A Formal Model of Social Relations for Artificial Companions

Florian Pecune, Magalie Ochs and Catherine Pelachaud

CNRS-LTCI, Telecom ParisTech
{pecune, ochs, pelachaud}@telecom-paristech.fr

Abstract. For virtual agents to be implied in a long-term interaction with users, they have to be able to establish and maintain long-term relationships with them. In order to give the capability to a virtual agent to reason on its social relation with a user, we propose a BDI formalization of social relations. Based on psychological theories, we characterize social relation over three dimensions, namely dominance, liking and familiarity. Liking (of the agent toward a user) is defined as a like-mindedness between the agent and a user in terms preferences. The degree of liking can be computed from the number of shared preferences and their related importance for the agent. The agent's dominance is defined as its power on the other's intentions achievement. The level of dominance depends also on the other's power. Finally, the agent's familiarity is defined by the value of intimacy with a user and relies on the quantity of knowledge about a user's beliefs and intentions. Based on the formal representation of social relation, the virtual agent may reason about the different strategies to adopt to deal with its relation with another agent (virtual or real).

Keywords: Virtual agents, artificial companions, social relations, BDI agent

1 Introduction

Machines nowadays own a prominent place in our daily lives and the amount of research projects featuring artificial companions has significantly increased over the past ten years. A companion can be defined as "*a robot or a virtual conversational agent that possesses a certain level of intelligence and autonomy as well as social skills that allow it to establish and maintain long-term relationships with users*" [1]. The growing interest in artificial companions may be understood in the light of the works in Social Psychology. Indeed, many theories emphasize the need for affiliation and its influence on human behavior [2]. Several works also highlight the impact of social relationships on the well-being of individuals [3]. If virtual companions are obviously not supposed to substitute human companions, they can still compensate the lack of social relations in particular situations. This can be, for example, the case for isolated seniors [4] in order to assist them in daily tasks, or simply to have small talks with them [5]. Studies conducted by Turkle [6] also show that users are becoming more comfortable with the idea of interacting with these kinds of agents. To create believable interaction between an artificial companion and a user, the social relation of the companion toward the user should reflect the social context of the interaction. Indeed, the companion's reactions should be different depending whether it interacts with a user to play a game or to teach a lesson of mathematics. These reactions should also reflect what happened during the interaction (for instance if the user abuses the companion). In order to give the capability to an artificial companion to reason about its own social relation toward the user, we propose a BDI-like model of social relations. Several dimensions of the agent are defined as particular combinations of beliefs and intentions. Moreover, functions to compute the values of the intentions are proposed. From this formalization emerge strategies that an agent can plan to try to modify its social relation with the user. The article is organized as follows. In Section 2, we present the related works in the field of artificial companions and relational agents. In Section 3, the formalization of the different dimensions of a social relation is introduced. The formal model is illustrated on a scenario Section 4. We conclude and discuss future work in Section 5.

2 Related Work

2.1 Artificial companions and relational agents

Started in 2004, the *Cogniron* project is one of the pioneers in this field [7]. The idea was to create a robot able to understand user's needs, and to assist him in various tasks. The robot must be able to perceive its environment, to analyze the various constraints inherent in achieving its task, and then to decide the best strategy to adopt. However, by focusing on the cognitive aspect of the companion, the relational part has been missed. The robot is finally more an assistant than a real companion. The COMPANIONS project is more focused on the relationship between the user and the companion. The purpose is to encourage users to converse regularly with a virtual agent by telling it his day [8]. The developed model allows the agent to respond to emotional cues provided by the user, making the interaction more believable. The LIREC project focused on long-term interactions by modeling agent's memory and the evolution of dynamic relationships. The objective of this project was to develop a companion able to recall past interactions and adapt its behavior accordingly [9].

In the domain of virtual agents, some research has focused on relational agents, which aim at building relationships over several interactions with their users. However, they are mainly developed for a specific purpose and in a specific context. In [10] the authors propose an overview of the different areas in which relational agents could be used. Most of them are used in applications related to health and behavioral change, but they can also be found in leisure and domestic applications, or even in video games. Some relational agents have also been developed. For instance, Laura [11] encourages users to exercise on a daily basis. Laura's behavior evolves over the everyday interactions, through pre-scripted dialogues. The relationship is based on a stage model, which could be related to the notion of intimacy.

Besides those research projects, there are many companions who have been or are still marketed today by industry. In addition to the *Tamagotchi*, there are robots like *Pleo* the dinosaur or *Furby*, responding to voice activation and learning behaviors over time. The virtual dogs *Aibo* and *Genibo*, as well as the virtual cat *NeCoRo*, try to mimic the behavior of the animals they represent. If *NeCoRo* is only able to purr and respond to cuddles, *Aibo* and *Genibo* are able to move and detect obstacles. Finally, the autonomous robot *PaPeRo* offers many interesting features such as facial recognition and voice synthesis, making interactions with the user more enjoyable. Although they are made for long-term relationships, few of these companions integrate an explicit social model for their relation with the user.

2.2 Social agents

Unlike relational agents and artificial companions, social agents do not focus on modeling long-term relationships. However, some of them model social relations and their dynamics. This is particularly the case of Avatar Arena, a scenario in which a user must negotiate a schedule appointment with several agents [12]. Before each session, the value of appreciation between agents is fixed (low, medium, high), and it evolves according to the Congruity Theory described by Osgood and Tannenbaum. This theory can also be found in FAtiMA [13], a computational model used in projects such as *FearNot!* or ORIENT. In Demeanour [14], relation also relies on the axes of appreciation and dominance. The relationship evolves according to the personality of the agent, its mood, and the social attitudes expressed by the other party. This relationship can influence the non-verbal behavior of the agent, particularly in terms of posture and gaze. In SCREAM, emotions play an important role, changing the relationship according to their type, as explained in [15]. The representation of social relation in SCREAM is based on the same two axes of appreciation and dominance: the social power represents the hierarchical relationship between two people and the social distance, the degree of appreciation. Prendinger adds to these two axes the notion of familiarity, also based on emotions, but evolving monotonically. Although it seems obvious that an agent's relationship toward a user will influence the agent behavior, there are few models considering the agent's belief of the supposed relationship of the user toward itself. PsychSim attempts to fill this gap by providing a framework in which the agent will take into account the others' relations before planning their own actions [16]. Finally, Alpha Wolf gives us the opportunity to play a Cub [17], and interact with other wolves through howls, growls and whines. The dimension of dominance is the only one used here, and it just changes the animation of our cub depending on which wolf it is facing.

Among these models, some of them formalize social relations using logical concepts. In [18], the authors try to team up humans with a group of synthetic characters, using a formal representation of liking and dominance. In [19], the author formalizes the five bases of power described by Raven [20] with four different categories. The author also models the agent's decisions, knowing the power relations in which he is involved. Logical formalization is also used to model emotions. Based on the work exposed in [21], Dastani and Lorini propose a formal model of emotions, and define different coping strategies [22]. Some computational models have been proposed to compute the social relations of an agent and to simulate the impact on their nonverbal behaviors. However, few of them integrate strategies that the agent can plan to try to modify its social relations.

3 Formalization of Social Relations

To give the capability to a companion to reason about its own relation, we formally describe a social relation. In this section, after introducing the dimensions used to model companions' relation, we present their formal representation.

3.1 Dimensions of a Social Relation

Social relations can be represented by different dimensions (or axes). Although the most widely used dimensions refer to notions of dominance and liking, it is important to note that definitions of each of these dimensions vary depending on the model. In our model, we define liking as "*a general and enduring positive or negative feeling about some person*" [23]. This notion is asymmetric [24,15] -so not necessarily reciprocal. However, in [14] as in [25], this notion is mixed with commitment. On dominance, there are also different interpretations. In [14] and [15] dominance is defined as a hierarchical framework. In our work, we rely on [24] to define this dimension as "*the capacity of one agent to affect the behavior of another*". This influence is itself characterized by the resources and strategies available to the agent [20].

Dominance and Liking are not sufficient to cover the whole domain of social relations. Svennevig sets a list of dimensions related to this field [26]. As mentioned above, liking is sometimes mixed with commitment [14,25], which represents the need to have a close relationship. Familiarity (or intimacy) describes the way in which the relationship evolves over time [5]. Another important notion, especially in close relationships, is trust[27]. Based on [5], trust might be defined as a combination of affect, familiarity and solidarity. Thus, as a first step, we will not consider trust as a dimension. In the case of artificial companions, we assume that the agent is committed and connected to the user. Thus its level of commitment is fixed in our model. On the other hand, familiarity may vary. In our work we consider the definition of familiarity as the "*mutual knowledge of personal information*" [26]. Therefore, as a first step, we consider in our model the three following dimensions: **dominance**, **liking** and **familiarity**.

3.2 The logical framework

Let $\neg$, $\wedge$, and $\vee$ represent respectively logical negation, conjunction and disjunction, $\exists$ and $\forall$ existential and universal quantifier. Let the symbols i, j and k

represent three different agents (virtual or real), ϕ and ψ logical formulas, E a set of events with e and e' part of this set. The agent's mental attitude of belief and intention are formalized respectively by the modal operator B and I.

- $B_i(\phi)$ can be read as the agent i believes that ϕ is true.
- $I_i(\phi)$ means that the agent i has the intention ϕ. That means that the agent i thinks ϕ is currently not true, the agent i desires persistently (until the achievement of ϕ or until the agent thinks that it is impossible to achieve) that ϕ will be achieved in the future, and the agent i desires that all the plans allowing to achieve ϕ will be realized. This definition of intention and belief is based on the formal theory of interaction proposed in [28].

The operators $Feasible$, $Possible$, $Before$ and $Agent$ are inspired by [28]

- $Feasible(e, \phi)$ means that the event e might happen, and ϕ will be true after that.
- $Possible(\phi) \equiv \exists e \in E, Feasible(e, \phi)$ means that there is at least one existing event e, and ϕ will be true after that.
- $Before(\psi, \phi)$ means that ϕ cannot be true if ψ is not true.
- $Agent(i, e)$ is true if and only if the agent i can be the author of the event e.

The operators are based on a Kripke's possible-world semantics with, for each operator, an accessibility relation [29]. Based on this logical framework, in the next sections, we propose a formalization of the following dimensions of a social relation: dominance (Sect. 3.3), liking (Sect. 3.4) and intimacy (Sect. 3.5).

3.3 Dominance

The theoretical background for the formalization of the dominance is based on the work of Emerson [30], and more particularly on his definition of dependence.

Dependence In the work of Emerson [30], the dependence of an agent A upon another agent B is "*(1) directly proportional to A's motivational investment in goals mediated by B and (2) inversely proportional to the availability of those goals to A outside of the A-B relation*". The motivational investment corresponds to the importance accorded by an agent i to an intention ϕ for which another agent j has an influence (positive or negative). For instance, an agent i has the intention to go to a party. Since he has no way to go there, his friend j offers his help to bring him to the party. The dependence level of i upon j depends on the importance accorded by i to attend this party. The availability of the goal, in the definition of Emerson, corresponds to the number of agents k different from j that also have a positive influence on the intention ϕ. For instance, if the agent i has another friend k who can lend him his car to go to the party, i will be less dependent toward j since he has an alternative solution. Thus, the first step to model dependence is to represent the way an agent can influence another agent's intentions. We define two cases, represented by two formulas: an agent can be helpful (positive influence) or threatening (negative influence). An agent i is helpful for another agent j if j has an intention ϕ and i can help j to achieve this intention. The achievement can formally be represented by an event e after what ϕ will be true. We formally define a helpful agent i for an agent j concerning an intention ϕ by the syntactic abbreviation:

$$Helpful_i(j, \phi) \equiv I_j(\phi) \wedge \exists e \in E(Feasible(e, \phi) \wedge Agent(i, e)) \qquad \text{(Def.Help)}$$

An agent i can also be threatening for an agent j if j has an intention ϕ and i can prevent an intention ϕ. Preventing an intention can formally be represented by (1) an event e after what ϕ will never be true or (2) an event e canceling a precondition ψ of ϕ. A threatening agent i for an agent j concerning an intention ϕ is formalized by the following syntactic abbreviation:

$$\begin{aligned} Threatening_i(j, \phi) \equiv & I_j(\phi) \wedge \exists e \in E(Feasible(e, \neg Possible(\phi)) \wedge Agent(i, e)) \\ & \vee \exists e' \in E(Feasible(e', \neg\psi) \wedge Agent(i, e') \wedge Before(\psi, \phi)) \end{aligned} \qquad \text{(Def.Threat)}$$

The sub formula $Feasible(e, \neg Possible(\phi)) \wedge Agent(i, e)$ means that the agent i may act to make the intention ϕ of j impossible to achieve. For instance, i can cancel the party, making the intention impossible to achieve. The sub formula $Feasible(e', \neg\psi) \wedge Agent(i, e') \wedge Before(\psi, \phi)$ means that the agent i may act to make a precondition to achieve ϕ false. For instance assuming that i gave his consent to j to go to the party, i can reconsider his decision. An agent j is **dependent** upon an agent i toward an intention ϕ if j believes that i can help to achieve an intention ϕ he cannot achieve by himself or prevent him to achieve this intention. Formally, the dependence of an agent j upon another agent i toward an intention ϕ is defined by the following syntactic abbreviation:

$$\begin{aligned} Dependent_j(i, \phi) \equiv & B_j(Helpful_i(j, \phi) \wedge \neg\exists e' \in E(Feasible(e', \phi) \wedge Agent(j, e'))) \\ & \vee B_j(Threatening_i(j, \phi)) \end{aligned} \qquad \text{(Def.Depend)}$$

The sub formula $\neg\exists e' \in E(Feasible(e', \phi) \wedge Agent(j, e'))$ means that the agent j cannot achieve the intention ϕ by himself. For instance, if the agent j gets his own car to go to the party, j will not be dependent upon i anymore.

The dependence value As explained in Sect. 3.3.1, Emerson quantifies the level of dependence of an agent i toward an intention ϕ according (1) to the importance accorded by i to this intention and (2) the number of agents having a positive influence on this intention. We then introduce the importance accorded to an intention by an agent. Based on [31], we distinguish the importance accorded by i to the success of the intention ϕ and the importance accorded by i to the failure of ϕ, formalized as follows :

$$imp\_succ_i(\phi) \in]0;1] \quad and \quad imp\_fail_i(\phi) \in]0;1]$$

A high value means that i accords a high importance to the intention ϕ. The importance to achieve an intention and the importance not to have this intention failed might be equal or negatively correlated. For instance, a student thinking that successfully passing his exams is a formality will give a low importance to graduating, while according a high importance to failing his exams. We suppose that these values are predefined for the agent, but they can change according to the situation. We also define a function representing the amount of agents susceptible to have a positive impact on this intention ϕ, using the following syntactic abbreviation:

$$helpers\_num(i, \phi) = card\{k \neq i | Helpful_k(i, \phi)\}$$

This function represents the number of agents k different from i able to achieve i's intention ϕ. For instance, if i has the intention to go to a party, the value depends on the number of agents that can bring him to the party. The dependence value of i upon j toward his intention ϕ depends if j is helpful or threatening. If i thinks that j can help him to achieve his intention ϕ (i.e. $Helpful_j(j, \phi)$) then the dependence value of i depends on the importance of success accorded to ϕ (i.e. $imp\_succ_i(\phi)$) and the number of potential helpers (i.e. $helpers\_num(i, \phi)$). Indeed, the more people might help i, the less j is powerful. Consequently, we propose the following function to compute the dependence value upon a helpful agent:

$if\ B_i(Helpful_j(i, \phi))$

$$then\ depend\_value_i(j, \phi) = \frac{imp\_succ_i(\phi)}{helpers\_num(i, \phi)} \quad \text{(P.Help)}$$

$with\ depend\_value_i(j, \phi) \in]0; 1]$

For instance, an agent i according a high importance attending to a party will be very dependent upon j if he is the only one able to bring i to this party. If i thinks that j can prevent him to achieve his intention ϕ (i.e. $Threatening_i(j, g)$) then the dependence value of i depends directly on the importance of failure accorded to ϕ (i.e. $imp\_fail_i(\phi)$). Formally, the dependence value of a threatening agent is defined as follows:

$if\ B_i(Threatening_j(i, \phi))$

$$then\ depend\_value_i(j, \phi) = imp\_fail_i(\phi) \quad \text{(P.Threat)}$$

$with\ depend\_value_i(j, \phi) \in]0; 1]$

For instance, an agent i will be a little dependent upon an agent j if j can cancel the party but i does not care so much about missing the party. Whereas the dependence value is defined for a specific intention, we introduce the mean dependence value to represent the global dependence of an agent upon another considering all the agent's intentions. Formally, the mean dependence value of an agent i over an agent j is defined as a mean of the dependence value for each intention of the agent.

$$mean\_depend\_value_i(j) = \frac{sum\{depend\_value_i(j,\phi)|I_i(\phi)\}}{card\{\phi|Helpful_j(i,\phi) \vee Threatening_j(i,\phi)\}}$$
$$with\ mean\_depend\_value_i(j) \in]0;1] \quad \text{(F.Dep)}$$

The sub formula $card\{\phi|Helpful_j(i,\phi) \wedge Threatening_j(i,\phi)\}$ represents the number of i's intentions for which j has an influence. $sum\{depend\_value_i(j,\phi)|I_i(\phi)\}$ represents the sum of the dependence values of i for each intentions ϕ. In some particular cases (i.e. when the standard deviation of dependence values is high), this mean value is replaced with the highest dependence value. For instance, if an agent i is very dependent upon j concerning a particular intention, but is just a little dependent concerning other intentions, his dependence value upon j will stay quite high.

Dominance In his work, Emerson [30] defines the power of an agent A over an agent B as *a potential influence* depending on A's dependence upon B and B's dependence upon A. Indeed, an agent i thinks that he is dominant toward another agent j if he thinks that he is less dependent than j (i.e. if i thinks that he can help or threaten the other agent j's intentions more that the latter can on i's intentions). Formally, an agent i is **dominant** toward an agent j if and only if i thinks he is less dependent upon j than j is upon him :

$$Dominant_i(j)\ iff\ B_i(sup(mean\_depend\_value_j(i), mean\_depend\_value_i(j)))$$
$$with\ sup(a,b) = true\ iff\ a > b \quad \text{(Def. Dominant)}$$

To compute the dominance value, we introduce the following function:

$$dom\_value_i(j) = mean\_depend\_value_j(i) - mean\_depend\_value_i(j)$$
$$with\ dom\_value_i(j) \in]-1;1[\quad \text{(F.Dom)}$$

Consequently, a negative value of dominance means that the agent thinks he is submissive whereas a positive value traduces a belief on his dominance over the other agent. Since the dominance value is based on beliefs of the agent (the agent i thinks he is less dependent than the agent j), an agent i might be dominant toward an agent j, and j might also be dominant toward i.

Strategies Based on this formalization of dominance, several strategies may be defined for an agent i to try to change his value of dominance on another agent j. The following strategies that we can infer from the proposed formalization are consistent with those described by the psychologist Emerson [30]

1. Emerson highlights that to increase dominance, one may change his motivational investment, which is the importance of an intention. Given our proposed formalization, the agent i can decrease the importance accorded to

one of his intentions that the agent j may achieve or threat. So, given that the dependence of i depends on the importance of the intention (P.Help and P.Threat), the total dependence of i will decrease (F.Dep) and the dominance of i on j will increase (F.Dom)

2. Emerson claims that one may decrease his dependency by finding other sources. In our formalization, the agent i can try to find another agent that provides him some help; indeed, in this case, the number of helpers will increase and then the dependence value of j will decrease (P.Help).
3. Emerson affirms that one might try to increase the motivational investment of another person. Given our formalization, the agent i can also try to induce a higher importance for a particular intention to the agent j on which the agent i has an influence. Increasing this importance will lead the agent j to be more dependent upon i and then i will be more dominant.
4. Another strategy highlighted by Emerson consists in denying alternative sources for the other agent. In our formalization, the agent i can try to prevent other agents to help j achieving an intention for which he can be helpful. Lowering this value will lead j to be more dependent upon i and then i will be more dominant.

In the two first strategies, the agent will modify his own mental state, whereas he will try to influence the other agent in the two following ones. Using this formalization, we can also define two other strategies, consistent with Raven's work on social power [20]. In those strategies, the agent i can try to influence the agent j beliefs and intentions:

5. Raven claims that reward power "*stems from the ability of the agent to offer a positive incentive, if the target complies*". In our formalization, the agent i can inform the agent j that he can help j to achieve one of his intentions (Def. Helpful). The agent j will be more dependent upon i, depending on the importance of success j accorded to the intention and the number of other agents able to help j.
6. Raven affirms that in coercive power, "*the agent brings about change by threatening the target with negative, undesirable consequences*". Given our formalization, the agent i can try to inform the agent j that he can prevent j from achieving one of his intentions (Def. Threatening). The agent j will be more dependent upon i, depending on the importance of failure j accorded to the intention.

3.4 Liking

The formalization of liking is rooted in the work of Heider and his proposed Balance Theory [32]. This theory can be represented as a triangular schema between a person P as the focus of the analysis, another person O and an impersonal entity X, which could be a physical object, an idea or an event (Fig.1). The edges of this triangular schema describe liking relations between the three vertices P, O and X: P's liking toward the agent O, P's liking toward the concept X and O's liking toward the same concept X.

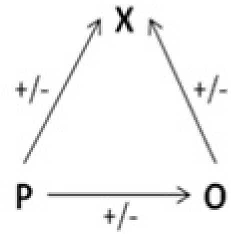

Fig. 1. Heider's Balance Theory

Sharing appreciation According to [33], the state between three entities is called balanced "*if all three relations are positive in all respects or if two are negative and one positive*". For instance, taking an agent i, another agent j and a concept X, the state between i, j and X is balanced if i likes j, i likes X and j likes X; or if i likes j, i dislikes X and j dislikes X as well. Since the Balance Theory assumes that people tend to seek balanced states, we define two scenarios: (1) if i and j share the same appreciation about X, then i likes j. (2) If i and j have contradictory appreciations about X, then i will dislike j. The first step to model liking between i and j is to formalize the way an agent i feels about a concept X. To formally represent the appreciation of an agent i about a concept X, we introduce the following formula:

$$Appreciate_i(X)$$

This formula can be read as "*the agent i appreciates the concept X*". Knowing this, we consider that i shares the same appreciations with j if i thinks that they both have the same appreciation about X. For instance, both i and j appreciate football, or both i and j dislike spinach.

$$Shared\_Appreciation_i(j, X) \equiv Appreciate_i(X) \wedge B_i(Appreciate_j(X)) \tag{Def. Shared}$$

On the opposite, agents i and j have contradictory appreciations if i thinks that i's and j's appreciations about X are different. For instance, i likes vegetables but thinks that j does not like it.

$$Contra\_Appreciation_i(j, X) \equiv Appreciate_i(X) \wedge B_i(Appreciate_j(\neg X)) \tag{Def. Contra}$$

Liking value To quantify the liking relation, we introduce the importance accorded by an agent i to a particular concept X, formally represented as follows:

$$importance_i(X) \in [0; 1]$$

The higher the value is, the more the concept X is important to the agent i. We suppose that this value is predefined for the agent. To compute the liking value between an agent i and an agent j, we represent the sum of the importance accorded to shared or contradictory concepts. The mean shared value refers to the mean value of the importance accorded to concepts for which i and j share the same appreciation. The mean shared value is formalized as follows:

$$mean\_shared\_value_i(j) = \frac{sum\{importance_i(X)|Shared\_Appreciation_i(j,X)\}}{card\{X|Shared\_Appreciation_i(j,X)\}}$$
$$with\ mean\_shared\_value_i(j) \in [0;1] \qquad \text{(F.Shared)}$$

The sub formula $card\{X|Shared\_Appreciation_i(j,X)\}$ represents the number of concepts X for which i and j share the same appreciations. For instance, if i and j like football and philosophy, the mean share value of i toward j is the mean between the importance i accords to these two concepts. The mean contradictory value refers to the mean value of the importance of these concepts for which i and j have contradictory appreciations. The mean contradictory value is formalized as follows:

$$mean\_contra\_value_i(j) = \frac{sum\{importance_i(X)|Contra\_Appreciation_i(j,X)\}}{card\{X|Contra\_Appreciation_i(j,X)\}}$$
$$with\ mean\_contra\_value_i(j) \in [0;1] \qquad \text{(F.Contra)}$$

The sub formula $card\{X|Contra\_Appreciation_i(j,X)\}$ represents the number of concepts X for which i and j have contradictory appreciations. For instance if i likes football and philosophy, but thinks that j doesn't like it, the mean contradictory value of i toward j is the mean between the importance accorded by i to these concepts. The final liking value of i toward j is the difference between mean shared value and mean contradictory value, represented as follows:

$$lik\_value_i(j) = mean\_shared\_value_i(j) - mean\_contra\_value_i(j)$$
$$with\ lik\_value_i(j) \in [-1;1] \qquad \text{(F.Like)}$$

A negative value means that i dislikes j while a positive value means that i likes j. The higher the value is, the more i likes j. For instance, if the concepts for which i share the same feelings with j are more important that those for which i and j have contradictory opinions, i's like value toward j will be positive. Formally, an agent i **likes** an agent j if and only if the like value of i toward j is positive.

$$Like_i(j)\ iff\ lik\_value_i(j) \in]0;1] \qquad \text{(Def.Like)}$$

Since the liking value is based on beliefs of the agent (the agent i thinks that j does not like the concept X), an agent i might like an agent j, but j might dislike i. This is consistent with the work presented in sect. 3.1 claiming that liking is not necessary mutual [24]. For instance, i can think that j does not like vegetables but this belief can be wrong (j actually does like vegetables). Then, if j believes that i likes vegetables, i dislikes j whereas j likes i.

Strategies Considering Heider's Balance Theory [32], an agent i has two different strategies to alter the liking relation between i and j:

1. The agent i can change its own feeling about a particular concept for which the agent j has an opposite appreciation. By this way, i and j will share preferences about the same concept (Def. Shared). This will increase the sharing value (F.Shared) and thus the like value of i toward j (F.Like).
2. The agent i can try to change j's feeling about a particular concept for which the agent i has an opposite appreciation. By this way, i and j will share preferences about the same concept. This will increase the sharing value and thus the like value of i toward j.

3.5 Familiarity

The theoretical background for the formalization of the familiarity is based on the Social Penetration Theory [34]. In this theory, the familiarity between a person A and a person B can be measured in terms of *breadth* and *depth*. The *breadth* represents the number of beliefs of A about B's feelings. For instance, A can know B's feelings about family, work, politics, etc. The *depth* represents the intimacy value accorded to a particular topic. For instance, A may have superficial beliefs about B's family while knowing very personal information about B's political orientation. Fig. 2 shows an example of familiarity relation, with different levels of breadth and depth.

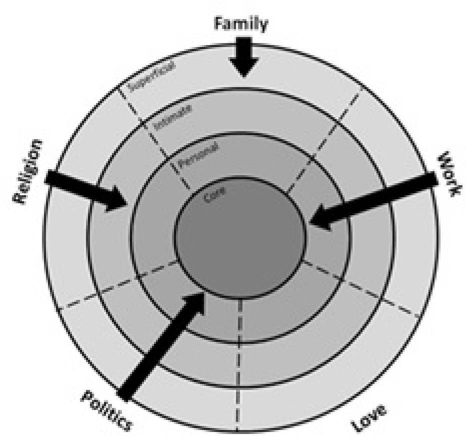

Fig. 2. An example of breadth and depth in familiarity relation

The familiarity value The first step to compute the degree of familiarity between an agent i and an agent j is to represent the ratio between the number of i's beliefs about js feelings and the number of possible conversational topics (breadth). That is, the number of possible conversational topics represents the number of concepts X for which i can have feelings. Given N the number of possible conversational topics, the breadth value is represented by the following formula:

$$\begin{aligned} & breadth\_value_i(j) = \frac{card\{X|B_i(Appreciate_j(X))\}}{N} \\ & with\ breadth\_value_i(j) \in [0;1] \end{aligned} \tag{F.Breadth}$$

The sub formula $card\{X|B_i(Appreciate_j(X))\}$ represents the number of i's beliefs about j's feelings. A high value means that i has a broad relation with j. For instance, an agent i can possibly discuss with j about five different topics (family, work, religion, politics and love) but only have beliefs about four of them. To compute the depth value, we first represent the intimacy accorded by an agent i to a concept X as follows:

$$Intimacy_i(X) \in [0;1]$$

A low value means that the concept is considered as superficial by the agent i, whereas a high value means that i considers the concept as intimate. We suppose that this value is predefined for the agent. Knowing this, the depth value is the mean of intimacy values and is represented by the following formula:

$$\begin{aligned} & depth\_value_i(j) = \frac{sum\{Intimacy_i(X)|B_i(Appreciate_j(X))\}}{card\{X|B_i(Appreciate_j(X))\}} \\ & with\ depth\_value_i(j) \in [0;1] \end{aligned} \tag{F.Depth}$$

$sum\{Intimacy_i(X)|B_i(Appreciate_j(X))\}$ represents the sum of intimacy values accorded to j's concepts for which i has beliefs. $card\{X|B_i(Appreciate_j(X))\}$ represents the number of j's concepts for which i has beliefs. The familiarity value of an agent i toward an agent j is formalized as the mean between breadth and depth.

$$\begin{aligned} & fam\_value_i(j) = \frac{breadth\_value_i(j) + depth\_value_i(j)}{2} \\ & with\ fam\_value_i(j) \in [0;1] \end{aligned} \tag{F.Fam}$$

A low value means that the agent i has a superficial relation with j, whereas a high value means that the relation is quite intimate. For instance, an agent i having numerous beliefs about j's feeling toward intimate concepts is very familiar with j. Using this formalization, a particular link can be drawn between breadth and liking. Indeed, new beliefs about j's feelings increase i's breadth (F.Breadth), but these beliefs will also change i's liking value (Def. Shared and Def. Contra). The more an agent i knows about another agent j, the more accurate i's liking relation will be.

Strategies In [35], the authors define four types of familiarity relations: (1) narrow breadth and superficial depth, where people have very limited contacts; (2) broad and shallow relationships, where people share superficial opinions about a large variety of topics; (3) narrow and deep relationship, involving personal disclosure about some specific topics and (4) a large and deep relationship, where people share intimate and personal information for a large variety of topics. Based on this formalization of familiarity, several strategies may be defined for an agent i to try to change his familiarity degree toward another agent j:

1. The agent i can ask j to give him superficial information about topics for which i does not have any beliefs. In this case, the number of beliefs of i about j's feelings will increase (F.Breadth), and the familiarity value will increase as well (F.Fam).
2. The agent i can ask j to give him more intimate information about particular concepts for which i already has superficial beliefs. In that case, the intimacy value will increase (F.Depth), and i will be more familiar with j (F.Fam).

According to the self-disclosure theory explained in [34], the disclosure process is reciprocal: giving intimate information to another person will lead him to disclose intimate information as well. We can thus define another strategy.

3. The agent i can share his feeling about a particular concept for which he wants to know j's feeling. Doing so, the agent j might share his feeling in return, increasing breadth or depth.

4 Conclusion and Future Work

In this work, we introduced a logical framework for artificial companions. The social relation between two agents is represented by three different dimensions: dominance, liking and familiarity. Dominance can be measured by the degree of influence on other's intentions. It also depends on the degree of dependence upon this same agent. The value of liking indicates the degree of like-mindedness between the two agents; an agent will like another agent if he believes that they share the same feelings (positive or negative) about particular concepts. The familiarity between two agents is represented by two values: the breadth is the number of beliefs that an agent has about another agent's feelings. The depth is the intimacy value accorded to these concepts. We also introduced dynamics of these dimensions. Agents may use different strategies to change their own mental state or try to influence other agent's relations. In future work, we will conceptualize the mechanisms used to form strategies and to reason about these three dimensions. Thus our next step is to implement a formal model of social decision making, leading a companion to reason about its own social relations, but also trying to influence other agents (virtual or human) considering their own beliefs and intentions. To simulate our model of social attitude, we will apply this model to a scenario where the agent, playing the role of a companion, could adopt different believable behaviors according to its social relation toward the user. This social relation might dynamically evolve depending on the interactions.

References

1. Lim, M.: Memory models for intelligent social companions. Human-Computer Interaction: The Agency Perspective Studies in Computational Intelligence **396** (2012) 241–262
2. Maslow, A.: A theory of human motivation. Psychological Review **50**(4) (1943) 370–396
3. Berscheid, E., Peplau, L.: The Emerging Science of Relationships. In: Close Relationships. W.H. Freeman, New York (1983) 1–19
4. Kriglstein, S., Wallner, G.: Homie: an artificial companion for elderly people. In: Conference on Human Factors in Computing Systems: CHI'05 extended abstracts on Human factors in computing systems. Volume 2. (2005) 2094–2098
5. Bickmore, T., Cassell, J.: Social Dialogue with Embodied Conversational Agents. In: Advances in Natural Multimodal Dialogue Systems. Springer Netherlands (2005) 23–54
6. Turkle, S.: In Good Company: On the treshold of robotic companions. In: Close Engagements with Artificial Companions: Key social, psychological, ethical and design issues. John Benjamins (2010) 1–10
7. Clodic, A., Alami, R., Montreuil, V., L., S., Wrede, B., Swadzba, A.: A study of interaction between dialog and decision for human-robot collaborative task achievement. (2007)
8. Cavazza, M., de la Cámara, R.S., Turunen, M., Gil, J.R., Hakulinen, J., Crook, N., Field, D.: How was your day? : an affective companion eca prototype. In: Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue, Association for Computational Linguistics (2010) 277–280
9. Gomes, P., Paiva, A., Martinho, C.: Episodic memory retrieval through re-appraisal of past events, Philadelphia (2010)
10. Campbell, R., Grimshaw, M., Green, G.: Relational agents: A critical review. The Open Virtual Reality Journal **1** (2009) 1–7
11. Bickmore, T., Picard, R.: Establishing and maintaining long-term human-computer relationships. ACM Transactions on Computer-Human Interaction (TOCHI) **12**(2) (2005) 293–327
12. Rist, T., Schmitt, M., Pelachaud, C., Bilvi, M.: Towards a simulation of conversation with expresive embodied speakers and listeners. (2003)
13. Lim, M.Y., Dias, J., Aylett, R., Paiva, A.: Creating adaptive affective autonomous npcs. Autonomous Agents and Multi-Agent Systems **24**(2) (2012) 287–311
14. Ballin, D., Gillies, M., Crabtree, I.: A framework for interpersonal attitude and non-verbal communication in improvisational visual media production, London, UK (2004) 203–210
15. Prendinger, H., Descamps, S., Ishizuka, M.: Scripting affective communication with life-like characters in web-based interaction systems. Applied Artificial Intelligence: An International Journal **16**(7-8) (2002) 519–553
16. Marsella, S.C., Pynadath, D.V., Read, S.J.: Psychsim: Agent-based modeling of social interactions and influence. In: Proceedings of the international conference on cognitive modeling, Citeseer (2004) 243–248
17. Tomlinson, B., Blumberg, B.: Social synthetic agents. Computer Graphics **26**(2) (2002)
18. Prada, R., Paiva, A.: Teaming up humans with autonomous synthetic characters. Artificial Intelligence **173**(1) (2009) 80–103

19. Pereira, G., Prada, R., Santos, P.: Conceptualizing Social Power for Agents. In: Intelligent Virtual Agents (IVA'13), Edinburgh (2013)
20. Raven, B.: The bases of power and the power/interaction model of interpersonal influence. Analyses of Social Issues and Public Policy **8**(1) (2008) 1–22
21. Rao, A., Georgeff, M.: Modeling Rational Agents within a BDI-Architecture. In: Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning. Morgan Kaufmann, San Francisco (1991) 473–484
22. Dastani, M., Lorini, E.: A logic of emotions: from appraisal to coping. In: Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'12), Valencia, Spain (2012)
23. Moshkina, L., Arkin, R.: On tameing robots. IEEE International Conference on Systems, Man and Cybernetics **4** (2003) 3949–3959
24. Prada, R., Paiva, A.: Social Intelligence in Virtual Groups. In: New Advances in Virtual Humans. Volume 140. Springer (2008) 113–132
25. Isbister, K.: Better Game Characters by Design, A Psychological Approach. Morgan Kaufmann (2006)
26. Svennevig, J.: Getting Acquainted in Conversation. John Benjamins, Philadelphia (1999)
27. Lewicki, R., McAllister, D., Bies, R.: Trust and distrust: New relationships and realities. Academy of management Review **23**(3) (1998) 438–458
28. Sadek, D.: A study in logic of intention. Proceedings of the 3rd international conference on Principles of knowledge representation and reasoning (KR'92) (1992)
29. Kripke, S.: Semantical considerations on modal logic. Acta Philosophica Fennica **16**
30. Emerson, R.: Power-dependence relations. American Sociological Review **27**(1) (1962) 31–41
31. Ochs, M., Sadek, D., Pelachaud, C.: A formal model of emotions for an empathic rational dialog agent. Autonomous Agent and Multi-Agent Systems **24**(1) (2012) 410–440
32. Heider, F.: The Psychology of Interpersonal Relations. Lawrence Erlbaum Associates Inc (1958)
33. Zajonc, R.B.: The concepts of balance, congruity, and dissonance. Public Opinion Quarterly **24**(2) (1960) 280–296
34. Altman, I., Taylor, D.: Social Penetration: The Development of Interpersonal Relationships, New York (1973)
35. Altman, I., Vinsel, A., Brown, B.B.: Dialectic conceptions in social psychology: An application to social penetration and privacy regulation. Advances in experimental social psychology **14** (1981) 107–160