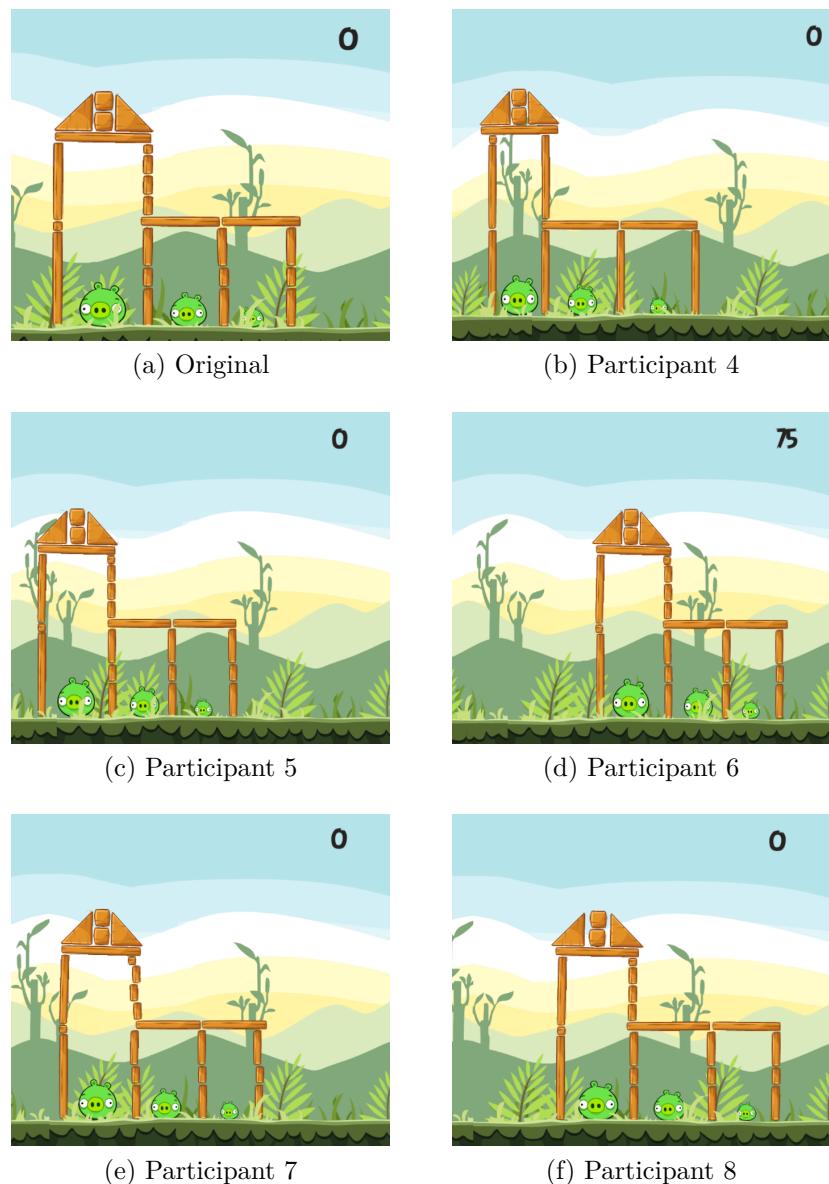


Figure A.2

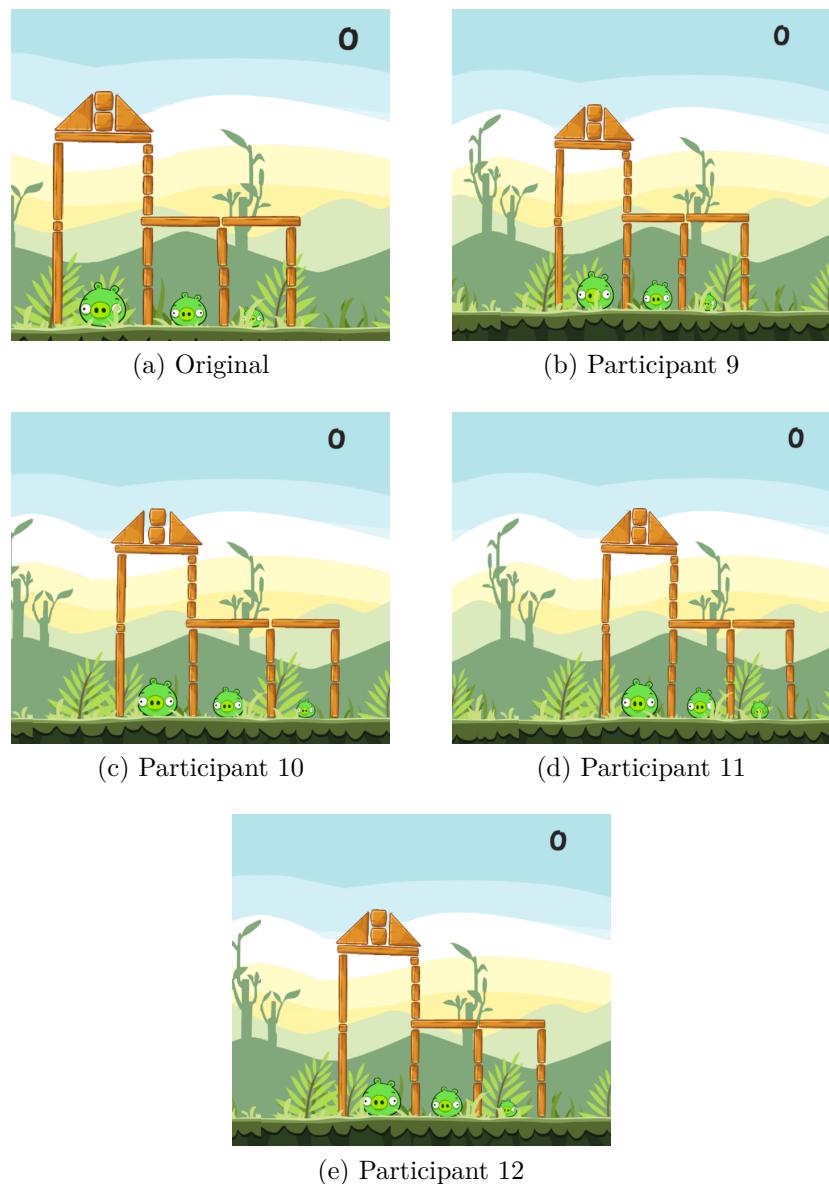


Figure A.3

A.2 Scenario 2

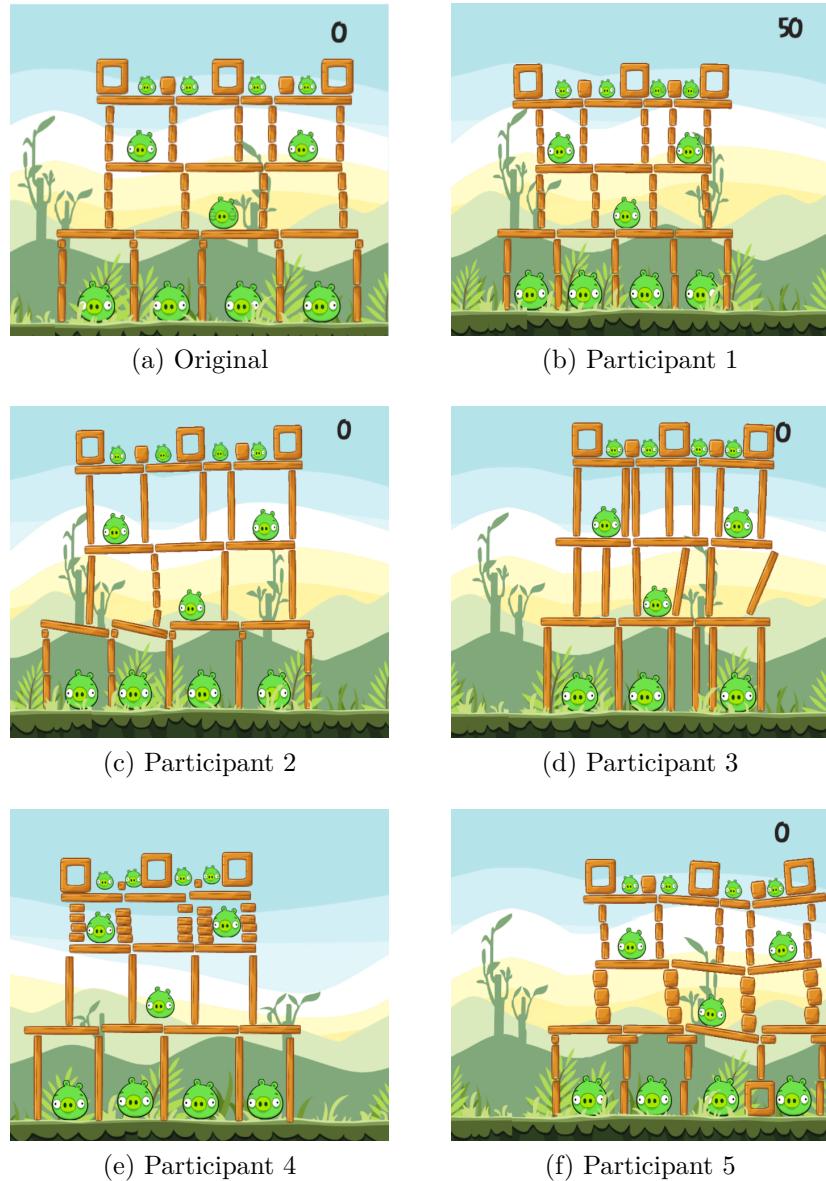


Figure A.4

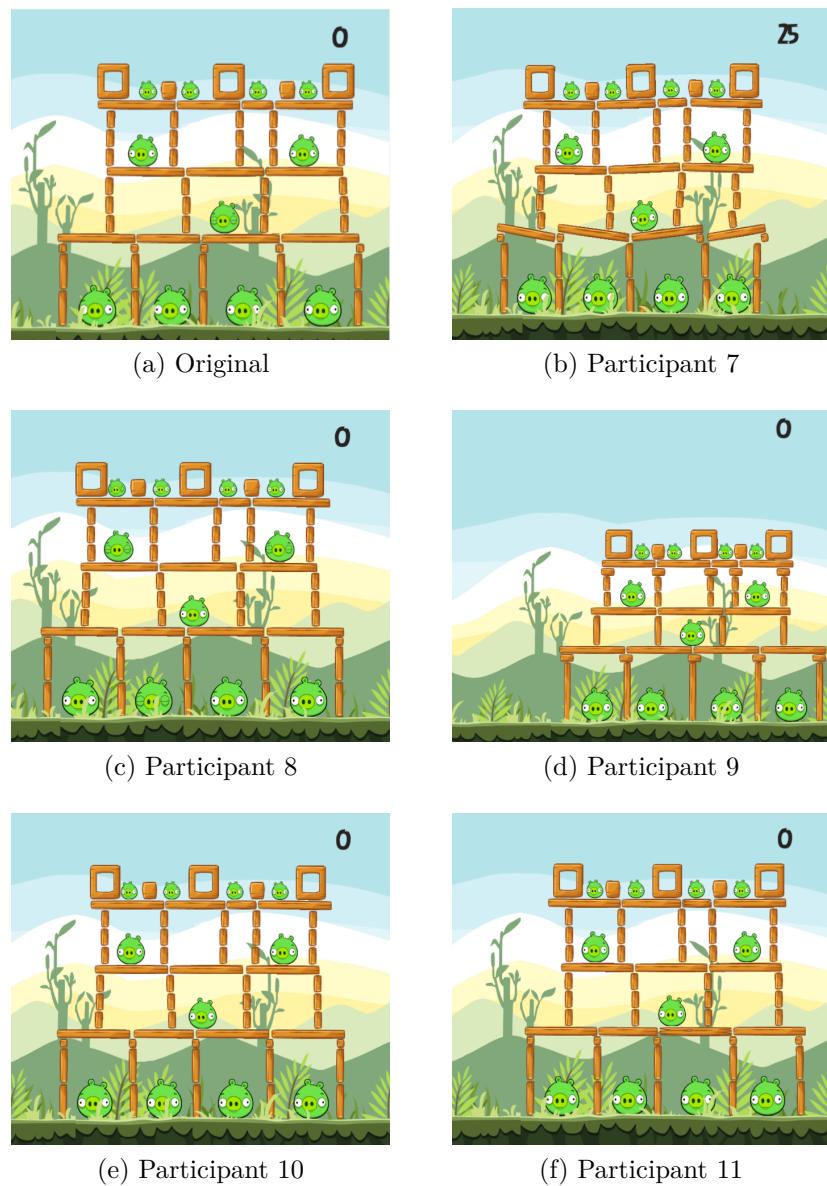


Figure A.5

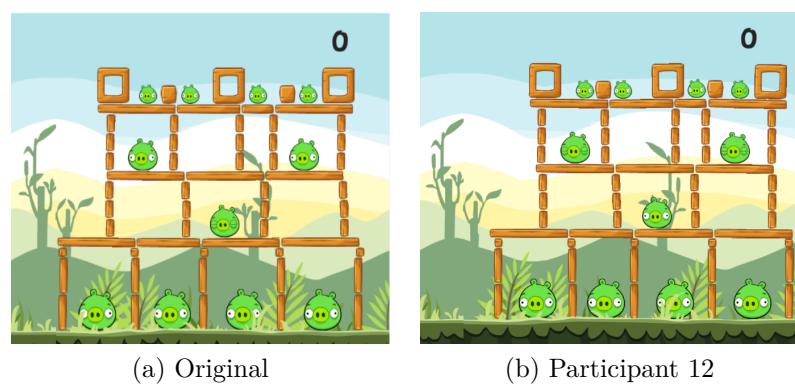
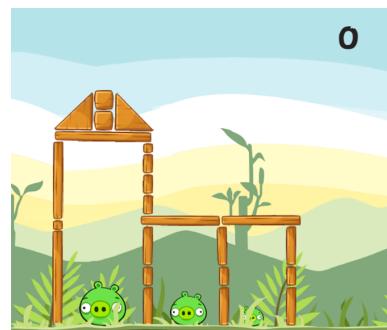


Figure A.6

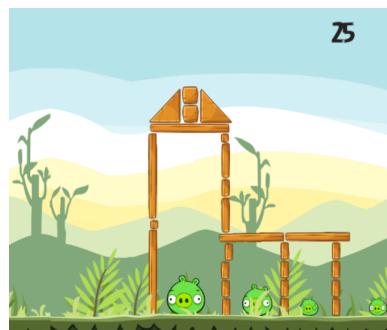
Appendix B

Mixed-Initiative Module Scenarios

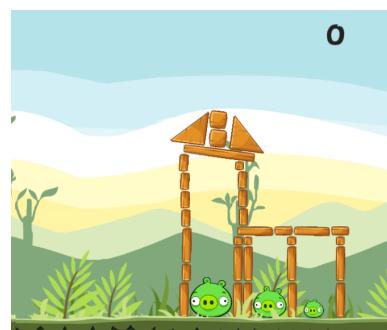
B.1 Scenario 1



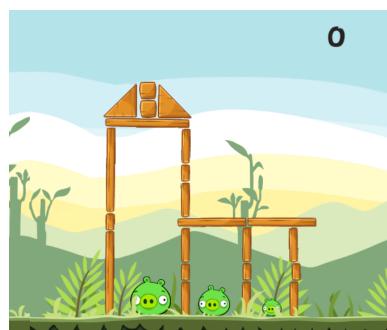
(a) Original



(b) Participant 1



(c) Participant 2



(d) Participant 3

Figure B.1

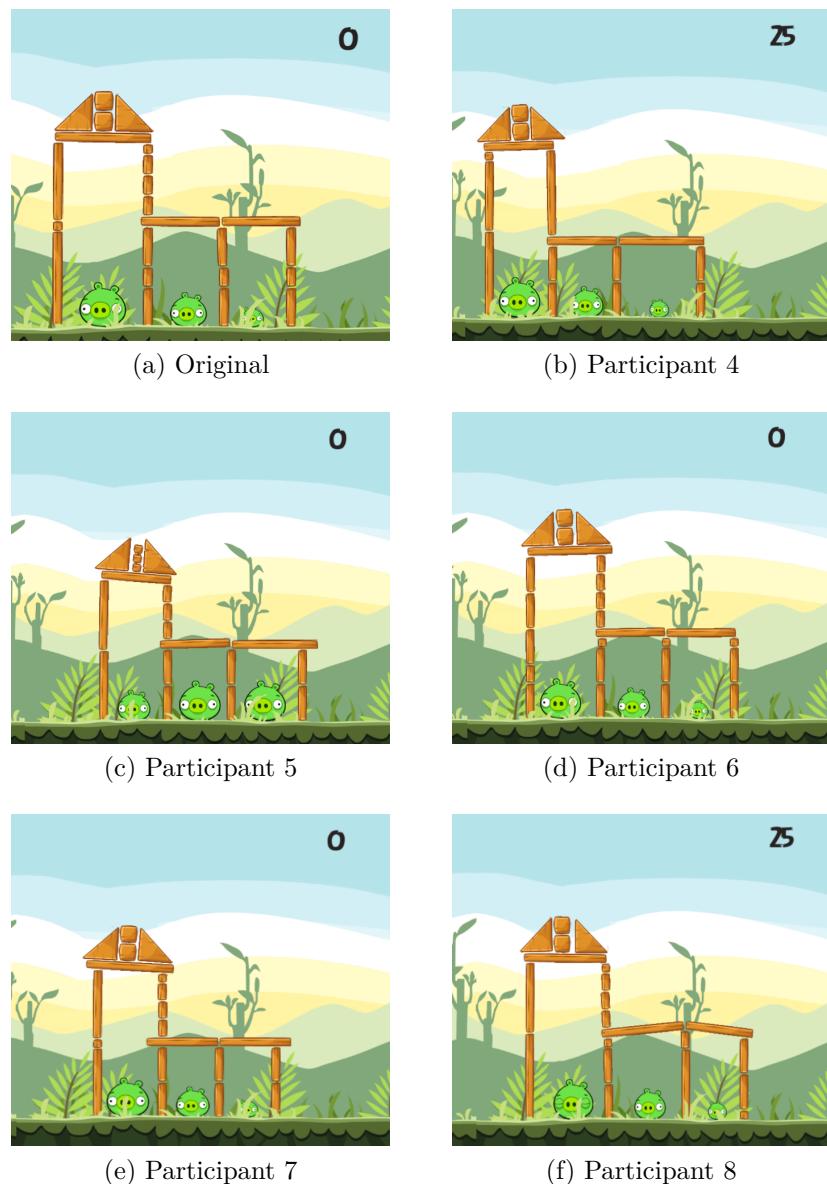


Figure B.2

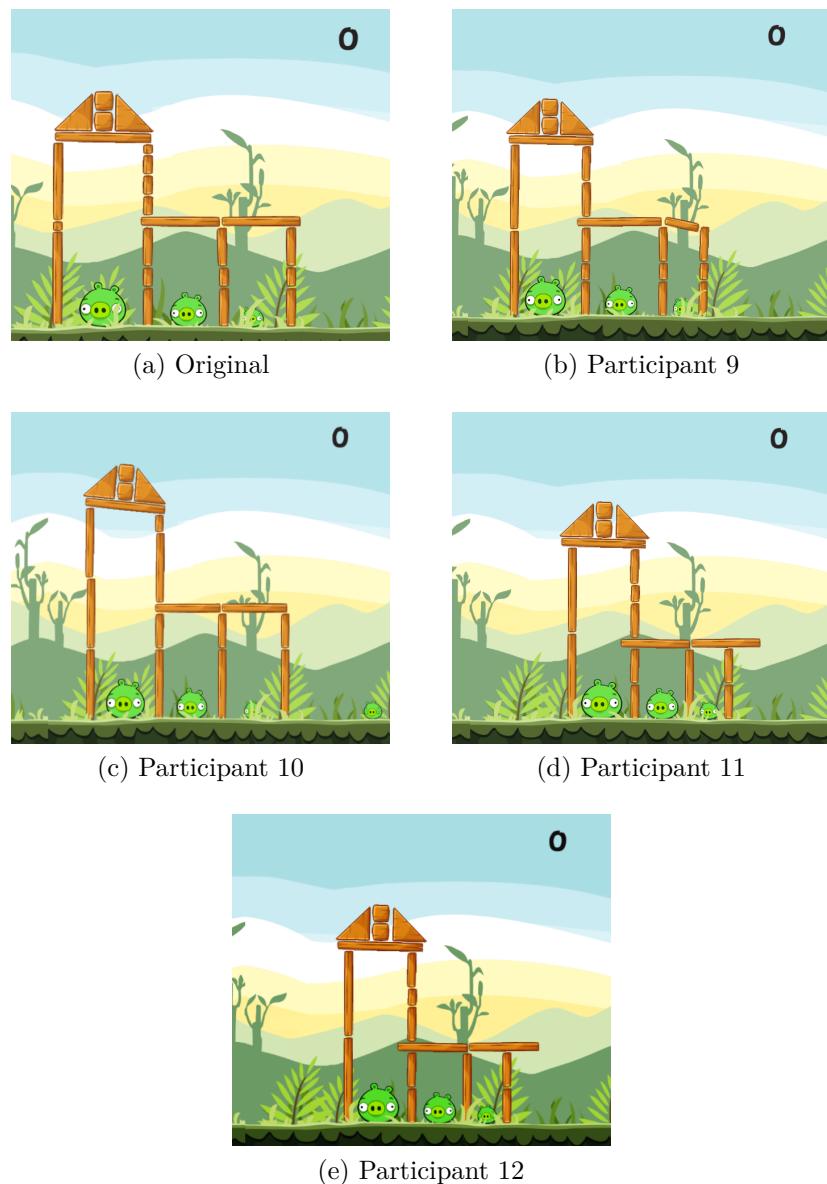


Figure B.3

B.2 Scenario 2

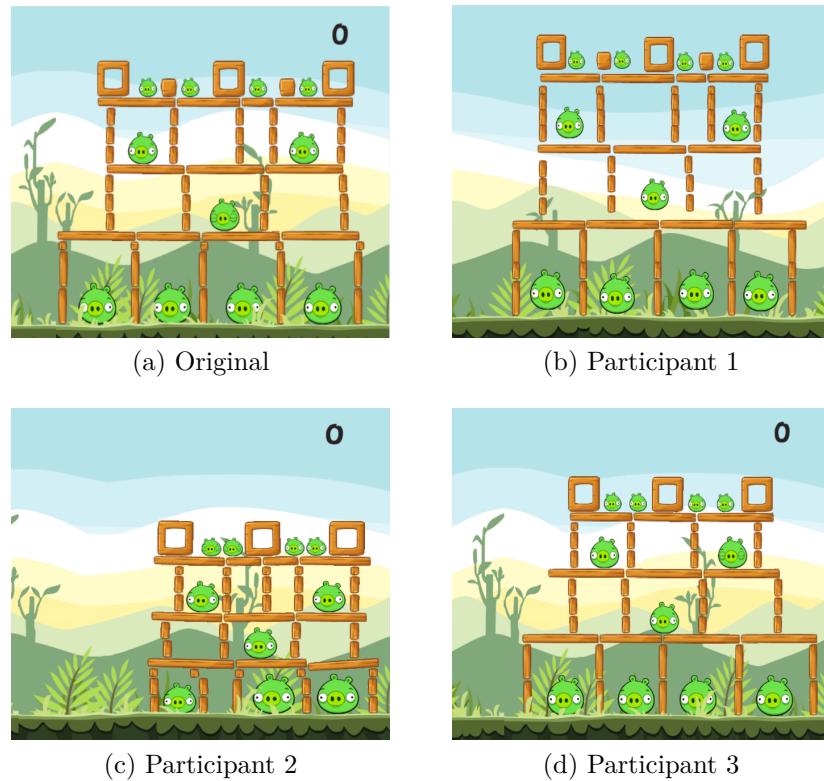


Figure B.4: .

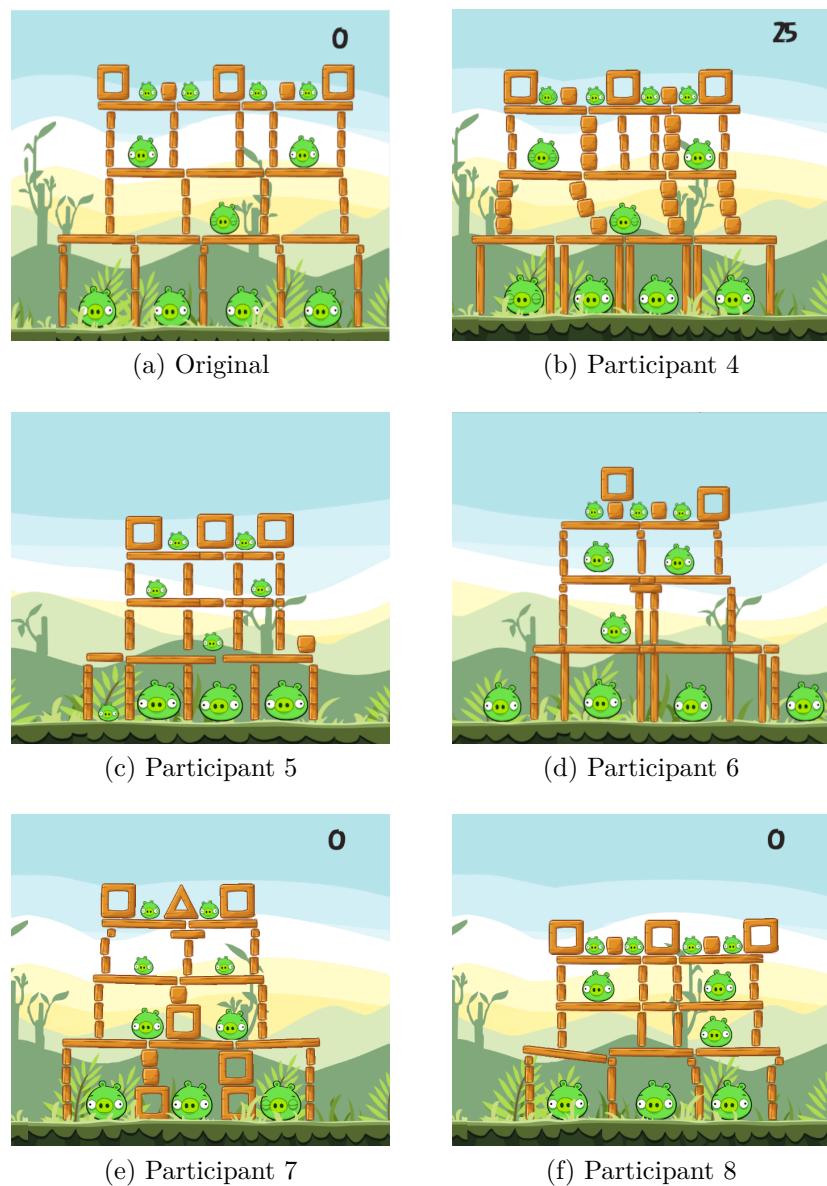


Figure B.5

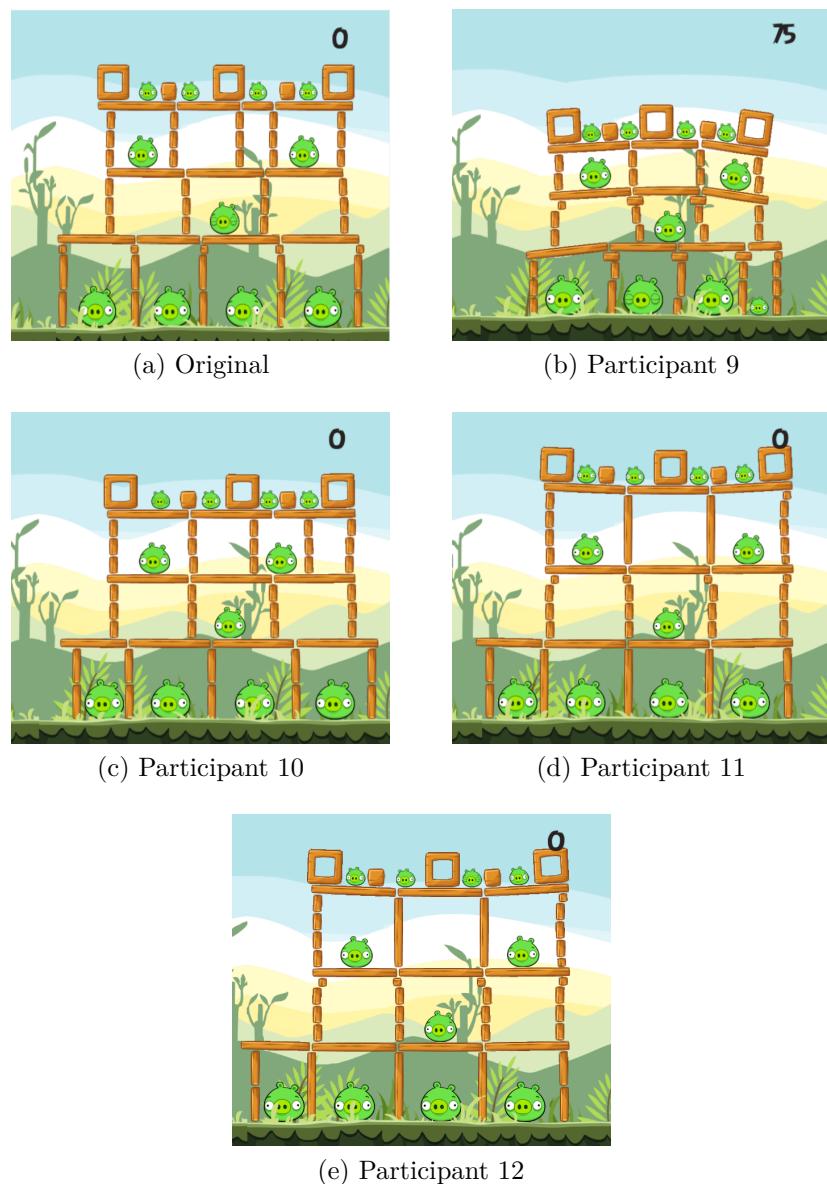


Figure B.6