From Individuals to Social and *Vice-versa*

André Campos [1,2], Frank Dignum [2], Virginia Dignum [2]

[1] Computer Science and Applied Mathematics Department
UFRN Lagoa Nova, 59072-970 Natal-RN, Brazil
andre@dimap.ufrn.br

[2] Information and Computing Sciences Department
Padualaan 14 3584CH Utrecht, The Netherlands
{dignum,virginia}@cs.uu.nl

Abstract. The concept of cyclical influence between individuals and society is widely accepted, but hard to understand in all details. This paper proposes the use of three processes of social influence as a way to study the link between social and individual levels of abstraction. These processes are used to design an agent architecture which tries to provide explicit links to its social context. In order to detail the impact of the social influence, the architecture also includes personality and emotional aspects.

Keywords: Social influence, BDI agents, personality, emotional agents.

1 Introduction

The study of social dynamics based on individual behaviors is not recent ([1] *apud* [2]). However, in the last years, this approach has been deeply boosted by the capacity of current models to represent in a more precise way the structures, norms, culture, and resources available in a society, as well as by the capacity to model in more detail the complexity of the human behavior. This includes the results from the interactions between emotions, personality traits, cognitive reasoning and, certainly, how the individuals relate to their social context.

These two levels of complexity (society and human behavior) are extremely interdependent to each other. Several works have already being conducted to study their relationships. However, those works normally address one of the two research directions, either: 1) by studying the emergence of collective patterns and structures from the individuals, or 2) by studying the influence of social elements (norms, policies etc.) over the individuals. Some researchers have also conducted analyses to evaluate the two pathways, composing the cycle macro-micro-macro. Nevertheless, for the sake of simplicity, these studies normally address only specific relations (e.g. norms and emotions [3]) and do not involves many aspects of the human being.

In order to facilitate the model translation between micro and macro levels of abstraction some authors (e.g. Dignum et. al [4]) propose the use of a mediation layer, named meso-layer. The latter connects the two other levels of abstraction by gathering information from the macro level elements (like global behavioral patterns)

and including new ones (like norms and organizational structures) that “influence” individual behavior. The meso-layer presented in Dignum et. al [4] aims to study the influence of policies over the individuals. These policies are defined by policy makers at the macro and meso levels in an agent-based simulation, and the resulting behavior is analyzed. Briefly, it is used in a top-down approach, which means that there is no change in the meso level arising from the bottom-up (the agents).

Although this is enough for studying policy making, it is not adequate to study the emergence of elements present in the meso level. An example is the dynamics of existing groups or the creation of new ones. Groups, as elementary social structures, are part of the meso layer. However, they should not necessarily be formed or imposed as part of a social or organizational policy. Individuals can spontaneously leave or take part of a group according to their interests, needs, preferences and objectives, following a bottom-up approach.

In order to make changes in the meso level that reflects human societies, it is necessary to use coherent theories linking psychological foundations with social behavior. This paper proposes an agent architecture addressing this need, based on the model of social influence proposed by Kelman [5]. The architecture also incorporates other theories in order to detail several human characteristics, more precisely: emotions, personality, personal and social values.

The following two sections introduce the elements of social influence proposed by Kelman and how they related to different levels of abstraction, respectively. The Section 4 presents other elements of human behavior that we consider important to configure the link between micro and meso, as well as the reasons for choosing specific models. After, the components and the cognitive process of the architecture of the social agent are described. Then, some final remarks are made, pointing out potential applications and the next steps of the work.

2 Social Influences

The behavior of a society can be seen as a consequence of the choices and actions performed by their members, but their choices are also influenced by the whole system. This recurrence generates a highly dynamic behavior cycle that may explain the resistance and/or desire for changes within a society.

In order to examine this cycle, it is necessary to study the mechanism that influences individual behavior from a social point of view. Kelman investigation on social influence, proposes a linking model between the individual and social systems. Although his studies were initially used as a mean to understand the mechanisms which allow a person to influence a target audience [6], the original model proved to be also useful in different contexts: from group psychotherapy to large social systems, involving organizations and very large social contexts, involving nations and its links to individual values (e.g. national identity) in the context of peacemaking (these works are summarized in [5]).

For Kelman, individual beliefs are not necessarily (fully) integrated into the person’s own value system. The latter is highly dependent on external influence. His investigation on this dependency led him to distinguish three processes of social influence: compliance, identification and internalization [7].

These three processes address the issue of "when" an individual accepts the influence of other person. The first one, *Compliance*, occurs when the individual wants to attain a favorable reaction from the other (like a child who adopts a behavior to be rewarded or to not be punished). The second one, *Identification*, occurs when the individual wants to establish or maintain a satisfying relationship to the other (like a husband who changes his attitudes to satisfy his wife expectations). The last one, *Internalization*, occurs in order to maintain the equivalent correspondence of actions and beliefs with his or her own value system (like a teenager who imitates other's attitudes for maximizing or confirming his or her own values).

Along with these processes, there is also a dimension orthogonal to the aspects of social influence, which is the distinction between two types of personal concerns. These concerns drive how an individual will react to social influence: either by instrumental concerns (e.g. assuring rewards or avoiding punishments) or by self-maintenance concerns (e.g. managing one's public image).

In large social systems, these elements of social influence are related to three distinct concepts: interests, relationships and identities [5]. These concepts are captured from tasks that all social elements (individuals, groups, organizations, societies) must perform as they negotiate their social environment:

- Protecting and promoting their interests: This task is related to the compliance process, where individuals and groups may influence each other to attain their own interests (or goals);
- Establishing and maintaining their relationships: This tasks is related to the identification process, where individuals and groups may establish the set of roles for their expectations;
- Affirming and expressing their identities: This tasks is related to the internalization, where individuals and groups share and exchange their values (or identities).

According to Kelman, "*In managing their interests, relationships and identities, individuals and groups must attend to the requirements of both social order and self-maintenance, and of ensuring the proper balance between them*" [5]. In the following section, we present how this is applied to the meso and individual layers by introducing the appropriate links between these two levels of abstraction.

3 Elements of the Meso Layer in the Social Influence

The meso level of abstraction, as described in Dignum et. al [4], refers to a intermediate level connecting and translating the elements found on the macro level to the micro level. It is composed of three types of components. The first one comes from the descriptive elements that was empirically validated in the domain, but are not in the focus of the simulation. They are regarded in the meso level as "law of nature", which the agents abide. The second component comes also from the macro level, but is in the focus of the simulation. This component is treated as a benchmark to which the agent behavior is compared. The third type of component tries to influence individual behaviors, through mechanisms that regulate their joint activities [8]. They include elements like norms, organizational structures and cultural backgrounds.

The next paragraphs present a correspondence between these elements (norms, structures, and cultural backgrounds) and the elements discussed in the previous section (interests, relationships, and identities). It establishes a means to express the social influence in a bi-directional way through a link between the micro and meso levels of abstraction.

3.1 From Interests, Relationships, and Identities to Rules, Roles, and Values

It is easy to identify that interests, relationships, and identities are inherent properties of both social and individual entities, as long as they refer to the link between them. As argued by Kelman, "*Individuals have interests, relationship and identities, which they pursue and express through the various groups and organizations with which they are affiliated. The groups and organizations – formed, essentially, to serve their members – in turn develop their own interests, relationships, and identities, which become personally important to the members and which the members are expected to support*" [9].

Interests reflect the goals that both individuals and groups have. In this aspect, groups establish a set of rules necessary to assure their member to attain the group's goals. Members are then influenced by these norms and rules through a mechanism of rewarding or punishment. For instance, the existence of a rule like "One should not drive faster than 100 km/h" reflects a specific goal of a social system, which tries to assure the behavior of their members in order to achieve it. On the other hand, individuals, when they cannot (or do not want to) leave a social system, may also influence it through their own internal rules and their respective rewarding or punishment strategies. Strikes are good examples of such strategies.

Rules, also described as norms in the meso level, are then a key concept which is present in both levels of abstraction. The social influence through norms is related to a compliance process, which represents adherence to them. In accepting influence via this process, members (and sometimes groups) assure themselves to have continuously rewards and approval (or also to avoid punishments).

Relationships reflect the roles assigned to members of a group and their respective expectations from them, as well as the role of the group itself to their members. Members are then influenced by the expectation from the others according to the role they are playing in the group. A teacher would, for instance, try to behave according to the expectation of the students (e.g. to give a good lecture). Groups have also their roles for their members. This means that individuals belong to a group only when there is a gain for them, i.e. when the group is playing its role as expected. Following the same example, a student may leave a college if the lectures are not given as expected by him or her.

Roles are then another key concept from the meso level that it is present at the individual level, and reflect the structure of the group, organization or society. The social influence through roles is related to an identification process. In accepting the influence via this process members and groups are meeting the expectations of their roles, thus maintaining their desired relationship to the group or to the members, as well as their wish to fully accomplish their roles. It is important to notice that the

influence here is not about changing the role of a member or a group, but to change their settings in order to better accomplish that role.

Finally, identities reflects the system of values that everyone possesses, from their past experiences or cultural backgrounds. Members are then influenced by values shared in their social contexts as a way to maximize or maintain their own value system. A foreigner may, for instance, incorporate a behavior present in a host country according to what he or she considers important (its own values), and the feeling of belonging to such a society would be manifested by the set of values that s/he share with the society. Group identity is also a product from the values shared by their members, even if it can be considered as being almost independent from the individuals. This notion of collective identity is intrinsically related to the culture of the social system.

Value system is another key concept present in both meso and individual levels. The social influence through value system is related to an internalization process. In accepting the influence via this process members assure the maintenance or maximization of the equivalence of actions and beliefs with his or her own value system.

3.2 Linking the meso and micro layers

The three processes of social influence proposed by Kelman suggest three different ways in which individuals and social systems are integrated: by adherence to its rules (or norms), by the involvement in its roles, and by the sharing of its values. The way in which rules and roles are exchanged between groups and individuals are, however, different. But in both cases, values serve as guides for the other elements.

For the groups, the rules for the individuals are a way to constraint individual behavior, while for the individuals, their rules are a way to assure that their roles (established by the group) does not deviate them from their values. For instance, let us consider a working environment where a rule forbidding couples to work together exists, and that there is couple whose group roles imply in working together. What should the couple do? Their personal values (to lie or to not lie, how much they are attached to each other, how much they are committed to the work etc.) will drive their standard action patterns (i.e. personal rules) toward the group, saying if they obey the rules (breaking the relationship), do not follow the rules (lying about the relationship), leave the group, or decide to try to change the rules of the group. Indeed, one can see the individual rules as his/her principles, ethic code or moral, which set what one should or should not do. They are based on the personal values.

For groups, the roles of the individuals are a way to structure and coordinate actions of their members, while for the individuals, the role of a group is to help them to maximize their personal values. Thus, the same kind of decision may happen when a role is assigned to an individual who does not see the role related to anything s/he considers important. Even worst, the role may be against some personal values. Let us consider, for instance, the obligation of an individual against a specific war to serve in the Army (it is an obligatory role in some societies). Although the process of Identification in a higher level of abstraction (National identity), if the role in such a group (Army) does not contribute to the personal values of the individual, why s/he

would want to belong to the group (Army)? Some enforcement rules (punishments) may try to change this picture, but it is the values which guide the link between group and individual.

The relationship between the goal of rules, roles, and values for groups or individuals is them enriched through the existence of different types of social attitudes. For instance, Dastani et. al [10] distinguish three basic types of role enactment: selfish enactment (the individual gives priority to its own goals), social enactment (the individual gives priority to its role's goals), and maximally social enactment (the individual ignores its own goals for the duration of the role enactment). Following this line, V. Dignum introduces another type: maximally selfish enactment (the individual ignores the role's goals) [11].

Figure 1 illustrates the links previously described showing an individual belonging to two different groups and the three processes of social influence between them. Rules, roles, and values connect the individual to the groups s/he belongs to.

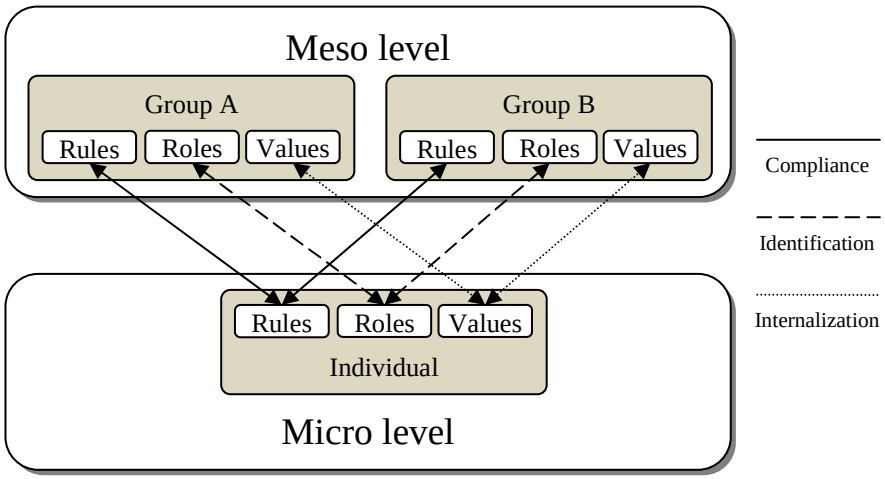

Fig. 1. Rules, roles, and values are the elements which interface the meso and micro levels of abstraction. Rules enforce the achievements of goals, roles define expected behaviors, and values constitute their identities. They reflect the three processes of social influence: compliance, identification and internalization.

Figure 2 illustrates (next page) how such a dynamics occurs explaining the vision of rules, roles and values from the perspectives of individuals and groups. This dynamics points out the social influence from the meso to the micro level, and *vice-versa*.

It is easy to observe that, although presented as three different concepts, rules, roles, and values are highly interdependent (as expressed in the legend of Figure 2). In order to structure how those dependencies should be treated internally by the individuals it is necessary to go deep on other aspects of human behavior. The following section explores some concepts of the individual behavior that we consider essential for configuring the whole picture of social influence: rationality, emotions, and personality.

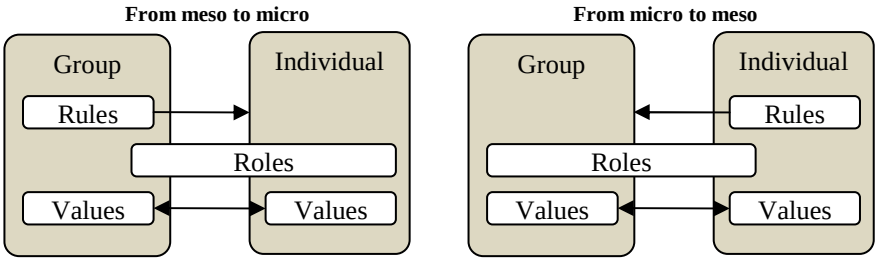

Fig. 2. The direction of social influence according to the elements involved. One can read the left-side block of the diagram as the group saying "*These are my rules that you must follow when playing your roles to me in order to bring or maximize our shared values*". The right-side block can be read as the individual saying "*These are my rules that I will follow as a member of group playing these roles to me in order to bring or maximize some shared values*".

4 The Behavior of Individuals

In the last years, several studies have been conducted in order to model human behavior..This includes physiological aspects (stress, fatigue,…), emotions (happiness, disappointment, …), behavioral preferences (personality) and socials (reputation, trust,…). Each of these domains regards human being from a particular perspective. Taken independently, they help to understand some specific aspects of human behavior. Taken in an integrative way, the aforementioned studies help to construct the whole panorama of the human being.

However, to integrate many different theories and/or models into a coherent architecture is a hard task with several potential validation issues (that is, to join the pieces of a puzzle does not mean to solve it). As any complex system, the resulting behavior from the interaction of its elements is very sensible to the way in which each element and the respective interactions are modeled. The complexity, already present in the individuals, still increases when they interact together. Indeed, the increasing number of agents in a simulation makes difficult to validate the model, assuring the expected macro behavior from an adequate internal representation.

Although such a difficulty as well as the practical issue of integrating different theories, we understand that to simulate real human social behavior it is necessary to incorporate different aspects of the human being, i.e. it is important to integrate the concepts from different areas and also to study the relationship between them. Following this line, prior to addressing the potential validation issues outlined here, it is firstly necessary to address how to integrate the several facets of human behavior, as well as to conceive the required links between them and the social framework presented in the previous sections. The elements that can be clearly identified as playing part in the social dimension are: the decision making process, the emotional component and the individual personality. The following paragraphs present rationales behind the choices made for modeling each of those dimensions.

4.1 Decision Making

Agents representing individuals within a social simulation should be able to reason about their surrounding environment, which encloses the social context where they are embedded. Several cognition architectures have already been proposed for such a purpose. Each one has its own strengths and weaknesses, according to where they are applied.

Cognitive models, such as ACT-R [12] and SOAR [13] aim at understanding how people organize knowledge and produce intelligent behavior based on numerous facts derived from psychology experiments, employing quantitative measures. However, these models lack realism since they do not incorporate demographics, personality differences, cognitive style, situational and emotive variables, group dynamics and culture. On the other hand, neurological oriented models that mimic the brain, such as neural networks, lack transparency to link observed behavior to the implementation. Realistic agent models should combine the characteristics of the different types.

The model of the human mind CLARION [14] aims to explore the interaction of implicit and explicit cognition, emphasizing bottom-up learning (i.e., learning that involves acquiring first implicit knowledge and then acquiring explicit knowledge on its basis). CLARION's goal is to form a (generic) cognitive architecture that captures a variety of cognitive processes in a unified way and thus to provide unified explanations of a wide range of data. The BDI model is also a generic cognitive architecture [15]. It has formal logic-grounded semantics and introduces well established concepts from the theory of intentionality [16].

The CLARION model and the BDI model are both excellent candidates for the extension as aimed for in this paper. However, although the fact that BDI requires extensive computational resources, it provides clear clues where the concepts presented in the previous sections can be applied. The intentional paradigm behind the BDI model fits well for modeling the agent interests, represented by its desires, and their matching with the group interests, represented as a commitment to the goals established by the agent's roles.

4.2 Emotions

Emotions can indeed deeply influence the interactions among the members of a group, by stimulating or inhibiting behavior and, as consequence, influencing in the behavior of the whole system. The inclusion of the emotional component in the agent architecture helps to set up the internal consequences of individual choices regarding the rules, roles and values. For instance, the emotional reaction when people find themselves deviating from standards in the domain of responsibility may be: 1) social fear, when they deviate from external social rules or norms; 2) guilt, when they deviate from role expectations; or 3) regret, when they deviated from social values [7].

Emotion is, however, a subjective concept, which means that several authors have already presented their own vision. It is not our purpose to explore all existing work in the area, but to focus on those approaches that integrate emotions into a cognitive

cycle of perception – action – perception. The model of emotions proposed in [17], also known as OCC model, addresses such a need. They support the idea that emotions are a product resulting exclusively from our cognition, generated from our perception and our expectations. As a cognitive outcome, emotions have their origins in our perception and involve some sort of positive or negative reaction to what was perceived. Following this line, they consider emotions as valenced reactions to events (e.g. pleased, displeased,…), agents (e.g. approving, disapproving,…), and objects (e.g. liking, disliking,…).

Some authors criticize such a theory for not including physiological aspects of emotions, like other theories (e.g [18]). However, the exclusively-cognitive approach proposed in the OCC model is enough for our purposes. Although we recognize the influence of physiological aspects to the emotions, the level of granularity for the model we want to address in this paper does not deal with issues ranging from social system to hormonal specificities at the individual level.

4.3 Personality

The use of personality models in agents in order to study the micro-meso relationship becomes essential as it facilitates the creation of realistic complex scenarios. It improves and reproduces in a more realistic way the autonomy of the agents. Indeed, autonomy is related to how the individuals behave and what make them to behave differently from each other, even when they face the same situation. Psychologists have tackled this issue for several decades through what they named Personality [19].

It is personality which enables individuals with the same roles, following the same constraints (rules), sharing the same values, and having (virtually) the same beliefs, to behave differently. It provides a clear mechanism of preferences for specific behavioral patterns which are independent from any other aspect (beliefs, values, etc.). More precisely, personality represents the structured and dynamic set of characteristics of an individual, normally acquired from the environment and personal experiences [20]. These theories can be divided into two categories: personality types and personality traits [21]. Personality traits theories express the human characteristics through quantification values (for instance, 0.9 represents a strong characteristic). Examples of this category are the Big Five model (also known as OCEAN or Five Factor model) [22] and the model proposed by Theodore Millon [23]. Personality type theories do not express characteristics through values, but rather through a set of category (or types) in which the individual belongs. An example is the MBTI model [24], proposed as an extension of the theory of personality preferences developed by Jung.

For our concerns, we are interested in representing personality in social simulations independently from the context where the simulation is running. Following this line, we need a model that can be more easily focused on the process rather than the contents. A personality type model fits well this requirement since it is not expressed as a continuous (like the personality traits). Thus, according to an individual's type, a different cognitive process, involving his/her emotions and decision making, can be performed. The MBTI model fits well the needs of our approach. Besides that, it has a

solid background through the several decades of use in organizational behavior studies [21].

5 The Architecture of the Social Agent

The approach used to present the agent architecture is to show the overall picture of the agent cognition taking into account the different aspects mentioned in the previous sections. The architecture is then presented in a high-level of abstraction rather than concentrating in a particular and focused problem.

Our approach to support the micro-meso interaction as a process of social influence uses: 1) a decision making process, based on the BDI model; 2) an emotional component, based on the OCC model; and 3) a personality-based mechanism, based on the MBTI model. These elements must connect to a social component which set up the individual standards (rules), the roles that the agent plays (as well as the groups where it plays), and its personal values.

The Figure 3 provides a general overview about the different components in the proposed architecture. The personality component is introduced over the others, since it does not save or process information, but rather it establishes the way in which the other components do that. The following subsections explain how the dynamics of the architecture.

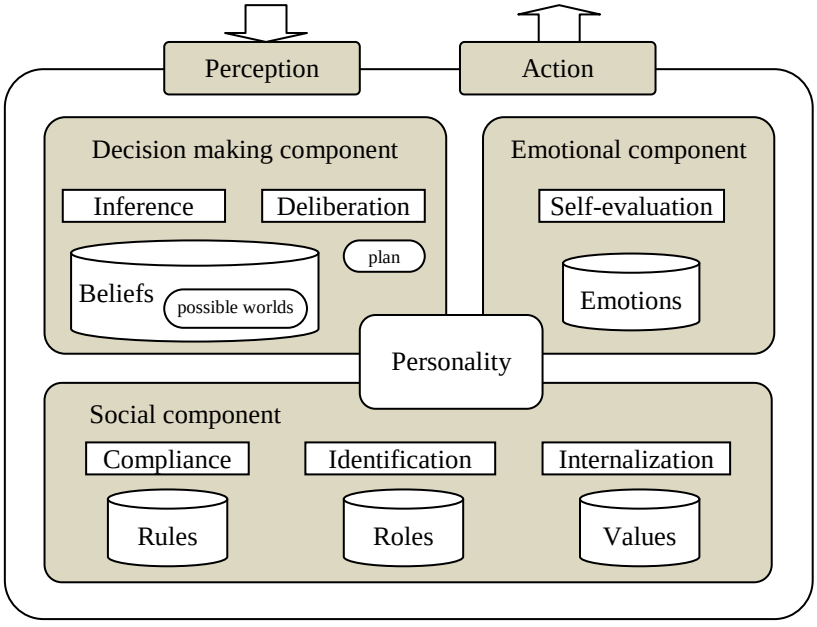

Fig. 3. The architecture of the proposed social agent. It incorporates four main components: The Decision Making, the Emotional, the Social and the Personality.

5.1 Components of the Architecture

The decision making component is responsible for defining the social action plan which the agent will be committed to, as well as for deciding when the agent should reevaluate its goals and, consequently, its plan. The desires (BDI element) in this model come from different components in the architecture: from its rules (not the group rules, since they represent constraints and not desires), from the objectives of the roles the agent in playing, from its personal values, from the plan it is committed to, and from the emotions. Those set of desires can be inconsistent and/or unachievable. Following a BDI mode of reasoning, the Deliberation element is responsible for filtering them into consistent and achievable goals, according to the agent beliefs. It will then choose the goal (or goals) to commit, defining the agent intention, and a plan for it is setup.

An approach used in this model that differs from existing BDI architectures is the separation of the inference mechanism from the deliberation process. In other words, in the Deliberation element, there is no reasoning about the possible worlds. No conclusion about which world state can be achieved is drawn. The inference about possible worlds are made in advance by the Inference element and put available into the agent beliefs. The reason for such a separation is to make the possible worlds structure not only dependent on the rational deliberation (goals), but also dependent and based on the agent emotional focus, the processes of social influence and its personality type (they are explained further on). The structure representing the possible worlds (for instance, a tree) is then not based only on rational choices. Instead, the rational choice will search for the best option according to only what it possible to foresee (what is present in the tree). An agent representing an individual in a panic state (emotional state), for instance, will not develop further its possible worlds' analysis. Its behavior would be almost reactive. This approach fits several studies involving cognition and emotions pointing up that emotions set boundaries for our decision-making process (otherwise, we could think about the best option of an issue forever). It also fits to the Damasio concept of "somatic markers" [25].

The Figure 4 (next page) illustrates the inference process taking into account not only a rational choice but the emotions which can be triggered by possible world states. The figure shows the current world state (as the agent believes) as W_0 and a set of possible worlds deriving from it. In the example, the agent believes that a desired state can be achieved ($W_{1.2}$), but an undesired state can also occurs (W_2). It is worth to explain here that desired and undesired states are derived from the agent values and its social influence (rules and roles). As an undesired state, W_2 can produce emotions (explained in the next paragraph) and define the agent emotional focus. The emotional focus is then explored in a next iteration in order to develop new possible worlds deriving from it. In other words, the agent will not infer in any possible direction, but in virtue of its goals and affective states.

The emotional component is responsible for managing how the represented emotions in the agent affect the other components. As mentioned in the previous paragraph, emotions can influence the decision making by restraining or expanding the vision of possible worlds. The latter can also influence the current emotional state of the agent. For instance, if a not desired world can be foresee in the possible worlds (W_2), the emotion "fear" is produced by the emotional component.

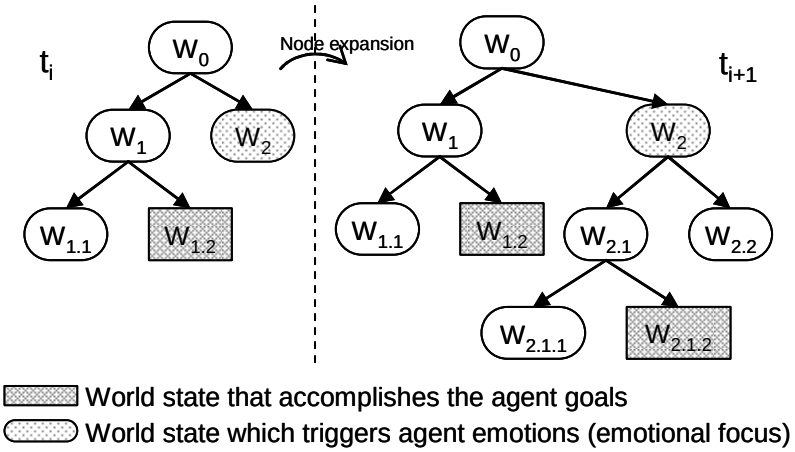

Fig. 4. An example of how the agent infers the possible worlds, taking into account the desired and undesired world states. States can trigger agent emotions which also drive the inference mechanism.

Following the OCC cognitive structure of emotions [17], agent emotions are represented as a tree of valences. In our model, each tree node has an intensity value, which represents how deeply the agent is affected by the perceived current world or by the possible future worlds. The emotion intensity of a world state is based on the elements from the social component, i.e. the agent values related to such a state, the roles the agent is playing, and its personal rules. The self-evaluation box presented in Figure 3 represents this process of updating the intensity values of the agent's emotions. The element which has produced the most intensive emotion in the agent becomes its emotional focus, which will in return influence the possible worlds' vision (as mentioned in the previous paragraph).

The social component is responsible for interacting with the outer social world through the three processes of social influence proposed by Kelman. Its interaction with the other components has already been briefly outlined. Although the three processes of social influence can be defined into the decision making component (they are part of a decision mechanism), we choose to define them as separated processes in order to explicitly identify them, being easier to adapt to different social contexts without having to redefine the deliberation mechanism. This separation also helps the management of the possible worlds' structure, since the compliance, the identification, and the internalization would alter it only when there are perceived changes in the rules, roles and values of the groups, respectively.

Finally, the role of the personality in this architecture is to define how some processes are performed. The configuration should reflect the four dichotomies presented in the MBTI model. For instance, according to the type of the agent in the Sensing-Intuition dimension, the way in which the possible world's structure is expanded should be different. A sensing agent would expand the tree in depth, trying to figure out the real and concrete consequences of its actions (they tend to be pragmatic), while an intuition agent would expand the tree in breadth, trying to figure out all the possible outcomes (they tend to want a whole overview of possibilities).

Another example is how the deliberation valuates a world state in the tree in order to choose the best option. Thinking agents would give more importance to world states that achieve the goals of its roles without deviating from its principles (rules), while Feeling agents would valorize the world states which satisfy its personal and its groups' values.

5.2 The cognitive process

This subsection provides an overview about how information is transformed from a perception stimulus to an action to be performed. Although there is no flux of information in Figure 3, there are several data dependencies between the components of the architecture. The choice of not explicitly represent them in the figure is to avoid surcharge of information, which could make it harder to understand. The following paragraphs present the process in a descriptive way, and after as an algorithm.

In short, the perception component receives stimuli from the outer world and updates the beliefs according to the agent personality. Indeed, the personality provides different ways to interpret a stimulus, which means that different agents facing the same situation may generate different beliefs. For instance, individuals with a *Thinking* personality type "*are direct to the extent of seeming insensitive to others*" (with *Feeling* type) [21].

After the beliefs have been update, the social component, through the processes of compliance, identification, and internalization, evaluates how this perceived world affects the agent's rules, roles, and values, respectively. This is, in fact, an evaluation about how the changes in the outer world impacts on the inner world, and it can be made by comparing the rules, roles, and values from both. Let us consider, for instance, the previously mentioned example of the individual against a war who has just invited to serve the Army. It is in this cognitive step that an agent representing this individual will contrast its values, roles, and rules to the society ones, and elaborate the impact of this new perceived outer world in its inner world. This impact can be considered as the individual level of deviation from social standards.

If the inner world is affected, for instance some personal values are infringed, emotions are raised from the current world state (W_0). The self-evaluation process updates then the emotions, following the OCC-based type of reactions (to events, agents, or objects), and an emotional focus is defined. In the previous example, *dissatisfaction* and *disappointment* emotions might arise from the impact caused by the invitation to serve the Army. If such emotions are those which contribute the most for the current emotional state of the agent (the sum of all its emotions), this event becomes then the agent's emotional focus.

The next step is performed by the decision making component, which expands or restraints the possible worlds' structure base on the plan the agent is currently committed (intention), the emotional focus, the level of deviation from social standards, as well as its personality type (as previously described) and its beliefs. Following the same example, the agent will expand the possible world's reasoning about the consequences and possibilities related to the Army invitation, which is its emotional focus. After a fixed number of expansions in the structure, the emotions are

reevaluated. Emotions like *fear* or *anxiety* may then appear as consequence of new possible worlds, and new inferences are made based on the new emotional focus.

The cycle may continue according to the personality type (*Sensing* persons will expand the possible worlds searching for concrete actions, while *Intuition* persons will expand to get an overview of all possibilities involved) and level of emotions raised (e.g. *panic*). After this, the deliberation process will filter the desires (spread in the other components) into consistent and achievable goals, reevaluate the current plan, and may trace a new plan to commit with. This plan can express a way of accepting the social influence (through compliance, identification, or internalization), as well as a way to not accept the social influence, trying so to change the social environment. In both cases, if there is a new plan, it is send to the action component, which will replace the last one. The following code specifies in a high-level of abstraction the description above.

Pseudo-code of the reasoning process of the social agent (in a high-level of abstraction), where: T is the personality type, Ru, Ro, and Va are the rules, roles, and values, P is the plan the agent is currently committed and A is the set of possible actions.

```
01: while alive
02:   S = get stimuli (outer world)
03:   B = update beliefs (S, T)
04:   Q = evaluate impact (B, Ru, Ro, Va)
05:   E = update emotions (B, Q, T)
06:   repeat
07:     W = review possible worlds (B, P, E, Q, T, A)
08:     E = update emotions (B, Q, T)
09:   until not(panic(E)) or has to react (B, T)
10:   if empty(P) or not(achievable(P,W)) or reconsider(P)
10:     D = gather desires (Ru, Ro, Va, E, P, T)
11:     I = define intention (B, D, I, T)
12:     P = generate plan (W, I, A, T)
13:   execute action (P, A)
```

6 Final remarks

In this paper we propose the use of three processes of social influence, namely *Compliance*, *Identification*, and *Internalization*, into an agent-based social simulation. Those processes are presented into an organizational approach (rules, roles, and values) where explicit links between individuals and the society are provided. In order to compose the general framework where rules, roles, and values influence and are influenced by individuals, other concepts related to human behavior were introduced and an agent architecture was conceived. We presented the main components of such an architecture and described its cognitive reasoning process.

The work described here is still in its early phase. The general concepts and abstractions were proposed and an architecture in a high-level of abstraction was designed. Connections from meso layer toward the micro were stressed and some hooks for the opposite direction were also stated. The latter is currently being developed based on the idea that, for instance, norms emerge through their

immergence in the agents' minds [26]. The present work includes norms in the agent cognition process, which can facilitate such a process.

Since the presented work employs an unusual approach (it embraces several aspects of social and individual at once), few works could be related in depth. Some agent architectures addressing the social component within the agent cognition do not address emotions and personality (e.g. B-DOING [27] and EMIL-A [26]). Others, which link social aspects of the emotions, do not integrate an organizational or normative approach (e.g. [28]). The PMFServ architecture [29] uses however the same approach, embracing a large number of concepts. PMFServ exploits however a quantitative way of modeling, where the agent decision-making is based on weighted sum of values and do not look further then the next world state. The architecture presented here is target to qualitative models and is the agent is able to reason about plans.

Although the lack of simulation results showing the impact of social influence over the individuals and *vice-versa*, the architecture was grounded in well established concepts. We foresee that the proposed model can be applied in several social contexts, ranging from the study of group formation to the study of emergence of insurgent movements, including its respective causes and consequences from and to the society.

The next steps are to provide formal specifications regarding the dependencies between meso and micro elements, as well as the dependencies inside the agent architecture, and to implement them in a simulation platform. We envisage implementing the agents in an extension of the 2APL language [30], since both are founded on the BDI model of reasoning.

Acknowledgments. This work is partially supported by the Brazilian National Research Council (CNPq).

References

1. Halpin, B.: Simulation in Sociology: A review of the literature. In: Proceedings of the Workshop on Potential of the computer simulation for the social sciences, University of Surrey (1998).
2. Conte, R., Gilbert, N., Sichman, J.S.: MAS and Social Simulation: A Suitable Commitment. In: Conte, R., Gilbert, N., Sichman, J.S. (eds.). MABS'98. LNCS, vol. 1534, pp. 1--9. Springer, Heidelberg (1998).
3. Staller, A., Petta, P.: Introducing Emotions into the Computational Study of Social Norms: A First Evaluation. Journal of Artificial Societies and Social Simulation. vol. 4, http://jasss.soc.surrey.ac.uk/4/1/2.html (2001).
4. Dignum, F., Dignum, V., Catholijn, J.: Towards Agents for Policy Making. In: Proceedings of the Multiagent-based Simulation, MABS'08, Estoril, Portugal (2008).
5. Kelman, H.C.: Interests, Relationships, Identities: Three Central Issues for Individuals and Groups in Negotiating Their Social Environment. Annual Review of Psychology. 57, 1--26 (2006).
6. Kelman, H.C.: Attitude change as a function of response restriction. Human Relations, 6, 185--214 (1953).

7. Kelman, H.C.: Social influence and linkages between the individual and the social system: Further thoughts on the processes of compliance, identification, and internalization. In: Tedeschi, J. (ed.) Perspectives on social power, pp. 125--171 (1974).
8. Feltovich, P., Bradshaw, J., Clancey, W., Johnson, M.: Toward an Ontology of Regulation: Socially-Based Support for Coordination in Human and Machine Joint Activity In O'Hare, G., O'Grady, M., Ricci, A., Dikenelli (eds.). ESAW'06. LNCS. Springer-Verlag (2007).
9. Kelman, H.C.: Continuity and change: My life as a social psychologist. In: Eagly, A.H, Baron, R.M., Hamilton, V.L. (eds.) The social psychology of group identity and social conflict: Theory, application, and practice. American Psychological Association Press, pp. 233--275 (2004).
10. Dastani, M., Dignum, V., Dignum, F.: Role-assignment in open agent societies. In: Proceedings of the 2nd international joint conference on Autonomous agents and multiagent systems, AAMAS'03, pp. 489--496, ACM Press. Melbourne, Australia, (2003).
11. Dignum, V.: A model for organizational interaction: based on agents, founded in Logic. PhD Thesis. Utrecht University, The Netherlands (2003).
12. Anderson, J. R.: Rules of the Mind. Lawrence Erlbaum Associates (1993).
13. Laird, J. E., Newell, A., Rosenbloom, P. S.: SOAR: an architecture for general intelligence. Artificial Intelligence, 33(1), 1--64 (1987).
14. Sun, R.: The CLARION cognitive architecture: Extending cognitive modeling to social simulation. In: Sun, R. (ed.) Cognition and Multi-Agent Interaction, pp.79--99. New York, Cambridge University Press (2006).
15. Rao, A., Georgeff, M.: BDI-agents: from theory to practice. In: Proceedings of the 1st International Conference on Multiagent Systems, 312--319 (1995).
16. Bratman, M.: Intention, Plans, and Practical Reason. Harvard University Press (1987).
17. Ortony, A., Clore, G. L., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press (1990).
18. McNaughton, N.: Biology and Emotion. Cambridge University Press (1989).
19. Revelle, W.: Personality Processes. Annual Review of Psychology. 1, 295--328 (1995).
20. Funder, D. C.: The Personality Puzzle. 4th Edition. W. W. Norton Press (2006).
21. Zeisset, C.: The Art of Dialogue: Exploring Personality Differences for More Effective Communication. Center for Applications of Psychological Type Press (2006).
22. Digman, J.: Personality Structure: Emergence of the Five-Factor Model. Annual Review of Psychology, 41, 417--440 (1990).
23. Millon, T.: Reflections on Psychosynergy: A Model for Integrating Science, Theory, Classification, Assessment, and Therapy. Journal of Personality Assessment, 72, 437--456 (1999).
24. Myers, I. B., McCaulley, M. H, Quenk, N. L., Hammer, A. L.: MBTI Manual: A guide to the development and use of the Myers Briggs type indicator. Consulting Psych. (1998).
25. Damasio, A.: Descartes Error Emotion, Reason, and the Human Brain. G P Putnam, (1994).
26. Andrighetto, G., Campenni, M., Conte, R., Paolucci, M.: On the Immergence of Norms: A Normative Agent Architecture. AAAI Fall Symposium. Emergent Agents and Socialities: Social and Organizational Aspects of Intelligence, 11--18 (2007).
27. Dignum, F., Kinny, D., Sonenberg, E.: From Desires, Obligations and norms to Goals. Cognitive Science Quarterly Journal, 2, 407--430 (2002).
28. Jiang, H., Vidal, J. M., Huhns, M. N.: EBDI: an architecture for emotional agents. In: Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems, AAMAS'2007, 38--40, ACM Press. Honolulu, Hawaii (2007).
29. Silverman, B., Johns, M., Cornwell, J., O'Brien, K.: Human Behavior Models for Agents in Simulators and Games. In Presence: Teleoperators and Virtual Environment, 15, 139--162 (2006).
30. Dastani, M.: 2APL: a practical agent programming language. Autonomous Agents and Multi-Agent Systems, 16, 214--248 (2008).