# Personas versus Clones for Player Decision Modeling

Christoffer Holmgård, Antonios Liapis, Julian Togelius, Georgios Yannakakis

# Personas versus Clones for Player Decision Modeling

Christoffer Holmgård[1], Antonios Liapis[1], Julian Togelius[1], Georgios N. Yannakakis[1, 2]

[1] Center for Computer Games Research, IT University of Copenhagen, Copenhagen, Denmark
{holmgard,anli}@itu.dk,julian@togelius.com
[2] Institute of Digital Games, University of Malta, Msida, Malta
georgios.yannakakis@um.edu.mt

**Abstract.** The current paper investigates how to model human play styles. Building on decision and persona theory we evolve game playing agents representing human decision making styles. Two methods are developed, applied, and compared: procedural personas, based on utilities designed with expert knowledge, and clones, trained to reproduce playtraces. Additionally, two metrics for comparing agent and human decision making styles are proposed and compared. Results indicate that personas evolved from designer intuitions can capture human decision making styles equally well as clones evolved from human playtraces.

**Keywords:** Decision making, player modeling, evolutionary computation.

## 1 Introduction

The current paper investigates how to create generative models of human player behavior or playing style in games. This can be seen as a method for understanding game-playing behavior. Generative models of playing behavior are also potentially useful in *procedural play testing* for procedural content generation, for simulation based testing, and within mixed-initiative game design tools for instant feedback during the design process. In other words, agents that play like humans can help understand content as it is being created, by playing it. This paper assumes that game players exhibit *bounded rationality*, i.e. they play to optimize some objective or set of objectives, but that they might not be very good at it. Playing style could then be characterized by how the players' in-game decisions differ from those of an agent that played rationally (given some set of objectives). We investigate this by using AI methods to train agents that behave rationally, and see to what extent they can predict human players' behaviors.

In previous work we have designed a simple turn-based, tile-based roguelike game which features monsters, treasures and potions in mazes [6]. 38 players played 10 levels of this game, and we recorded their every action. Next, we analyzed the design of the game to extract a number of possible affordances which we translated into partially conflicting objectives that a player might seek to fulfill (e.g. kill all monsters, avoid danger or get to the exit quickly). Using these *affordances* we trained agents to play the game rationally for each objective. We call these agents *procedural personas*. Both Q-learning [6] and evolutionary algorithms [5] were used to train high-performing agents; the evolved agents have the benefit that they generalize to levels they were not trained on. The agents' behavior was compared to playtraces of the human players through a metric we call the *action agreement ratio* (AAR) which compares agents and humans at the action level. But is this really the right level of analysis for comparing players to agents? It could be argued that the microscopic level of comparing actions gives a biased view of how well an agent's behavior reproduces player behavior, and that it is more interesting to look at behavior on the level of conscious decisions. Further, are we right to assume boundedly rational behavior given some set of objectives? It might be that with the same agent representation, we could train agents that reproduce player behavior better by using the actual playtraces as training data.

The current paper tries to answer these two questions. We propose a new playtrace comparison method (*tactical agreement ratio*) that instead of asking whether an agent would perform the same action as the player in a given state asks whether it would choose to pursue the same affordance in that state. We also train agents to behave as similarly as possible to human players using playtraces as objectives; we call such agents *clones*. Clones are compared to personas on both seen and unseen levels, using both action-level and affordance-level comparison. In the following we briefly outline the relations between persona theory, decision theory, player modeling, and the resulting concept of procedural personas. We briefly describe our testbed game, MiniDungeons, and the methods we used to create game playing personas and clones, before we present the results from comparing the resulting agents to the human players.

## 2 Related Work

In this section we outline our concept of procedural personas, relating it to its roots in decision theory and the use of personas for game design.

**Decision Theory and Games.** The personas used for expressing designer notions of archetypical player behavior in MiniDungeons are structured around the central concepts of decision theory. Decision theory states that whenever a human makes a *rational decision* in a given situation, the decision is a result of an attempt to optimize the expected *utility* [7]. Utility describes any positive outcome for the decision maker and is fundamentally assumed to be idiosyncratic. This means that in principle no definite assumptions can be made about what can provide utility to the decision maker. The problem is further complicated by the fact that the effort a decision maker directs toward attaining maximum utility from a decision can be contingent on the expected utility itself. For problems that are expected to provide low utility even in the best case, humans are prone to rely more heavily on heuristics and biases for the decision making process [11, 4]. In practice, however, for structured, well-defined problems, insights from e.g. psychology or contextual information about the decision maker or the decision problem may provide us with opportunities for assuming which decisions are important and which outcomes may be of utility to the decision maker.

As decision spaces, most games are special cases since the available decisions and their consequences are highly structured by the game's mechanics and evaluation mechanisms. Games, through their design, often provide specific affordances [3] to the player, and suggest utility for various outcomes. This perspective forms the basis for our understanding of player behavior in our testbed game, as we assume that players are interacting with the game in accordance with the rules, understanding and responding to the affordances of our game. This, in turn, motivates our use of utility for attaining game rule based affordances as the defining characteristics of the personas we develop. Similar theoretical perspectives have been described by other authors, notably Dave Mark in [8].

When attempting to characterize player decision making styles in games using utilities, it is important to consider the level of decision making relevant for the game, as described in [1]. Here, we model players at both the individual action level as well as at the more tactical level of game affordances. Below we describe how we apply simple utility based agents by using linear combinations of utilities to define personas that represent archetypical decision making styles in our testbed game at two levels of abstraction.

**Player Modeling.** The concept of personas was first adapted to the domain of (digital) games under the headline of *play-personas* by Canossa and Drachen who define play-personas as "clusters of preferential interaction *(what)* and navigation *(where)* attitudes, temporally expressed *(when)*, that coalesce around different kinds of inscribed affordances in the artefacts provided by game designers" [2]. Our long term research agenda is to operationalize the play-persona concept into actual game playing procedural personas, by building generative models of player behavior from designer metaphors, actual play data, or combinations of the two.

Generative models of player behavior can be learned using a number of different methods. A key dichotomy in any player modeling approach lies in the influence of theory (vs. data) for the construction of the player model [14]. On one end, *model-based* approaches rely on a theoretical framework (in our case persona theory or expert domain knowledge) and on the other hand, computational models are built in a *model-free*, data-driven fashion. In this paper, personas represent the model-based approach while what we term *clones* represent the data-driven approach. Model-free player modeling can be done by imitating the player directly, using supervised learning methods on the playtraces, or indirectly using some form of reinforcement learning to train agents to behave in a way that agrees with high-level features extracted from the playtraces [12]. Evolutionary computation can be used to optimize an agent to behave similarly to a playtrace or optimize it to exhibit the same macro-properties as said playtrace [9, 12, 13]. Direct imitation is prone to a form of overfitting where the agent only learns to cope with situations which exist in the playtraces, and might behave erratically when faced with new situations. Indirect imitation to a large extent solves this problem by learning a more robust, general strategy, which could be termed a decision making style. In the following section we give a brief introduction to our testbed game.

## 3 MiniDungeons

The testbed game, *MiniDungeons*, implements the fundamental mechanics of a roguelike dungeon exploration game where the player navigates an avatar through a dungeon containing enemies, powerups, and rewards. The turnbased game puts the player in a top-down viewed tilebased 12 by 12 dungeon containing monsters, potions, and treasures. Impassable tiles constitute the walls of the dungeon, while passable tiles contain enemies or items for the player. All of the level is visible to the player who can move freely between passable tiles. When the player moves to a tile occupied by a monster or item, immediately the monster is fought or the item is collected and applied. The player has a 40 *hit point* (HP) health counter and dies if this drops to zero. Monsters randomly deal between 5 and 14 HP of damage while potions heal 10 HP up to the maximum value of 40 HP. Treasures have no game mechanical effect other than adding to a counter of collected treasure. The game contains one tutorial level and 10 "real" levels. For further details on the test-bed game and discussion of its properties, we refer to [6]. The necessary data for developing and evaluating the agents was collected from 38 anonymous users who played MiniDungeons online; this resulted in 380 individual playtraces on the 10 MiniDungeons levels provided. The data was subsequently used to evolve clones and baseline agents as described below. Fig. 1 shows a selected level from the game, along with human playtraces from the level, exemplifying the diversity of human decision making styles expressed in even a simple game like this.

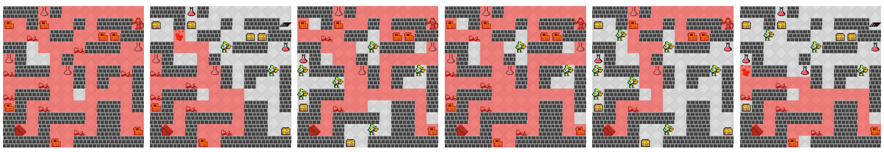

**Fig. 1.** Heatmaps of six selected human playtraces in Level 2 of MiniDungeons, showing a diversity of player decision making styles

## 4 Method

This section describes the two *agreement ratio* metrics used to evaluate persona and clone likeness to humans, the fitness functions for personas and clones, and the evolutionary approach used. The metrics address the problem of decision characterization at two different levels of abstraction.

**Action Agreement Ratio.** The first metric used to evaluate agent to human likeness is the *action agreement ratio* (AAR). AAR considers each step of a human playtrace a distinct decision. To produce the AAR between an agent and a human player, all distinct game states of the human playtraces are reconstructed. For each game state, the agent being tested is inserted into the game state and queried for the next preferred action, essentially asking: "What would you do?". If the action is the same as the actual next human action, the agent is awarded one point. Finally the AAR is computed by dividing the points with the number of decisions in the human playtrace.

**Tactical Agreement Ratio.** The second metric used for evaluating the likeness between agents and humans is the *tactical agreement ratio* (TAR). TAR only considers reaching each distinct affordance in the level a significant decision, ignoring the individual actions in between. For MiniDungeons the affordances considered relevant are: fighting a monster, drinking a potion, collecting a treasure, or exiting a level. For each affordance reached in the human playtrace, the resulting game state is reconstructed and the agent being tested is inserted into the game state. The agent is then allowed as many actions as necessary to reach the next affordance, again asking the question "What would you do?", but at the tactical affordance level. If the next encountered affordance matches the actual next human one, the agent is awarded a point. Finally the TAR is computed by dividing the points with the number of affordances reached in the human playtrace.

**Evolved Agent Controllers.** The controllers of the game agents are represented as seven linear perceptrons. Each perceptron takes 8 inputs describing safe and risky path distances to the nearest affordances in the map. Further details of the controller representation is given in [5]. Controllers are evolved using a ($\mu$+$\alpha$) evolutionary strategy without self-adaptation. For each generation the top 2% performing elite individuals remain unchanged, the lowest performing half of the remaining population is removed, and single-parent offspring from the remaining individuals are produced to maintain the population size. Finally all individuals not existent in the elite are mutated. Mutation is accomplished by changing each connection weight in the network with a random number drawn from a Gaussian distribution centered around zero with a standard variation of 0.3, a value confirmed as useful for this game by informal experimentation. All experiments are done using a population size of 100 individuals, trained for 100 generations. Controllers are initialized with random connection weights for all connections in the linear perceptrons.

**Personas.** For the purpose of the experiments 5 individual personas with different utility configurations were defined, based on designer interpretations of likely gameplay in MiniDungeons. The personas were intended to represent five hypothetical extreme decision making styles in interacting with the game: an *Exit* (E) persona who simply tries to escape the level, a *Runner* (R) persona who tries to escape the level in as few steps as possible, a *Survivalist* (S) persona who tries to avoid risk, a *Monster Killer* (MK) persona who tries to kill all monsters and escape the level, and a *Treasure Collector* (TC) persona who attempts to collect all treasures and escape the level. The decision making styles are defined by the utility weights presented in Table 1 and serve as a metaphor for the relative importance of the affordances to the archetypical player represented by the persona. When assigned to personas, utility points from a level are normalized by the maximally attainable utility for the same level. Personas are evolved by, for each generation, exposing them to 9 of the 10 levels of MiniDungeons, yielding 50 agents in total. For each generation, their fitness is computed as the average of the normalized utility scores from the seen levels. All subsequent evaluations presented in this paper are done using 10-fold cross validation, i.e., a persona is evaluated on the level which it was not exposed to during evolution.

**Clones.** Clones, like personas, are evolved by exposing them to 9 of the 10 levels of MiniDungeons. Their fitness value is computed as the average normalized AAR across all 9 seen levels. One clone per player per map is evolved, yielding 380 agents in total. All subsequent tests are done using 10-fold cross validation, evaluating the clones on unseen levels.

**Baseline Agents.** In order to evaluate the limits of the perceptron-based representation, a set of baseline agents is evolved, one agent for each human playtrace, 380 total. These are exposed to a single level of MiniDungeons. Their fitness scores are computed directly from AAR in an attempt to establish the closest fit to each human player that the representation can achieve.

## 5 Results

This section compares the two presented evaluation metrics, and compares the ability of personas, clones, and baseline agents to represent human decision making styles in MiniDungeons. Table 2 shows the mean of the agreement ratios for each kind of agent evolved, using both the AAR and TAR metrics. The ratios indicate that all agents achieve higher agreement with human playtraces when evaluated with the AAR metric than with the TAR metric. Additionally, they indicate that when using AAR clones perform only slightly better than personas ($t = -3.23$, $df = 753.00$, $p < 0.001$) while when using TAR the clones perform substantially better than the personas ($t = -39.26$, $df = 721.51$, $p < 0.001$), as tested using Welch's t test. Using AAR, the baseline agents perform significantly better than both personas and clones ($df = 2$, $F = 62.59$, $p < 0.001$), but when using TAR they perform significantly worse than the clones ($df = 2$, $F = 59.1$, $p < 0.001$), as tested using ANOVA. Table 3 shows which personas exhibited the best ability to represent human playtraces, for each MiniDungeons level and in total. For each human playtrace, the personas with the highest AAR and TAR, respectively, are identified. Both metrics generally favor the Treasure Collector persona as the best match for most playtraces, although there is some discrepancy between the two measures in terms of which personas represent the human playtraces best.

**Table 1.** Utility weights for the five designed personas

| Affordance | E | R | S | MK | TC |
|---|---|---|---|---|---|
| Move | -0.01 | -0.02 | -0.01 | -0.01 | -0.01 |
| Monster | | | | 1 | |
| Treasure | | | | | 1 |
| Death | | | -1 | | |
| Exit | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |

**Table 2.** Agreement ratios for personas, clones, and baseline agents

| Agent | Metric | Mean | SD |
|---|---|---|---|
| Personas | AAR | 0.75 | 0.08 |
| Clones | AAR | 0.77 | 0.08 |
| Baseline Agents | AAR | 0.81 | 0.09 |
| Personas | TAR | 0.62 | 0.11 |
| Clones | TAR | 0.66 | 0.13 |
| Baseline Agents | TAR | 0.61 | 0.13 |

**Table 3.** Best persona matches based on Action Agreement Ratio (AAR) and Tactical Agreement Ratio (TAR), respectively

| | AAR | | | | | | | | | | | TAR | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Total | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Total |
| E | 0 | 2 | 5 | 1 | 0 | 1 | 5 | 1 | 2 | 3 | 20 | 0 | 1 | 2 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 6 |
| R | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 3 | 1 | 0 | 3 | 6 | 0 | 1 | 0 | 16 |
| S | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 6 |
| MK | 8 | 8 | 0 | 2 | 3 | 1 | 7 | 2 | 2 | 0 | 33 | 5 | 3 | 1 | 4 | 0 | 1 | 2 | 3 | 4 | 0 | 23 |
| TC | 30 | 27 | 33 | 35 | 35 | 36 | 26 | 35 | 34 | 35 | 326 | 33 | 31 | 28 | 33 | 38 | 33 | 30 | 35 | 32 | 36 | 329 |
| Total | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 380 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 38 | 380 |

## 6 Discussion

It seems that the AAR metric achieves higher agreement ratios than the TAR metric. The two metrics aren't directly comparable, however, as the level's specific layout has a higher influence on the AAR value than the TAR value. Additionally, clones and baseline agents were evolved toward AAR, rather than TAR, for these experiments. Evolving toward TAR might have yielded different results. Other external playtrace comparison metrics could advantageously be used for calibration such as aggregated statistics of in-game event occurrences or other action/edit-distance based methods such as the *Gamalyzer* metric [10]. The fact that personas and clones perform roughly equally well, when measured by AAR, suggests that the persona method is a viable approach to modeling player decision making styles from expert knowledge. The method is less playtrace-dependent and computationally expensive than the cloning method, but needs an expert game designer. Still, some players may exhibit decision making styles that cannot be captured by the designer's intuition, and would be captured better by the cloning approach, as suggested by the higher agreement ratios obtained from the clones. In order to address this issue, we would propose using observed deviation from initial persona behavior to guide the evolution of new utility configurations for subsequently derived personas, combining the persona and cloning approaches.

## 7 Conclusion

This paper presented two methods of modeling player decision making styles. One was based on *personas*, evolved from designer expert knowledge, the other was based on *clones*, based on human playtraces. Two metrics were used to evaluate the agents' abilities to represent human decision making styles. The methods were shown to perform almost equally well when compared at the action level, while clones performed better than personas when compared at the affordance level.

**Acknowledgements.** We thank the players of the game. The research was supported, in part, by the FP7 ICT project C2Learn (project no: 318480) and by the FP7 Marie Curie CIG project AutoGameDesign (project no: 630665).

## References

1. Canossa, A., Cheong, Y.G.: Between Intention and Improvisation: Limits of Gameplay Metrics Analysis and Phenomenological Debugging. DiGRA Think, Design, Play (2011)
2. Canossa, A., Drachen, A.: Play-Personas: Behaviours and Belief Systems in User-Centred Game Design. In: Human-Computer Interaction–INTERACT 2009, pp. 510–523. Springer (2009)
3. Gibson, J.: The Concept of Affordances. Perceiving, Acting, and Knowing pp. 67–82 (1977)
4. Gigerenzer, G., Gaissmaier, W.: Heuristic Decision Making. Annual Review of Psychology 62, 451–482 (2011)
5. Holmgård, C., Liapis, A., Togelius, J., Yannakakis, G.N.: Evolving Personas for Player Decision Modeling. In: IEEE Conference on Computational Intelligence and Games (2014)
6. Holmgård, C., Liapis, A., Togelius, J., Yannakakis, G.N.: Generative Agents for Player Decision Modeling in Games. In: Foundations of Digital Games (2014)
7. Kahneman, D., Tversky, A.: Prospect theory: An Analysis of Decision under Risk. Econometrica: Journal of the Econometric Society pp. 263–291 (1979)
8. Mark, D.: Behavioral Mathematics for Game AI. Course Technology Cengage Learning (2009)
9. Ortega, J., Shaker, N., Togelius, J., Yannakakis, G.N.: Imitating Human Playing Styles in Super Mario Bros. Entertainment Computing 4(2), 93–104 (2013)
10. Osborn, J.C., Mateas, M.: A Game-Independent Play Trace Dissimilarity Metric. Foundations of Digital Games (2014)
11. Rubinstein, A.: Modeling Bounded Rationality, vol. 1. MIT Press (1998)
12. Togelius, J., De Nardi, R., Lucas, S.M.: Towards Automatic Personalised Content Creation for Racing Games. In: Computational Intelligence and Games, 2007. CIG 2007. IEEE Symposium on. pp. 252–259. IEEE (2007)
13. Van Hoorn, N., Togelius, J., Wierstra, D., Schmidhuber, J.: Robust Player Imitation using Multiobjective Evolution. In: Evolutionary Computation, 2009. CEC'09. IEEE Congress on. pp. 652–659. IEEE (2009)
14. Yannakakis, G.N., Spronck, P., Loiacono, D., André, E.: Player Modeling. In: Artificial and Computational Intelligence in Games, pp. 45–55. Dagstuhl Publishing, Saarbrücken/Wadern (2013)