A socio-emotional model of impoliteness for Non-Player Characters

Sabrina Campano
Université Pierre et Marie Curie
4, place Jussieu
75252, Paris, France
sabrina.campano@etu.upmc.fr

Nicolas Sabouret
Laboratoire Informatique de Paris 6
LIP6, 104 av du Président Kennedy
75016, Paris, France
nicolas.sabouret@lip6.fr

Abstract

Due to its important role in dialogue, politeness has been widely studied both from a theoretical point of view and at a practical level in virtual agents. However, few attention has been given to impoliteness, even though virtual characters, such as Non Player Characters in video games, should be able to use it to improve their believability. In this paper, we describe a model for utterance selection based on impoliteness that takes into account emotions, personality, and social relations.

1. Introduction

Some researchers have considered politeness as a key element in dialogue, and worked on this subject to improve the believability of virtual characters [1, 7, 18, 19]. However, no model has been proposed for impoliteness, despite virtual characters such as Non Player Characters in video games should be able to be impolite for a more believable behaviour. Impoliteness consists in performing verbal attacks towards an interlocutor [6], while politeness consists in minimizing verbal offenses towards an interlocutor [3].

Our work aims at providing a model of impoliteness for Non Player Characters in video games, in order to improve their believability. Non Player Characters are virtual agents controlled by the game, that interact with other NPCs, or with Player Characters controlled by humans. To fulfill this goal, we must take into account personality and emotions of a NPC, along with its social relations with the other characters in the game. Given an utterance set available for a turn taking of a NPC, we want him to select the appropriate utterance relative to impoliteness, according to the socio-emotional context. Our approach is mainly inspired from work in sociolinguistics, and behavioural studies on personality and emotions. Based on this work, we propose a set of rules to describe the socio-emotional influence on impoliteness, and we test this model through experimentation.

2. Related Work

Studies on politeness take their origins from the concept of “face” introduced by the sociologist Erving Goffman [9]. The face is a positive image that one wants to show to others. During an interaction, one of the goals of each participant is to preserve the face of their interlocutor, in order to pursue the interaction. Moreover a person can feel the emotions that he will cause to somebody [10], and human beings generally avoid to provoke negative emotions in their interlocutor because of this empathy.

Inspired by the work of Erving Goffman, Brown and Levinson went in detail with politeness strategies and produced a work still considered as a reference in politeness modelization [3]. They distinguished four super strategies, ordered by their risk of face loss for the interlocutor. The less polite strategy is doing a Face Threatening Act (FTA) “baldy”, that is to say without taking care of protecting interlocutor’s face, and the more polite strategy is doing a FTA “off record”, which corresponds to say something indirectly, by implied message. The level of politeness to use, which allows to select the appropriate strategy, is calculated from three factors : the power that the hearer has over the speaker, the social distance between the hearer and the speaker, and the rank of imposition of the FTA. The more the power, the distance, and the imposition rank are important, the more the speaker has to be polite.

In his work on impoliteness, Jonathan Culpeper refers to the social variables of Brown and Levinson’s model [6]. The greater the imposition of the speaker’s act, the more powerful and distant the interlocutor is, the more face-damaging the speaker’s act is likely to be. Moreover Culpeper mentions that an aggressive personality might reinforce impoliteness, and refers to a defensive reaction. This defensive reaction leads a speaker to be impolite if he is attacked by impoliteness. The work of Culpeper, along with Brown and Levinson’s model, led us to define the set of social variables we use in our model.

The defensive reaction mentioned by Culpeper is likely
to be caused by fear and anger emotions, because such emotions can lead to an aggressive behaviour [16]. Verbal aggression is a kind of aggression [4], and thus impoliteness, as far as verbal aggression, is concerned. This is the reason why the model described in this paper considers particularly fear and anger emotions at the core of impoliteness. This “defensive mechanism” is a key element for computing the NPC reaction to the impoliteness of its interlocutor.

Elisabeth Andrè et al. took into account the emotional state of the interlocutor to adjust the politeness level of their virtual guide [1]. The more the interlocutor is in a negative emotional state, the more the guide has to be polite. However this model is not sufficient for video games, since NPCs can also be impolite with their interlocutor.

One model tightly related to the emotional part of our problem is the one of Fleischman and Hovy [8]. In this model the words used by an agent corresponds to its emotional state provoked by the events of its virtual world. For example, “A smashed into P” does not express the same emotions than “A collided with P” : in the first one, there is a strong negative emotion directed towards A. Even if the researchers don’t mention it, this negative emotion can be related to impoliteness. If a speaker expresses his anger against someone, he risks making his interlocutor lose face. This approach led us to consider in our model the emotions that are susceptible to be provoked in the interlocutor by an utterance.

As Culpeper pointed out, personality might be related to impoliteness. An agreeable person produces efforts to maintain good social relations with others, and thus moderates the expression of his anger [12]. Agreeableness is one of the four major dimensions of personality structure [15]. Consequently, our model will focus on this personality trait, since it can prevents someone from causing negative emotions to others, and so from being impolite. Besides emotions, personality has yet been used to modulate goals and behaviour in a cognitive architecture [11].

Another dimension is susceptible to play a key role in impoliteness : the language register. A language register is a variety of language used in a specific social setting, and politeness relies in part on the correct selection of an appropriate language register according to the dialogue situation. Culpeper mentions that the use of taboo words (swear or abusive language) is an impoliteness strategy [6]. Therefore we think that the use of an inappropriate language register may be seen as impoliteness, although this statement is more or less relevant across different languages.

Under the light of this related work, our model takes into account social variables as dominance and liking, the personality trait of agreeableness, the fear and anger emotions, and language registers. All the traits considered have been chosen for their important influence on impoliteness, and their possibility to be associated with each others.

3. Proposed Model

3.1. Principle of the model

Our model is based on four major components : emotions, register, personality, and social relations.

**Emotions** : although no work seems to establish this direct connection, we make the hypothesis that spontaneous impoliteness, that is to say impoliteness which is not based on reasoning but on a fast processing, is provoked most often by fear and anger emotions. Indeed these two emotions, as defensive strategies, are at the origin of aggressive behaviour. We consider that impoliteness is equivalent to verbal aggression, which is a kind of aggression. Consequently in our model the more a character is angry or afraid, the more he can be impolite. Fear can also conduct to a second type of behaviour, not being impolite, when an aggressive behaviour is considered as too risky. It corresponds to the “fight or flight” reaction [5].

**Register** : we make the assumption that using a language register different from the expected register is impolite if in addition the speaker provokes interlocutor’s anger or fear. With fear and anger emotions, language register is the second component of impoliteness level in our model, and is both an output and input variable. Previous models on politeness used language register as an input variable [7], based on the fact that an interlocutor adapts to the formality detected in speaker’s utterances.

**Personality** : personality has an influence on emotions, and this problem has already been addressed [17]. However, we think that the agreeable personality trait has a direct influence on impoliteness, and this is why we take it into account in this model. As an agreeable person moderates the expression of his anger [12], we assume that this principle can be extended to impoliteness. Thus an agreeable NPC moderates its impoliteness level towards its interlocutor.

**Social Relations** : among the social variables involved in impoliteness, we selected the dominance that a speaker has over his interlocutor, and the liking of a speaker towards his interlocutor. The social distance is already represented through language register, since there is a language register imposed by the dialogue situation that takes into account the social distance. It seemed important to take liking into account, because it has an influence on anger [14], and we think that it plays an important role in the modulation of anger. Thus the more a speaker likes an interlocutor, the more anger’s anger is attenuated, and the more he dislikes an interlocutor, the more speaker’s anger is reinforced. When a speaker is strongly dominated by his interlocutor, dominance lowers speaker’s impoliteness in case of anger, or provoke the flight reaction in case of fear. If a speaker is strongly dominant, he can be impolite in a limited extent.

In this model impoliteness is triggered by fear and anger emotions, or depends on social relations. The higher the
intensity of anger or fear emotions felt by a NPC is, the more it provokes impoliteness from this NPC. When a NPC is impolite, he uses utterances that provoke anger or fear in the interlocutor, and a language register different from the expected register. It is possible that fear does not trigger impoliteness, since this emotion can induce a fight or flight reaction, which means that when threatened, a character can become offensive or try to escape [5]. If the language register used by a speaker towards a NPC is different from the language register expected, it reinforces the anger of the NPC. The principle of the model is illustrated in figure 1.

It is noticeable that this model is related with a fast, reactive processing, since individual traits and primary emotions triggers the interlocutor’s response. One might say that other factors than social relations, personality and emotions could trigger impoliteness, considering for example the case of a person being in a hurry. This situation can surely lead someone to be less polite, but we think it is not sufficient in itself to attack his interlocutor’s face.

3.2.2 Language Register

A language register is an element of the set \( \mathcal{R} = \{ \text{coarse, familiar, casual, consultative, formal, frozen} \} \). This set is inspired from Joos classification for the english language [13]. A total order relation \( \prec \) is defined on \( \mathcal{R} \), such as \( \text{coarse} \prec \text{familiar} \prec \text{casual} \prec \text{consultative} \prec \text{formal} \prec \text{frozen} \). We note \( R + n \) the nth successor of \( R \in \mathcal{R} \) w.r.t. the relation \( \prec \), and \( R - n \) the nth predecessor. If \( R \) has no successor, then \( R + n = R \), and if \( R \) has no predecessor, then \( R - n = R \). For example, \( \text{coarse} + 2 = \text{casual} \), and \( \text{coarse} - 1 = \text{coarse} \).

The register variables are : the natural character register \( NR_i \in \mathcal{R} \), which is the register that \( i \) uses the more often, the mastered character registers \( MR_i \subseteq \mathcal{R} \), representing the registers that \( i \) is able to use, and the imposed register for the dialogue situation \( IR(ds) \in \mathcal{R} \cup \{ \square \} \), which is the register participants should use if they want to be polite. The \( \square \) value represents the case there is no imposed register.

3.3. Expected Register Setting

Each character \( i \) NPC or PC has an expected language register for the dialogue situation, \( ER_i(ds) \in \mathcal{R} \), which is the register to be used by \( i \) in order to keep on being polite. This register is supposed to be known by \( i \) and by his interlocutor. \( ER_i(ds) \) is calculated from the imposed register \( IR(ds) \), the natural character register \( NR_i \), and the mastered character registers \( MR_i \).

If \( IR(ds) = \{ \square \} \), there is no imposed register, then \( ER_i(ds) = N R_i \). If \( IR(ds) \neq \{ \square \} \) there is an imposed register. In this situation there are two possible cases : i) \( IR(ds) \in MR_i \), the character \( i \) masters the imposed register, thus \( ER_i(ds) = IR(ds) \) ii) \( IR(ds) \notin MR_i \), the character \( i \) does not master the imposed register, thus the expected register is the closest register to \( IR(ds) \) that belongs to \( MR_i \). The order relation \( \prec \) is used to find the closest acceptable register.

For example, in a video game if a NPC which is a young man from New Jersey who wants to trade with a businessman from New York, the register imposed by the dialogue situation \( IR(ds) \) can be fixed to \( \text{formal} \). If the young man denoted \( i \) masters this register, then \( ER_i(ds) = \text{formal} \). Else, the young man is expected to use the closest register to \( IR(ds) \) among his mastered registers. If he masters \( \text{coarse} \), \( \text{familiar} \) and \( \text{casual} \) registers, then \( ER_i(ds) = \text{casual} \).

3.4. Computation of felt emotions

3.4.1 General Principle

When an utterance is pronounced by a speaker \( i \), it may cause emotions on his interlocutor \( j \) and modify his current felt emotions. As the model of Fleischman et Hovy is quite difficult to configure, we decided to tag utterances with an
emotional vector instead of tagging words. The emotional vector for the emotions that can be provoked by an utterance $u$ in $j$ is denoted as:

$$provoked_{emotions}_j(u) = \left( \begin{array}{c} fear_j(u) \\ anger_j(u) \end{array} \right)$$

as each vector value is included in $[0,1]$. The more the value is close to 1, the more the emotion is provoked by the utterance, the more the value is close to 0, the less the emotion is provoked by the utterance.

The way that felt emotions $anger_j(t)$ and $fear_j(t)$ are updated at turn $t$ depends on four elements:

- $anger_j(t-1)$ and $fear_j(t-1)$, the emotions felt by $j$ at turn $t-1$.
- $liking_j(t)$ and $fear_j(u)$, the emotions provoked by the utterance pronounced by $i$ intended at $j$.
- $liking_j(t)$ the liking of $j$ for $i$. The more $j$ likes $i$, the more the anger provoked in $j$ is lowered, the less $j$ likes $i$, the more the anger provoked in $j$ is aroused.
- $r_{diff}(R(u), ER_j(ds))$, a function that returns the rank difference between the expected register of $i$ $ER_i(ds)$ and the register of the utterance $R(u)$ that $i$ pronounced. For example, if $ER_i(ds) = casual$ and $i$ pronounced the utterance $u$ with $R(u) = familiar$, then $r_{diff}(R(u), ER_i(ds)) \rightarrow 1$. If $R(u) = coarse$, then $r_{diff}(R(u), ER_i(ds)) \rightarrow 2$.

3.4.2 Computation

Four weights are defined : $\beta, \gamma, \delta, \varphi \in [0,1]$, where $\beta$ is the weight for the emotions provoked by an utterance, $\gamma$ is the weight for the emotions of $j$ felt at turn $t-1$, $\delta$ the weight for the liking of $j$ for $i$, and $\varphi$ the weight associated to the difference between the register expected from $i$ and the register he really used. A simple linear function $d(i)$ is also used to represent the progressive decrease of emotions intensity when they are not aroused.

The computation of felt emotions is the following:

- if $fear_j(u) = 0$ then $fear_j(t) = d(fear_j(t-1))$,
  else $fear_j(t) = \gamma fear_j(t-1) + \beta fear_j(u)$.
- if $anger_j(u) = 0$ then $anger_j(t) = d(anger_j(t-1))$,
  else $anger_j(t) = \gamma anger_j(t-1) + \beta (anger_j(u) - \delta liking_j(t)) + \varphi r_{diff}(R(u), ER_i(ds))$.

Felt emotion intensity is bounded from 0 to 1.

For example, when a speaker $i$ pronounce an utterance that does not provoke fear or anger in interlocutor $j$, current fear or anger of $j$ decrease if they were aroused by previous events. If the utterance pronounced by $i$ provokes fear or anger in $j$, then the current emotions of $j$ are updated taking into account emotions provoked by the utterance. The more $j$ dislikes $i$ and the more the register of $u$ is different from the expected register of $i$, the more the anger of $j$ is reinforced.

3.5. Utterance selection

3.5.1 Starting utterance set

In response to an utterance two kinds of behaviours are possible: emotional reactions which are spontaneous, and planified reactions which are more related to a deliberative processing. The architecture of the DEEP project contains an inference engine that returns all the possible responses that correspond to a reactive or a planified processing [2]. Our model treat the emotional part of the selection. This selection is made among the set $U_i(t)$, which represents the utterance set proposed by the inference engine for the NPC $i$ to address the character $j$ at turn $t$.

3.5.2 Definition of Impoliteness Couple

The impoliteness level computed for a NPC $i$ that talks to $j$ at turn $t$ is represented by an impoliteness couplet $IC_{i,j}(t) = (R \in R, a \in [0,1])$. The difference between $R$ and $ER_i(ds)$, along with the intended anger $a$ at the interlocutor $j$, define the impoliteness level of $i$. This impoliteness couplet allows to select an utterance among $U_i(t)$. This utterance selection can be the result of a socio-emotional process or a voluntary intention from $i$.

Given an utterance $u$ pronounced by $i$, the stronger the provoked anger of $u$ is, and the greater the difference between the register of $u$ and the expected register $ER_i(ds)$ is, the more the impoliteness of $i$ is important.

3.5.3 Construction of Impoliteness Couplets

Speaker behaviour depending on his emotions, personality, and social relations Behaviours listed above correspond to hypotheses made for the model.

**Personality**
- agreeable personality trait : an agreeable character $i$ intends to avoid interlocutor’s anger because he wants to preserve his social relations [12]. Consequently, he tries to stay close to $ER_i(ds)$, and to not use an utterance provoking too much anger in the interlocutor.

**Social Relations**
- $i$ feels dominated by $j$ : a dominated character $i$ intends to avoid interlocutor’s anger because if he does not he can get in trouble. Consequently, $i$ tries to stay close to $ER_i(ds)$, and to not use an utterance provoking too much anger in the interlocutor $j$.
- $i$ feels dominant over $j$ : a dominant character $i$ doesn’t have to take too much care on his politeness towards $j$. He can use a register directly superior to $ER_i(ds)$, or directly inferior to $ER_i(ds)$, and use utterances that provoke a limited anger.
Emotions

\[\text{anger} : \text{if a speaker is angry, he has tendency to use a register inferior to } ER_i \text{ and to pronounce utterances that provoke interlocutor’s anger.}\]

\[\text{fear} : \text{if a speaker is afraid, he can stay polite in case he feels dominated, or respond with verbal aggression.} \]

Either he uses \(ER_i(ds)\) and intends to not provoke interlocutor’s anger, or he uses a register inferior to \(ER_i(ds)\) and intends to provoke interlocutor’s anger.

Algorithm The set of impoliteness couplets \(ICS(t)\) is initialized with \(IC_{i,j}(t) = (ER_i, 0)\). This couplet is neutral, that is to say not impolite. Then, \(ICS(t)\) is subject to modifications depending on thresholds fixed upon emotions, personality and social relations. When \(IC_{i,j}(t) = (R, a)\) is not a neutral couplet (i.e. the register \(R\) is different from the expected register, or the intended anger \(a\) is superior to 0), it corresponds to an impolite behaviour, of which level is defined by the algorithm below.

Step 1: Anger and Fear. This step constructs an impoliteness couplet depending on anger or fear emotion felt by a speaker \(i\). This is the emotion with the higher intensity which is taken into account. The more the emotion intensity is important, the more the register is shifted and more the intended anger is important. For example if the neutral couplet is \(IC_{i,j}(t) = (\text{casual}, 0)\), a strong anger intensity equal to 0.8 will create an impoliteness couplet \(IC_{i,j}(t) = (\text{coarse}, 0.8)\).

At the end of this step, if \(\text{anger}_i(t) < \text{fear}_i(t)\), the new impoliteness couplet \(IC_{i,j}(t)\) is added to \(ICS(t)\) which contains yet a neutral couplet. Otherwise, \(ICS(t)\) is reinitialized with the new impoliteness couplet \(IC_{i,j}(t)\). Indeed, fear is associated with two behaviours, ‘be not impolite’, or ‘attack verbally’, while anger is only associated with the ‘attack verbally’ behaviour.

Step 2: Agreeableness. Depending on agreeableness, this step modifies the impoliteness couplet of Step 1. An agreeable character moderates the expression of his anger to maintain good social relations, thus he lowers his impoliteness. A threshold was defined to estimate the agreeableness level: \(a_i^{\text{agreeableness}} \geq a_i^{\text{agreeableness}}\), \(i\) is very agreeable, and a function reduces the intended anger and the register difference.

Step 3: Dominance. This step constructs or modifies impoliteness couplets depending on dominance. As for agreeableness, thresholds were defined to estimate the dominance level. In case \(i\) is very dominant over \(j\), is unaffected by his fear or anger emotions, and is not particularly agreeable, he can be impolite in a limited extent if he wants to. He can use a register directly superior or directly inferior to his expected register, and can have a limited intended anger at \(j\). These impoliteness couplets are added to \(ICS(t)\). If \(i\) is very dominated by \(j\) and his fear felt emotion has created an impoliteness couplet, then this impoliteness couplet is removed from \(ICS(t)\). This will conduct \(i\) to keep on being polite with his interlocutor. If \(i\) is very dominated by \(j\) and his anger felt emotion has created an impoliteness couplet, then the impoliteness level of this couplet is lowered in the same way that in step 2.

At the end of the three steps described above, the set of impoliteness couplets \(ICS(t)\) is constructed. To use a register a NPC \(i\) has to master it. Consequently for each impoliteness couplet \(IC_{i,j}(t)\), \(i\) is not included in the speaker mastered registers \(RM_i\), then \(R\) is replaced with the closest mastered register among \(RM_i\).

3.5.4 Utterance selection

The utterance selection is made among \(U_{i,j}(t)\), which contains the utterances returned by the inference engine. It is supposed that \(U_{i,j}(t)\) contains at least one utterance with a register contained in an impoliteness couplet \(ICS(t)\). This guarantees that there is at least one utterance that will be returned by the function \(\min(U_{i,j}(t), ICS(t))\) corresponding to the Algorithm 1 below. This function selects utterances \(u \in U_{i,j}(t)\) that are the closest to a couplet \(IC_{i,j}(t) \in ICS(t)\). It means that the function selects the utterances \(u\) where for a given \(IC_{i,j}(t) = (R, a)\):

\[- R(u) = R, \text{ the utterance register is equal to the register of the impoliteness couplet } IC_{i,j}(t).\]

\[- \text{the distance between the utterance provoked anger and the impoliteness couplet intended anger } | \text{anger}_j(u) - a | \text{ is the minimum distance for all the couplets } IC_{i,j}(t) \in ICS(t) \text{ and for all the utterances } u \in U_{i,j}(t).\]

Algorithm 1 Min Function

Require: \(U_{i,j}(t), ICS(t)\)
initialize best
for all \(IC_{i,j}(t) = (R, a) \in ICS(t)\) do
for all \(u \in U_{i,j}(t)\) do
if \(R(u) = R\) then
\(d = \text{anger}_j(u) - a\)
if \(d < \text{best}\) then
\(\text{selected}_u = u\)
\(\text{best} = d\)
return \(\text{selected}_u\)

If more that one utterance is selected by the \(\min\) function, an utterance is chosen randomly among this set. The final utterance will be pronounced by the NPC \(i\) to the character \(j\) at turn \(t\).
4. Conclusion

This work is a first step towards the definition of a model for impoliteness that depends on emotions, personality, and social relations. The model was implemented in Java, and our first evaluation on ad-hoc examples showed coherent behaviors. Fear can provoke impoliteness or provoke the ‘be not impolite’ behavior depending on the setting of social variables and personality, and the expression of anger is modulated by these variables. Our next goal is to validate the impact of our model on NPC believability in a real video game. This work is under progress. We are currently integrating our model in the scene-definition platform in the DEEP project and the next step is to have this tested and used by game designers.

Moreover our model open future prospects, and we want to add other aspects related to previous works. Brown and Levinson’s politeness strategies [3], at the core of most of the work on politeness modelization, have to be taken into account, as the work of Culpeper concerning impoliteness strategies [6]. We plan to extend our model to politeness and to integrate strategies in it.

Initially our model focused on primary emotions as fear and anger, but the range of human emotions is much more rich, and other emotions play a important role in impoliteness as frustration, disgust, contempt, and boredom. The next version of the model should include these secondary emotions.

Another problem that has to be addressed is that manually tagging utterances is quite difficult to implement, and does not provide a dynamic frame that can adapt to contextual changes. Nevertheless, the computation of the emotional impact of an utterance is quite challenging, and many models still use manual tagging.

Currently, the utterance selection depending on the goals of the speaker is supposed to be held by an inference engine. We want to work on this inference engine, since it is possible to do meta-reasoning on emotions and social relations. For example, if a speaker wants to be liked by his interlocutor, knows that an utterance will cause joy emotion in the interlocutor, and knows this will lead the interlocutor to like him, he will use this utterance to achieve his goal of being liked. The inference engine should also determine if fear induce a flight or fight behaviour. The current model only uses dominance to this purpose, but this is not a sufficient criteria. We plan to study emotional responses and their impact on utterance selection in a more detailed way.

5. Acknowledgments

We thank Vincent Corruble and Magalie Ochs for their contribution to our work, and Florent Boutin, Christophe Gohmann and Karim Sehaba for their work on the DEEP project.

References