An Expressive Dilemma Generation Model  
for Players and Artificial Agents

Sarah Harmon  
University of California, Santa Cruz  
1156 High Street  
Santa Cruz, CA 95064 USA

Abstract

Dilemma scenarios and knowledge of social values support engaging and believable narrative gameplay. Previous work in dilemma generation has explored scenarios involving utility payoffs between two to three people. However, these models primarily require character relationships as preconditions, and do not extend to more complex choices that relate only to causes and values. This paper builds upon past work to create an expressive model of dilemma categorization for player and non-player characters.

Introduction

A dilemma is a literary device which forces a person to make a difficult decision between two or more alternatives. Dilemma scenarios have been used extensively in classical and modern literature, and have become an important component of digital games and game design (Zagal 2009; Mateas, Mawhorter, and Wardrip-Fruin 2015). These scenarios may be encountered or observed. A player that encounters a dilemma during play is able to analyze real-world conflicts while avoiding stereotypical thought patterns (Kemmer 2014). Conversely, someone who observes a second party undergo a dilemma personalized to their moral values may gain empathy and respect for that person (Upright 2002). The development of a system which is able to reason about and generate personalized dilemmas with identifiable values in real time would support these dynamics.

In narrative conflict generation, dilemmas may be categorized as a type of internal conflict, which arises when a character’s plan prevents (or fails in an attempt to prevent) another one of their plans from succeeding (Ware and Young 2012). Related systems which consider the motivations of individual characters have focused on the planning aspects of agent or player conflict resolution, but often rely on arbitrary hand-authored moral values (Si, Marsella, and Pyna- dath 2006; Swanson and Jhala 2012; Szilas 2003; Ware and Young 2014). Only one known system, GADIN, specifically creates interactive narratives via adaptive dilemma generation (Barber 2008; Barber and Kudenko 2009). GADIN recognizes five core dilemma categories, which are described in Table 1: betrayal, sacrifice, greater good, take down, and favor. Each dilemma is meant to represent a conflict of utilities, which are domain-specific “scores” that reflect the assumed positives and negatives of any state in the story-world for each character. For example, betrayal is a conflict between a high utility for the self and a low utility for a friend. Notably, favor is the only category where there is no discernible benefit to the self. In such a scenario, the decision impacts the utilities of the two characters involved.

The same authors describe a user model for adaptive dilemma generation, which takes the following player traits into consideration: honesty, responsibility, faithfulness, selfishness, strength of character, morality, values related to character relationships, and the strength with which the player holds each of their principles (Barber 2008). Because it can adjust weights for these values based on user actions, the authors claim their system will present dilemmas likely to result in the most difficult decision for the user.

The GADIN system provides an important foundation for adaptive dilemma generation. However, while this framework may generate dilemmas similar to those found in soap operas, it fails to map to some morality decision points featured in narratives and contemporary games. For example, one might have to decide whether to save their life’s work, or protect their home from destruction. This is not necessarily a conflict between character utilities, but between values. One way to model this situation might be to extend the favor scenario to include story entities beyond characters.

Another limitation of the GADIN system is that conflicts are described or reduced to be in terms of two people: either the self and an enemy or a friend, or between two people besides the self. Barber and Kudenko claim that “a decision on a dilemma involves only two recipients of direct differing utility payoffs. Other dilemmas can be reduced to this form” (Barber and Kudenko 2009). However, when considering dilemmas between more than two individuals, this strategy fails to represent all possible scenarios without information loss. Additionally, there is no means of representing a choice between how a person - whether enemy or friend - is helped, or destroyed. The possibility of unclear relationships further hinders accurate dilemma categorization.

1Principles are individual-specific values. For instance, a character may have a “not stealing” or “not drugging” principle that influences whether they decide to commit a crime. Only one crime is associated with each principle.
Finally, there is no given justification in current dilemma generation systems for the modeled user values. This approach is in direct contrast to psychology theory which has identified universal values that are consistently recognized across cultures (Schwartz 2012). Because the values and categories presented in the GADIN framework appear to be system-specific, it is unlikely to accommodate all individuals and cultures.

Mawhorter and others have been working on another interactive choice generation system, Dunyazad, which “dynamically builds choices with the goal of achieving specific poetic effects (Mateas, Mawhorter, and Wardrip-Fruin 2015). The goal of Dunyazad is to estimate the poetic effects of the choices it creates, by assuming that the player will be invested in certain goals and values. This type of system would also benefit from an underlying model of universal player values.

Battaglino et al. have investigated how narrative characters might be driven by values, and how factors like emotions and social appraisal might influence decision-making (2012; 2014). While developing a complete representation for elements such as agent planning, emotional state, personality, mood, and so on is beyond the scope of this paper, we aim to explore scenarios and factors in existing game dilemmas that have not been fully recognized. For example, the perceived dependability of the agent presenting the dilemma or that of calculated utilities could play a role in one’s decision (e.g., “Tim says I might lose my job, but I don’t trust him”). By considering dilemma as an intricate narrative device, we hope to contribute toward a comprehensive dilemma model. The goals of this paper are to (1) develop a more expressive model for dilemma generation in digital games grounded in social psychology theory and (2) conduct a preliminary evaluation to determine whether this model represents a wider variety of dilemma choices in hand-authored games.

Background

Schwartz identified ten basic values that all cultures recognize, and developed a theory which explains how these values are related (Schwartz 1992; 1994; 2006). This “Theory of Basic Human Values” has been found to be consistently reliable, unlike other value systems which have been criticized and have fallen into disuse (Beatty et al. 1985). Schwartz values can be organized on a motivational continuum, which provides a means of insight as well as predictive power for individual decision making, attitudes, and behavior (Figure 1).

More recently, Schwartz and others have proposed a nineteen-value system by expanding some of the original ten (Schwartz 2012). These values are:

- **Self-direction-thought**: Freedom to cultivate one’s own ideas and abilities.
- **Self-direction-action**: Freedom to determine one’s own actions.
- **Stimulation**: Excitement, novelty, and change.
- **Hedonism**: Pleasure and sensuous gratification.
- **Achievement**: Success according to social standards.
- **Power-dominance**: Power through exercising control over people.
- **Power-resources**: Power through control of material and social resources.
- **Face**: Security and power through maintaining one’s public image and avoiding humiliation.
- **Security-personal**: Safety in one’s immediate environment.
- **Security-societal**: Safety and stability in the wider society.
- **Tradition**: Maintaining and preserving cultural, family, or religious traditions.
- **Conformity-rules**: Compliance with rules, laws, and formal obligations.
- **Conformity-interpersonal**: Avoidance of upsetting or harming other people.
- **Humility**: Recognizing one’s insignificance in the larger scheme of things.
- **Benevolence-caring**: Devotion to the welfare of ingroup members.
- **Benevolence-dependability**: Being a reliable and trustworthy member of the ingroup.
- **Universalism-concern**: Commitment to equality, justice, and protection for all people.
- **Universalism-nature**: Preservation of the natural environment.
- **Universalism-tolerance**: Acceptance and understanding of those who are different from oneself.

Conflicts arise when similarly-ranked values are revealed to be incompatible. For instance, consider how GADIN’s dilemma types can be modeled using this system. Betrayal and sacrifice may be conflicts between benevolence and a value of self-interest (such as security-personal). Similarly, greater good and take down scenarios may be conflicts between a value of self-interest (e.g., face) and the threat posed by the enemy (e.g., security-societal). A favor may be simply a conflict between relationship values, but it could also be one of competing Schwartz or individual values.

GADIN’s player trait variables may be represented with Schwartz values as well. Honesty and faithfulness are components of benevolence-dependability, while responsibility for actions aligns with conformity-rules and universalism-concern. Strength of character and morality are not formally defined, but are assumed to be measures of the user’s value ranking consistency, and the user’s alignment toward society’s value ranking, respectively. The latter is particularly abstract, and may simply be linked to the author’s idea of what morality means (which is likely a combination of conformity-rules, benevolence, and universalism).

Intriguingly, the Theory of Basic Human Values provides a means to create dilemmas beyond the capability of the GADIN system. An example might be the choice between helping any entity which supports one of your cherished values at the expense of yourself. This entity does not need to
Figure 1: Continuum of related motivations formed by the original ten universal values. The proximity between values indicates similarity of underlying motivations.

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betrayal</td>
<td>Help self at expense of friend</td>
</tr>
<tr>
<td>Sacrifice</td>
<td>Help friend at expense of self</td>
</tr>
<tr>
<td>Greater Good</td>
<td>Help self and enemy</td>
</tr>
<tr>
<td>Take Down</td>
<td>Destroy enemy at expense of self</td>
</tr>
<tr>
<td>Favor</td>
<td>Choose between two characters</td>
</tr>
</tbody>
</table>

Table 1: Dilemma types used in the GADIN system (Barber 2008).

be a friend or enemy; it may be an acquaintance, or not even a person. If we make use of the Theory of Basic Human Values, character relationships remain useful but are not always necessary to establish a dilemma, and a larger variety of personalized dilemma scenarios that resonate across cultures may be generated as a result.

**Approach**

**Overview**

This paper extends Barber’s dilemma categories by applying the Theory of Basic Human Values, expanding the definition of a cherished entity, and by providing a means for a character to be adverse to an entity, value, or action. The model presented here, EGAD (the Expressive Game Dilemma model), enables both encountered and observed dilemma generation, and thereby supports more believable, compelling, and useful story worlds.

Prior work has established that creative literary devices possess key properties, such as coherence, novelty, and usefulness (Newell, Shaw, and Simon 1962; Boden 1998; Harmon 2015). Building on these ideas, EGAD assumes that a successful literary dilemma should be:

- **Novel:** the generated scenario should be a fresh set of circumstances to ensure continued engagement. To ensure novelty, an EGAD system never uses the same exact scenario at the same level of tension twice. The same values may, however, be involved in multiple dilemmas.

- **Coherent:** the reader must be able to clearly understand the choices presented, and why the situation is difficult to resolve. EGAD systems build each dilemma by referring to universal values, and generate a maximum of four choices in each scenario to minimize cognitive load.

- **Apt:** the dilemma should appropriately fit the context within which it lives. An EGAD system builds each dilemma using input from the story world.

- **Meaningful:** the dilemma should reveal something important about the character making the choice. EGAD systems always generate choices between values. Once a decision is reached, the decider thus reveals what they truly care about.

Unpredictability, while not essential to a successful dilemma, may also be an important part of keeping literary devices like dilemma scenarios interesting (Barber 2008). By modeling hidden values and choices (described in the next section), EGAD provides the possibility of surprising both the person who encounters the dilemma, and anyone who observes the encounter.

A subset of unpredictability is suspense, or “uncertainty about a particular outcome (on the part of the audience), where the outcome is significantly desirable or undesirable” (O’Neill and Riedl 2014). An EGAD system generates dilemmas with regard to the current level of narrative tension and expected suspense. Uncertainty about individual factors (e.g., choice trust and giver trust) is also emphasized as an important part of the dilemma’s design.
The evil sorcerer laughs, and begins to sprint away. You look at Vance, and see that he’s trembling.

What do you do?
• Comfort Vance
• Pursue the sorcerer
• Wait and see what happens

Figure 2: Sample output from the implementation example.

Model

Definitions As in (Karsdorp et al. 2012), we will refer to agents that have intentions and can perform actions as actors. An entity is defined as any actor, prop, location, or concept within the context of the story-world. Let Z be our main character, an entity of type actor. Z will be the decisioner: the entity that resolves the dilemma.

Z may cherish or be adverse to values. Values include Schwartz values, a goal that Z wants to see achieved, or any entity. From the perspective of Z, each entity \( x_1, x_2, \ldots, x_n \) in the story-world (including Z) may embody any number of additional values.

We will refer to Z’s cherished values as \( I \), Z’s despised values as \( D \), and values of unknown importance to Z as \( U \). If the summed lengths of the lists \( I, D, \) and \( U \) are zero, Z does not value anything, and a dilemma cannot be generated for Z.

An action is a response by an entity to some initiating event, and a dilemma is when Z must choose between two or three proposed choices (consisting of one or more actions). When making a decision in a dilemma, any action by Z serves to preserve or reject a value or set of values, which may be in the form of an entity or another action. The net weight of an action (or choice) is the quantity which represents what is at stake and what is to be gained relevant to one or more deciders for that action (or choice). Z can perceive a different net weight than others, but might still account for the appraisal of others in the decision.

Tension is a measure of dilemma severity. The greater the tension, the higher the stakes of the choices. Over the course of a narrative, tension may rise and fall.

The giver is the entity who presents the dilemma, and giver trust represents the giver’s perceived dependability according to Z. Z’s perception of the giver’s dependability, of course, may be inaccurate. Hidden values are values which the giver hides from Z, which may be revealed later in the story. Hidden choices are options that may be hidden from Z or the reader of the story. Choice trust represents the expected probability that the net weight for a given choice is accurate. A negative, zero, or positive choice trust implies the perceivable net weight is actually lower, an exact match of, or higher than the actual value.

EGAD’s Dilemma Types There are three basic types of choice for a decider: help, harm, or ignore. A help (harm) choice supports (rejects) one or more values. The decider may also ignore/refuse the giver or the perceived choices.

Using these basic unit choices, we can build compound choices. For instance, if a dilemma consists of two help choices for distinct elements \( a, b \in L \), then Z must decide which of \( a \) or \( b \) to help. This may be analogous to Barber’s favor (Barber 2008). However, in EGAD, the method of help is another factor which may influence Z’s decision. If \( a = b \), Z must choose how to help something that Z cherishes. If \( a \neq b \), Z must choose who to help. How Z helps each character may be an additional factor in the decision.

If we assume dilemmas consist of two choices, the possibility space then becomes:

- Help, for \( a \in [L, U] \)
- Harm, for \( d \in [D] \)
- Ignore/refuse the giver or choices
- Choose which \( a \) to help, for \( a \in [L, U] \)
- Choose which \( d \) to harm, for \( d \in [D] \)
- Choose how to help \( a \), for \( a \in [L, U] \)
- Choose how to harm \( d \), for \( d \in [D] \)

Dilemmas consisting of three choices may be formed by combining a help (harm) compound choice with a harm or ignore (help or ignore) unit choice, or by combining any of the three unit choices with replacement. Dilemmas of four or more options may similarly be generated by combining these base choices. The ignore option must be used only once, if at all, in a generated dilemma with a single giver.

It is unlikely that Z would want to harm an element of \( L \) (or \( U \), depending on the context). It is also unlikely they would want to help an element of \( D \). However, some situations may require the decider to do so. In such forced scenarios, the decider may be faced with additional dilemma options:

- Harm, for \( a \in [L, U] \)
- Help, for \( d \in [D] \)
- Choose which \( a \) to harm, for \( a \in [L, U] \)
- Choose how to harm \( a \), for \( a \in [L, U] \)
- Choose which \( d \) to help, for \( d \in [D] \)
- Choose how to help \( d \), for \( d \in [D] \)

Forced choices are special, because a non-chaotic Z would ordinarily never pursue them. It would not be a dilemma if the decider simultaneously confronted forced and free (unforced) choices.

Lastly, it is important to note that these base components of dilemma may overlap. Helping a cherished entity, for example, may cause the implicit support of a despised value. A fully-developed implementation of the EGAD model should
Ideally acknowledge relations between actions, states, and events.

**Implementation Example**

It is possible to construct many dilemma generation systems which rely on the EGAD model. Each may have unique requirements, such as a constraint on which values a character may have to choose between over the course of a story. The following section demonstrates one such system, but does not attempt to encapsulate all EGAD-based systems.

In this example implementation, we will demonstrate a simple, generic dilemma generation system for a player Z. Dilemma information is encoded using a knowledge representation framework that enables actor reasoning over hypothetical timelines. Dilemmas are generated by way of three phases: actualization, elaboration, and realization.

**Actualization** To run, the generator must be provided with story context. This includes actor information and goals, current locations of all entities, information specific to the decider (such as the lists L, D, and U), as well as information related to the current tension and scene.

During actualization in this example, the system decides the actions Z may take in response to the dilemma, the entities and values that will be involved in the dilemma, and the giver. The giver is decided randomly between Z and any existing actor in the D list. If the giver is an element of D, the situation may be considered forced; otherwise, it is free.

This example system decides, at random, whether the dilemma will consist of two or three choices. The types of choices that can be generated depend on the size of L, D, and U. If all lists are empty, no dilemma is generated. If U and L are empty, no actions which involve elements of these lists may be taken (such as free help actions). If D is empty, no action which involves D elements may be taken (e.g., actions labeled as free harm). Choices are then selected based on the available action types, and populated with values for a or d at random.

**Elaboration** During the elaboration phase, we add detail to each choice. A choice need only involve one, but may include multiple actions. Any action that an entity can take should contain a list of associated consequences which specify the supported or rejected values. (A more complex system might incorporate branching consequences using a tree-like structure.)

The story context and the dilemma outline created during actualization indicate which actions might be chosen to respond to a new initiating event. For instance, for a decider to encounter a dilemma over having an affair, they must have at least one partner as a precondition.

Finally, this example system adjusts the severity of the dilemma to match the current narrative tension. Part of the story context may include a list of weighted Schwartz values. These weights are adjusted relative to some audience’s cultural context. As the tension rises, actions with values that increase in cultural importance may have a greater probability of being chosen. We can also change the stakes to match the tension. As an example, dilemmas of honesty might occur when tension is low, while life-or-death situations may be more likely to occur at the story climax.

**Realization** During realization, the generator constructs a full scene based on the generated dilemma information. Sample output for the system described here is shown in Figure 2, which describes a scenario with a friend (“Vance”) and an enemy (the sorcerer). This system relies simply on grammar-based text generation for realization, but more elaborate realizers are possible.

**Evaluation**

To analyze the expressive power of the EGAD model, two hundred dilemmas were identified in contemporary games. A random sample (n=20) of these dilemmas was chosen, which contained dilemmas from the following games: *Far Cry 2*, *Final Fantasy VI*, *Heavy Rain*, *Life is Strange*, *Mass Effect* (1, 2, and 3), *Oni*, *The Elder Scrolls V: Skyrim*, *Sudden II*, *The Walking Dead*, *The Witcher* (1 and 2), and *Under- tale*. Refer to Figure 3 for an example of a dilemma chosen in the random sample.

Ten participants who had attained at least a college-level education were recruited for the evaluation, and split into groups of five individuals each. Group 1 was trained to understand the GADIN model, while Group 2 was trained to understand the EGAD model of dilemma categorization. The judges each took an individual survey, in which they were asked if the model they learned could adequately describe each of the twenty dilemmas. Overall, the majority determined six of the twenty dilemmas could not be represented using GADIN, but found the EGAD model was sufficient in each case. More disagreement occurred between judges when trying to fit the GADIN model to each scenario. The pairwise percentage agreement values are shown in Table 2.

Certain scenarios were found to be especially difficult for participants when trying to fit the GADIN model. Most

<table>
<thead>
<tr>
<th>Judges</th>
<th>GADIN</th>
<th>EGAD</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1, 5)</td>
<td>0.60</td>
<td>1.00</td>
</tr>
<tr>
<td>(1, 4)</td>
<td>0.60</td>
<td>1.00</td>
</tr>
<tr>
<td>(1, 3)</td>
<td>0.40</td>
<td>0.85</td>
</tr>
<tr>
<td>(1, 2)</td>
<td>0.80</td>
<td>1.00</td>
</tr>
<tr>
<td>(2, 5)</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>(2, 4)</td>
<td>0.50</td>
<td>1.00</td>
</tr>
<tr>
<td>(2, 3)</td>
<td>0.50</td>
<td>0.85</td>
</tr>
<tr>
<td>(3, 5)</td>
<td>0.70</td>
<td>0.85</td>
</tr>
<tr>
<td>(3, 4)</td>
<td>0.50</td>
<td>0.85</td>
</tr>
<tr>
<td>(4, 5)</td>
<td>0.80</td>
<td>1.00</td>
</tr>
<tr>
<td>Average</td>
<td>0.59</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table 2: Percent agreement between each pair of judges, and the average pairwise percent agreement. Statistics such as Fleiss’ kappa could not be used in the evaluation due to high agreement between judges in Group 2.
Figure 3: This dilemma from *The Walking Dead* was chosen as part of the random sample. The player must choose whether to give or not give a gun to Irene, a non-player character. Giving Irene a gun fulfills her request and gives her the ability to kill herself, as she desires. It is also likely that she has been infected, and her being alive thus poses a threat to the group (although the main character seems to think there may be a chance she can get help). However, she appears to be emotionally unstable, and her actions are unpredictable. She may attract more enemies with the sound of the gunshot, or betray you and fire at you, or the group. You may choose to ignore the choice, but hesitation may further aggravate Irene. This is a complex scenario which forces the player to consider not just who, but how to help. It also emphasizes the need for modeling ignore actions, giver trust, and choice trust.

frequently, GADIN failed to represent dilemmas between methods, rather than characters. Other problematic scenarios occurred when a decision was being made about a character of unknown relationship value, or when GADIN utilities misrepresented or were insufficient in representing all possible positives and negatives in the story-world.

**Discussion**

EGAD is the first attempt at applying the Theory of Basic Human Values to dilemma generation. A random sample of twenty game dilemmas was used to compare the expressive power of the EGAD and GADIN models. Although preliminary, the results suggested that EGAD is able to represent a wider variety of player dilemmas than the GADIN system. The target application of the present study is narrative games, rather than interactive narrative generation. One direction for future work might be to determine how players perceive dilemmas generated by each model.

While Schwartz values have consistently been found to resonate across cultures, they may not necessarily contribute to all game dilemmas. There may also be game-specific values that appear only in the context of digital world interaction. (One example of a game-only value might be a player’s devotion to portraying a character accurately in a role-playing game.) Further, EGAD emphasizes values as primary components in dilemma categorization. In the future, EGAD might be extended to more thoroughly support additional dilemma components.

Previously, GADIN was used to generate adaptive dilemmas for players by simply incrementing or decrementing user traits after each decision was made. At present, it is unknown whether this procedure is sufficient to model dynamic player values. Some players may make more complex decisions, such as choosing an option only when certain values are activated together. While an improved adaptive dilemma generation system for players is beyond the scope of this paper, the results presented here and work in psychology theory suggests that modeling Schwartz and narrative game values may expand the domain of generated dilemmas and provide a higher confidence of universal applicability. Future work should explore implementations of this model, and build toward an improved understanding of the factors involved in game dilemmas.

**Acknowledgments**

Special thanks to Arnav Jhala for supporting the pursuit of this research.

**References**


Ware, S. G., and Young, R. M. 2014. Glaive: A state-space narrative planner supporting intentionality and conflict. In *Proc. of the 10th Conference on AI and Interactive Digital Entertainment (AIIDE)*.