HAL
open science

Aliveness metaphor for an evolutive gesture interaction based on coupling between a human and a virtual agent

Pierre de Loor, Romain Richard, Julien Soler, Elisabetta Bevacqua

To cite this version:

Pierre de Loor, Romain Richard, Julien Soler, Elisabetta Bevacqua. Aliveness metaphor for an evolutive gesture interaction based on coupling between a human and a virtual agent. 29th International Conference on Computer Animation and Social Agents, CASA'2016, Nadia Magnenat Thalmann, Marina L. Gavrilova May 2016, Genève, Switzerland. pp.147-155, 10.1145/2915926.2915932 . hal-01322244

HAL Id: hal-01322244

https://hal.science/hal-01322244v1

Submitted on 30 May 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Aliveness metaphor for an evolutive gesture interaction based on coupling between a human and a virtual agent

Pierre De Loor
Lab-STICC, ENIB, UBL
deloor@enib.fr
www.enib.fr/~deloor

Romain Richard
Lab-STICC, ENIB, UBL
romain.richard@enib.fr

Julien Soler
Virtualys
julien.soler@virtualys.com
www.virtualys.com

Elisabetta Bevacqua
Lab-STICC, ENIB, UBL
bevacqua@enib.fr
www.enib.fr/~bevacqua

ABSTRACT

This paper presents a model that provides adaptive and evolutive interaction between a human and a virtual agent. After introducing the theoretical justifications, the aliveness metaphor and the notion of coupling are presented. Then, we propose a formalization of the model that relies on the temporal evolution of the coupling between participants and the existence of phases during the interaction. An example on a fitness exergame is provided and some illustrations show the behavior of the model during an interaction. A video complements this example.

Keywords

cognitive architecture, human-agent interaction, nonverbal communication, coupling

1. INTRODUCTION

An important challenge to face when designing virtual interactive agents is the issue of believability, not as much in the physical aspect as in the behavior performed during the interaction with a human user. We reckon that, to sustain natural and believable interactions, the agent should be able to adapt to the behavior performed by the human. Indeed, several researchers have shown that human-human interaction is a dynamic process in which interactants influence each other's behavior [1] and they tend to adapt both their verbal and non-verbal behavior [2]. This mutual influence appears on different levels, such as temporal, acoustic, non-verbal, verbal, and linguistic. Studies have shown, for example, that during an interaction individuals tend to synchronize their movements [3] and to mimic their partner's non-verbal behaviors (such as facial expressions, body and leg positions, gestures [4]). Such a mutual influence is then a crucial element in human-human interactions and people are able (consciously or not) to actively perceive and regulate it. As argued in [5], it is when the coordination is not working that people become consciously aware of it.

Previous works which tackle the implementation of believable human-agent interaction focused mainly on conversational interactions and the behavior the agent should perform, for example, while listening [6, 7, 8] or managing the turn [9, 10]. These works are concerned more with spotting the right time the agent should react than with the type and the quality of the behavior the agent should perform according to the human's behavioral signals. Mutual influence appeared just on a temporal level. Moreover, most of the existing approaches are generally static and based on a cognitivist perspective: to the perception of a signal corresponds a response generally described via some rules [6, 9, 10, 8] or probabilistic mechanisms [7]. The problem is that the mutual influence of the interacting entities and their temporal regulation are not the starting point of the definition of these rules. The main focus is on how to model the choice of the behavior to perform according to the input but not on the evaluation of the on-going interaction. Then, regulation and mutual influence are not guaranteed to appear and if they do, it is an indirect consequence of the rules. If an unexpected reaction occurs, regulation will certainly be broken and so will be the believability of the virtual character. The problem is that it is very hard to anticipate all the combinations of input which could brake the regulation if this regulation is not explicitly expressed into rules which evaluate the interaction.

In this work we would like to present a different approach that focuses on the evolution of the interaction itself by taking into account the correlation between the behavior performed by the human and that performed by the agent. Differently from previous works, we follow the enactive perspective, which considers that it is not possible to separate perception and action as traditional methods do when they compute output from input [11]. The difference between classical input/output approaches and our proposition is illustrated in figure 1. Due to the complexity of reproducing a full human-human interaction, here we focus only on bodily

CASA 2016, May 23-25, 2016,

ISBN: 978-1-4503-2138-9.

DOI: http://dx.doi.org/10.1145/2915926.2915932

interaction and we do not consider verbal communication at all. We will describe a model which enables a virtual agent to adapt to its partner's behavior in scenarii where speech is not mandatory to establish an interaction. A similar work was presented by Pugliese and colleagues [12]. Their approach, based on enaction too, mapped the human's quality of motion onto a bi-dimensional space where motion capture animations for the virtual agent where placed according to their motion style. So the agent played back the animation that fitted the human's behavior the most. Differently, we suggest to select the agent behavior not only according to the quality of that performed by the human but also according to an evaluation of the on-going interaction.

In the next section we describe the principles our work is based on. In section 3 a formalization of the model is presented and section 4 shows a real application of this model. We conclude this paper by offering some perspectives for future improvements.

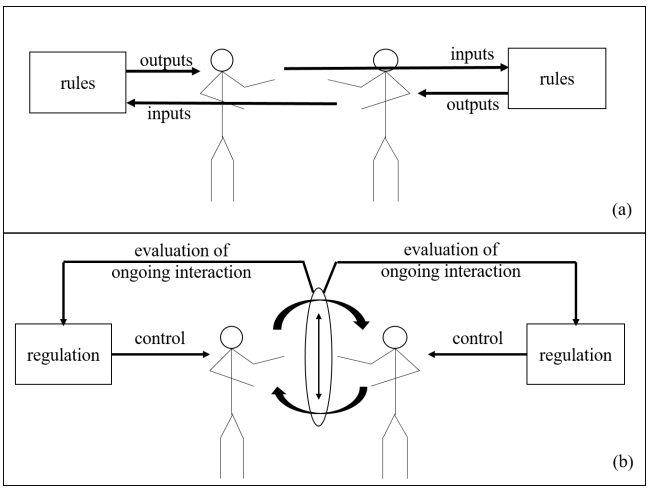

Figure 1: cognitivist (a) versus enactive (b) approaches

2. PRINCIPLES: ALIVENESS METAPHOR AND COUPLING

To design our model we were strongly inspired by the theoretical method to study human-human communication suggested by Fogel and Garvey [13]. Their method is based on the alive communication metaphor (or aliveness metaphor) which stresses the analogy between living phenomena and communication as their temporal self-regulated and continuous changing aspects. According to this method, a communication can be described through three states, called *co-regulation*, *ordinary variability* and *innovation*. Co-regulation appears when interactants are open to mutual influence and their coordinated co-actions create constantly new information constructing the communication. However, even when two people involved in an interaction experienced co-regulation, their co-actions are not always produced in the same manner. This is because in human-human interaction there is an intrinsic variability. According to Burgoon et al. [14, page 269], for example, variability increases when the interaction runs smoothly and people are less restrained by social and cultural norms and are freer to express their individual differences, such as behavioral style, goals and desires. This *ordinary variability* makes interactions 'never the same' but never nonsense. From time to time, extraordinary variability (called *innovation*) can appear, obliging the interactants to make an effort to integrate the novelty in order to keep the communication going. Innovation is important also because, without important changes, an interaction could get less interesting, thus the interactants have the tendency to introduce something new in order to maintain their engagement [15].

We introduce a model able to analyze the low-level of the on-going interaction between a virtual agent and a human in order to determine the state it is passing through: Are the interactants co-regulated? Are they showing variability or innovation? Then, this information is used to decide the agent behavior. To determine the running state of the interaction we use the concept of *coupling*. Coupling is defined as the continuous mutual influence between two individuals [16]. It possesses the capability to resist to disturbances, and compensates them by making the interaction evolve. Disturbances come both from the environment and from within the individuals. This definition is recursive since coupling exists because of the interactants' effort to "recover" it as its quality decreases. In [16] this effort is considered as an explanation of the improvement of the feeling of co-presence and engagement. Interactants measure coupling unconsciously and unconsciously they regulate it. We suggest here to use coupling as the variable that informs the model about the on-going interaction. One can make the link with the notion of ultrastability in cybernetic [17] which considers that biological systems regulate themselves through the observation of essential variables. Here, the essential variable is represented by the value of the coupling. Decisions are then described in terms of their ability to modify the coupling in real-time. The agent behavior is chosen according to its potential to make the interaction pass in one of the previous cited states: co-regulation, variability or innovation.

Another important characteristic of our model is the ability to evolve together with the interaction itself. The behavior performed by the agent is always influenced by what happened before and, whenever the innovation state is attained, the introduced novelty is injected and assimilated to become ordinary interaction.

3. FORMALIZATION

Relevant symbols of this formalization are defined in Table 1.

3.1 Aliveness states

The state of the interaction is modeled by a probability distribution of a discrete random variable S. This variable can take 3 values that represent respectively the fact that the interaction is in *co-regulation* state, in *variability* state or *innovation* state. The probability distribution is defined by the vector $PR = [pr_1, pr_2, pr_3]$, where pr_1 is the probability $\mathbb{P}(S = regulation)$, pr_2 the probability $\mathbb{P}(S = variability)$ and pr_3 the probability $\mathbb{P}(S = innovation)$. These probabilities evolve in time according to the coupling.

3.2 Evaluation of the coupling

As explained in section 1, the present work focuses on gestural behavior and gestural expressive quality shown by both interactancts. For such a reason we compute coupling

Table 1: Symbols Table

$\beta_i(t)$	tendency to innovate (at time t)
$\beta_v(t)$	tendency to vary (at time t)
$f_i(p,t)$	value of feature i of participant p at time t
$C(p_1,p_2,t)$	coupling at time t, between participants p_1 and p_2
$C_W^{p_1p_2}(T)$	mean coupling between participants p_1 and p_2 on a temporal windows T
th_{β_v}	threshold between successive coupling to increase variability probability
α_i	rate of innovation tendency increase
α_v	rate of variability tendency increase
$F_W^p(T)$	mean value of weighted features on a time interval T for participant p
initiator	the participant whose features evolve the most
S	discrete random variable that represents the state of the interaction (among (*co-regulation, variability, innovation*))
PR	vector $[pr_{cr}, pr_v, pr_i]$ of probability distribution of S
pr_1	probability to be in co-regulation
pr_2	probability to be in variability
pr_3	probability to be in innovation
δ_{p_i}	variation of weighted features of p_i between two temporal windows
τ	positive real that represents the strength of the influence of tendencies on probabilities pr_i

as the measure of the difference between some relevant behavioral features. The most relevant features depend on the nature of the interaction. For example, they are different for cooking together, playing tennis or dancing.

These features are real values extracted from the skeleton's movements of the interactants. Many features can be considered, such as hands and feet position, speed, energy, spatial extent, directness, impulsiveness, jerkiness of movements and more [18, 19, 20]. Here we introduce also one symbolic feature which is the identifier of the recognized gesture performed by the human. It is determined with the real-time gesture recognition algorithm described in [21]. This symbolic feature can be utilized for interactions where the shape of the gesture is important, such as dance for example.

Let us consider n behavioral features for each participant to the interaction: $f_i(p,t)$, where p is the participant, t the discrete time (in practice, the time step is 30ms), and i the index of the feature $(0 < i \leq n)$

We define a normalization function:

$$nm : \left| \begin{array}{lcl} \mathbb{R} & \longrightarrow & [0,1] \\ x & \longmapsto & |x|/(1+|x|) \end{array} \right.$$

that is used to compute $C(p_1,p_2,t)$, the global coupling of two participants p_1 and p_2 at time t, as:

$$C(p_1,p_2,t) = 1 - \left\{ \frac{\sum_{i=0}^{n} w_i * (nm(f_i(p_1,t) - f_i(p_2,t)))}{\sum_{i=0}^{n} w_i} \right\} \quad (1)$$

The weight w_i represents the relevance of the ith feature for computing the global coupling. It depends on the type of the interaction and its value w_i is limited between 0 and 1.

For example, in an imitation game, where each participant has to imitate the posture of the other, the weights of the features relative to the positions of the joints of the skeletons should be set to 1 while the other weights to 0. In a fitness exergame, where the quality and the dynamic of the movements are important, the weights could be distributed among features such as the spatial extent, the speed and the identifier of the performed gesture.

Formula 1 guarantees that $0 \leq C(p_1,p_2,t) \leq 1$. Note that for the symbolic feature corresponding to the *id* of a gesture, the difference is set to 1 if the gesture performed by p_1 and p_2 is the same, to 0 otherwise.

Now, let T be a temporal window of 30 frames, such as $T = [t_j, t_k]$, with $t_j < t_k$. We compute the mean of the coupling as follows:

$$C_W^{p_1p_2}(T) = \frac{1}{T} * \sum_T C(p_1,p_2,t) \quad (2)$$

We compute the mean value of coupling on a temporal window in order to filter noise in the data, such a noise is generally due to sensor flaws or to gesture recognition errors.

3.3 Temporal evolution

As mentioned in section 3.1, the probability distributions of S evolves in time according to the history of the coupling. To represent this dynamical evolution, we use two variables $\beta_v(t)$ and $\beta_i(t)$ that correspond respectively to the tendency to introduce variation and to innovate. Tendencies represent the psychological part of the metaphor in the sens that they model the need for changing (by introducing variation or innovation) to maintain the aliveness of the interaction. Considering two consecutive time windows $T_1 = [t_j, t_k]$ and $T_2 = [t_k, t_l]$, where $t_j < t_k < t_l$, the increase of $\beta_v(t)$ and $\beta_i(t)$ is computed as follows:

if $|(C_W^{p_1p_2}(T_1) - C_W^{p_1p_2}(T_2))| \leq th_{\beta_v}$ **then**
  $\beta_v(t) \leftarrow \beta_v(t-1) + \alpha_v * C(p_1,p_2,t) * (1 - \beta_v(t-1))$
else
  $\beta_i(t) \leftarrow \beta_i(t-1) + \alpha_i * (1 - \beta_i(t-1))$
end if

th_{β_v} is a tolerant threshold used to determine when the means coupling computed on two consecutive temporal windows are similar enough to increase the tendency to introduce a variation. α_v is the rate which this tendency increases at. Similarly, α_i is the rate which the tendency to innovate increases at.

In short, if the mean coupling does not change during

time, the tendency to make some variations increases. Otherwise, it is the tendency to innovate ($\beta_i(t)$) that increases (but more slowly since $\alpha_i < \alpha_v$). At each time step, these tendencies are evaluated and so does the probability distribution vector PR (see section 3.1). To do that, we use a *softmax* function that "squashes" the vector $[z = (C(p_1, p_2, t), \beta_v(t), \beta_i(t)]$ in PR:

$$PR(z)_j = \frac{e^{z_j}/\tau}{\sum_{k=1}^{3} e^{z_k}/\tau}; j = 1...3 \quad (3)$$

τ is a real that enables the adjustment of the influence exerted by z on PR. Using a *softmax* function offers two advantages: The first one is the guaranty that the sum of the probabilities is equal to 1, the second one is that it allows the adjustment of the influence level of the variables of z on the probability values of PR. For example, for a value of $\beta_i(t)$ and $\beta_v(t)$ close to 1, if $C(p_{r1}, p_{r2}, t)$ is close to 0, then p_{r1} and p_{r2} are close to 0.5 and p_{r3} is close to 0. However, for the same values of $\beta_i()$ and $\beta_v(t)$, but $C(p_1, p_2, t)$ close to 1, p_{r1}, p_{r2} and p_{r3} are all close to 0.33. The value of τ allows to set this adjustment for intermediate values.

After the computation of PR, the state S of the interaction, is determined. We will see in section 3.4 that when the value of S changes, the behavior of the virtual character changes too. In this case, the value of the tendencies $\beta_i(t)$ and $\beta_v(t)$ is modified:

when S passes to *variability*
  $\beta_v(t) \leftarrow 0$
  $\beta_i(t) \leftarrow \beta_i(t-1) + \alpha_i * (1 - \beta_i(t-1))$

when S passes to *innovation*:
  $\beta_i(t) \leftarrow 0$

In short, each time the interaction passes in *variability* (because something new appears in term of coupling), the tendency to variate is reset and the probability to innovate increases. The same principle is used for the tendency to innovate. Indeed, when an innovation (or variation) is triggered, the need to reach the corresponding state is satisfied and then the tendency to reach it again is reset. After some time, the tendency to innovate or to introduce a variation will increase again, according to the evolution of the coupling. At any moment, when the interaction passes to variability state or to innovation state, we can determine which interactant causes the change. We call them the *initiator*. To explain how we determine the initiator, firstly we need to introduce the formula $F_W^p(T)$, that is the mean value of the weighted features of a participant p on a time windows $T = [t_j, t_k]$ with $t_j < t_k$.

$$F_W^p(T) = \frac{1}{T} * \sum_{t \in T} \left\{ \frac{\sum_{i=0}^{n} w_i * nm\left(f_i(p,t)\right)}{\sum_{i=0}^{n} w_i} \right\} \quad (4)$$

Let p_1 and p_2 be two participants and T_1 and T_2 two timed windows partially overlapped, such as $T_1 = [t_j, t_l]$ and $T_2 = [t_k, t_m]$ with $t_j < t_k < t_l < t_m$, we determine the initiator as follows:

$\delta_{p_1} \leftarrow (F_W^{p_1}(T_1) - F_W^{p_1}(T_2))$
$\delta_{p_2} \leftarrow (F_W^{p_2}(T_1) - F_W^{p_2}(T_2))$
if $(\delta_{p_1} < \delta_{p_2})$ **then**
  initiator $\leftarrow p_2$
else
  initiator $\leftarrow p_1$
end if

In other words, the initiator is the participant whose relevant features have evolved the most.

Whenever the interaction passes through one of the three states, the behavior that the agent should show depends strongly on the type of scenario which the interaction takes place in. For such a reason, the programmer must define three generic functions, called $BehavCo$, $BehavVar$ and $BehavIn$, which describe the behavior that can be performed respectively during the states of co-regulation, of variability and of innovation. Section 4 will provide examples about the definition of these three functions.

3.4 Toward a more evolutive model

The previous part of our model enables an adaptation of the interaction through the three states of the aliveness metaphor, according to the value of the coupling between the two participants. Since this change is relative to the coupling, it appears very adaptive. It can be described as follows: interaction states change in time and so does the behavior of the virtual agent. Reciprocally, the behavior of the other interactant (human or virtual) evolves and by doing so it changes the state of the interaction. However, it is not enough. We want our model able to make the variability evolve as well, that means that each time the interaction passes to the variability state, the model does not compute always the same kind of variability in the agent behavior. This variability should depend on the type of interaction and on the behavior performed by the other. We are well aware that the variability depends also on the agent characteristics (such as behavioral style, goals, desires, mood and so on), but in this work we do not take that into account. We want also our model able to assimilate innovations to make them part of the ordinary interaction. An unexpected behavior is not surprising anymore when it appears a second time, it becomes part of the interaction. To sum up, we want that, even if the behavior that the agent can show in any state of the interaction is defined by the programmer in the three functions described above ($BehavCo$, $BehavVar$ and $BehavIn$), these functions do not produce always the same output for the same input and that they can be parameterised according to the evolution of the interaction.

To reach this goal, we use a classifier system. A classifier system is a set of weighted rules (which are a couple (condition-action)) [22]. The probability to choose one rule depends on the value of its condition and its weight. Generally, weights are computed through a reinforcement learning algorithm and rules are generated through a genetic algorithm. However in this work, we do not introduce reinforcement learning and rules are handwritten. The condition of a rule takes into account the other interactant behavioral features $f_i(p,t)$), the level of coupling $C(p_1, p_2, t)$ and the *initiator*; the actions change some parameters of the behavior the agent will perform, for example they will increase the speed of movements or their amplitude. Thus, the rules can modify, in real-time and according to the interaction,

```
<classifier>
<!-- Intern parameters -->
    <param [dynamic] />
<!-- Rules -->
    <rule [weight]>
       <condition/>
       <action/>
    </rule>
<!-- Rewards -->
    <reward [value]>
       <condition/>
    </reward>
</classifier>
```

Figure 2: structure of the XML file classifier

the parameters of the three generic functions ($BehavCo$, $BehavVar$ and $BehavIn$) which describe the behavior that can be performed during each interaction state.

Concretely, the user has to configure an `XML` file with customized tags. The tags are shown in Figure 2. Through the tag `param` one can define the variables that are implied as parameters for the conditions. If a variable is declared dynamic (`dynamic=true`), it is possible to known its temporal evolution to evaluate rules condition. Any variable of the full program could be used. The tag `rule` is used to define the rules. Conditions can use different operators of comparison (i.e `greater, lower, equal ...`) or evolution (like `increasing, decreasing and steady`) for dynamical variables. Actions contain operators to modify the parameters (i.e. `add, sub, divide, multiply, set`).

4. APPLICATIONS

In order to use our model in real human-agent interactions, we need to parameterize it according to the specific scenario which the interaction will take place in. We must specify:

1. the most important features that are needed to compute the level of coupling: the values of all w_i in formula 1,

2. the three functions $BehavCo$, $BehavVar$, $BehavIn$ and their parameters which can evolve during the interaction,

3. the rules of the classifier which describe how these parameters evolve.

It could seem as an heavy task, however it is not so cumbersome. To show that, we used our model to make a virtual companion interact with a human within an exergame.

4.1 Fitness exergame

Exergames for sport are flourishing thanks to the more and more affordable body sensors used in edutainment. However, the coupling between a virtual agent and a human during a training session is far from being simple and correctly simulated. The goal of this interaction is to have the virtual agent playing the role of a fitness companion who shows some fitness movements that the human must reproduce. The movements that the agent can show were

Figure 3: three screen shots from the video of the exergame. The human is tracked using a kinect.

collected through motion capture. A set of 12 fitness movements was captured. The agent can regulate its expressivity (in term of speed) in real-time according to that showed by the human. He can also offer an innovation by introducing a new fitness movement; as soon as a movement has been performed by both the agent and the human, it becomes a "known movement" and cannot be considered as innovating any more. Variability is offered by the agent either by varying its expressivity (always in term of speed) or by changing the current movement with another known one.

To make a virtual agent play the fitness game, we parameterise our model as follows:

1. *What are the important features used to compute the coupling?* Fitness practice relies on predefined body movements and on the manner they are performed. Then, the weights of the features corresponding to the recognized gesture, and some features relative to the effort provided by the performer (such as speed) are set to a positive value. The weight of the other features is set to 0.

2. *How do we specify the three functions $BehavCo$, $BehavVar$, $BehavIn$ and their parameters?* The role of $BehavCo$ is that, when the interaction is in co-regulation state, the virtual agent, who is performing a fitness movement, modulates its speed to be similar to that showed by the human. The interaction can move to variability state whenever an important change appears in the expressivity of the movement. Such a change could be offered by the agent or by the human; as we have seen in section 3.3 we can determine who initiates the modification in the interaction and we called it the *initiator*. So, whenever the interaction passes to variability state, if the initiator is the agent, the $BehavVar$ chooses either to vary the speed of a random amount selected within a given range, or to change the current movement with another known one. Otherwise, if the human is the initiator, the virtual agent modulates the movement it is performing according to the value of speed computed by the classifier system.

3. *How can the parameter of speed evolve?* There are two basic classifier rules: one specifies that if the coupling

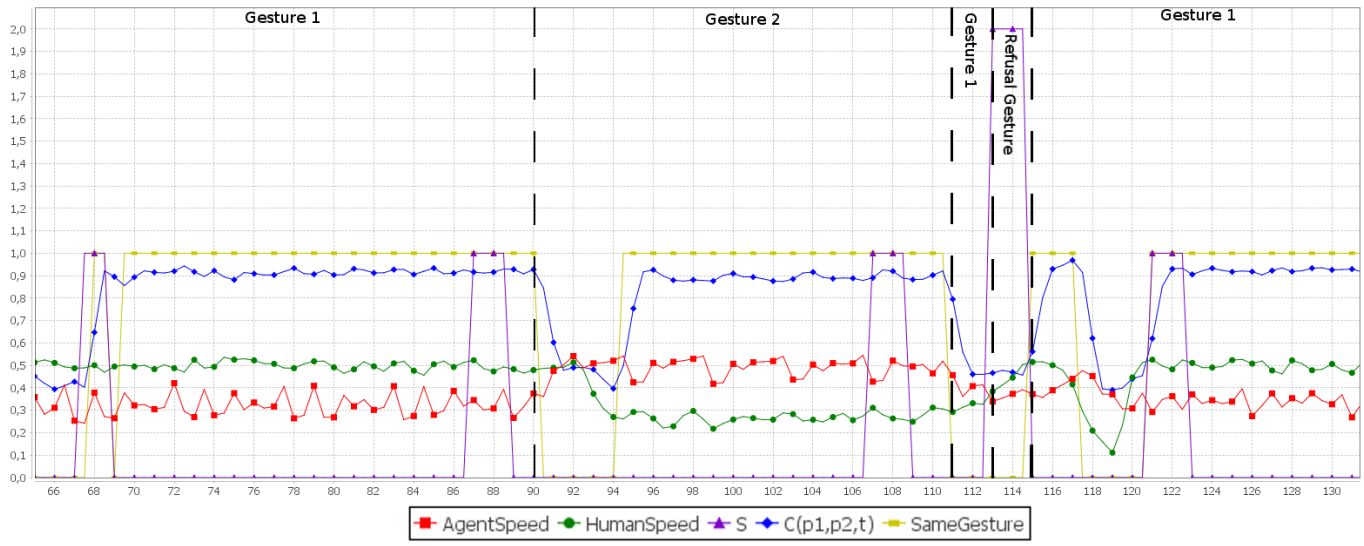

Figure 4: Traces of the execution of an episode of interaction. The value of S is 0 for *co-regulation*, 1 for *variability* and 2 for *innovation*

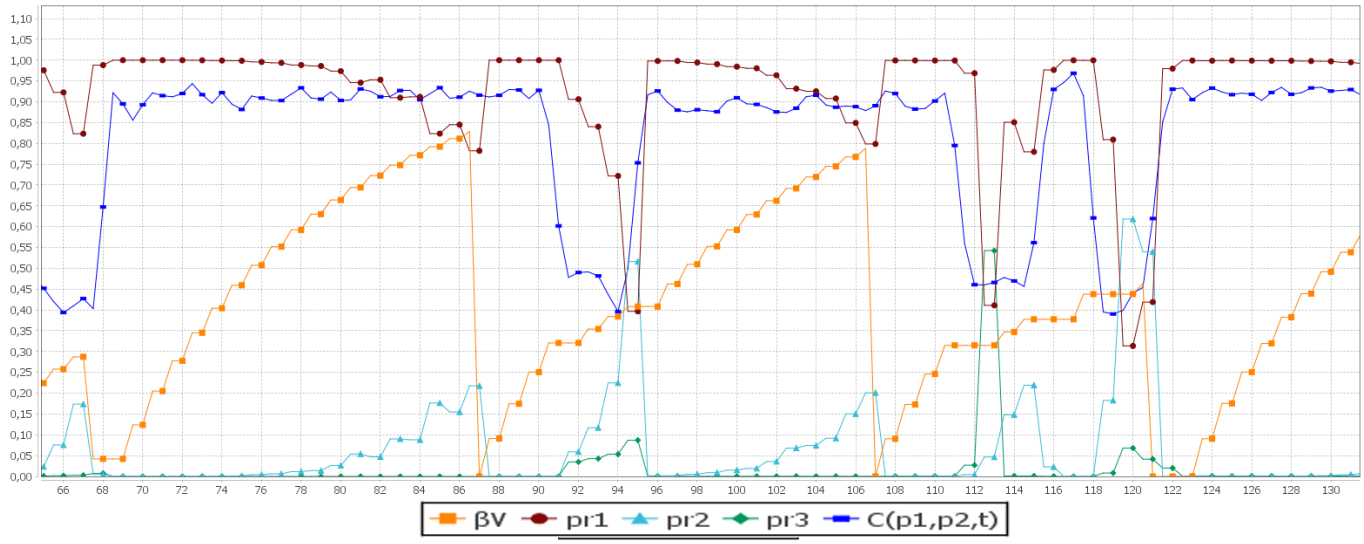

Figure 5: Same episode as Figure 4 for the evolution of probability distributions according to the coupling and the tendency to pass to variability state

is good for awhile, the range of the possible values for the random change of speed will grow. Reciprocally, another rule indicates that if the coupling is low, this range will decrease. Of course, several other rules could be imagined. The purpose here is to illustrate the possibility to introduce some evolution of the behavior during time.

Figure 6 shows the rules in the classifier which specify the variability of the speed. `Speed` and `deltaS` correspond respectively to the range of speed variability of the virtual agent and the difference between the speed of the human and the agent. Whenever the interaction passes to innovation state, if the agent is the *initiator*, the $BehavIn$ function introduces a new fitness movement, that is a movement that has never been performed before. If the human is the *initiator*, the agent refuses the proposition by randomly maintaining its current movement or playing a gesture of refusal. In this game just the agent knows all the movements that can be performed. Such a limitation assures that the human does not introduce movements that the agent could neither recognize nor reproduce.

4.1.1 Results

Figure 3 shows the technical realization of our example (we encourage the reader to have a look at the video posted

```
<classifier>
<!-- Intern parameters -->
 <param tag="deltaS" dynamic="true" value="0.0" />
 <param tag="speed" dynamic="true" value="0.0" />
<!-- Rules -->
 <rule weight="1">
     <condition param="deltaS" compare="greater"
     to="0.0"/>
     <action type="add" param="speed" value="0.01"/>
 </rule>
 <rule weight="1">
   <condition param="deltaS" compare="lower"
   to="0.0"/>
   <action type="sub" param="speed" value="0.01"/>
 </rule>
</classifier>
```

Figure 6: Example of the classifier rule that modifies the range of speed variability.

on the project web site[1]): Two girls played the fitness game with the agent. Figures 4 and 5 show the temporal evolution of different variables of the model during an episode of the interaction. For this example, we chose τ=0.07, α_v=0.1, α_i=0.08 and th_{β_v}=0.15.

This episode exemplifies the possibilities of our model. In figure 4, we trace the features that correspond to the speed of the agent and the human. We can observe that their evolution are linked together: sometime the agent moves faster than the human but, some times, it is the opposite. However, even if it is quite subtle, the speed variation of the agent follows that of the human. It also appears that if the gesture of the agent and the human is not the same (curve *sameGesture* from time 90 to time 94), the coupling decreases. When the coupling is high (for instance from time 70 to time 90), the $\beta_v(t)$ increases (on figure 5) and so increases pr_2. Then after awhile, S enters in *variability* state (time 86) and the agent offers another fitness movement (gesture 2 starts effectively at time 90, after the end of gesture 1). The coupling starts decreasing, but, at time 94, the human plays gesture 2 too and the coupling increases anew. One can observe that each time S enters in the variability state, $\beta_v(t)$ is reset to 0 (and so is $pr_1(t)$). However, it can come true that $\beta_v(t)$ increases while pr_2 decreases. It is due to the effect of the *softmax* function (see section 3.3). For instance, at time 96, as the coupling and pr_1 increases sharply, pr_2 decreases, even if $\beta_v(t)$ does not really evolve.

Another interesting phenomenon appears at time 112 when the probability to innovate is strong enough to trigger an innovation. Since the *initator* (which is not traced for clarity of the figure) is the human, $BehavIn$ plays a gesture of refusal. So the girl is surprised and stops her movement. The speed of the human decreases, the coupling too, but, as soon as the virtual agent enter in co-regulation, it slows down to get closer to the girl's pace. Then, the human resumes the interaction by performing the right gesture (at time 120).

Figure 7 illustrates, on a longer period of the interaction, that the number of the "known movements" increases. We remind that a known movement is a movement which has been performed by the interactants at least once. Each novel gesture is progressively introduced whenever the interaction entered in *innovation* state and the agent was the initiator (that is at times 90, 120, 320, 390, 450). Figure 7 shows also that during an interaction there are generally some variability and co-regulation cycles before an innovation. Statistically, the interaction is in co-regulation 80% of time, in variability 13% of time and in innovation 7% of time. The introduction of all 12 fitness gestures takes generally 12 minutes.

These different traces show that our model generates an evolving and self-adapting behavior for the agent, as suggested by the aliveness metaphor.

5. PERSPECTIVES

This paper has presented an architecture able to induce behavioral coupling between a virtual character and a real human. Its originality relies on the introduction of the aliveness metaphor that considers the interaction as a self-regulated system. The "dynamic equilibrium" between the participants reproduces real life interactions. We consider that the problem of believability, feeling of co-presence and engagement of people faced to virtual agents could be improved through this proposition. The next step of this work is to perform some evaluations. We also plan to improve the definition of features. For instance, [23] introduces a multi-level analysis of motor actions of coaching that could be relevant for a high performance virtual coach (for the moment, coaching is used as an illustrative example of the principles of our architecture without claiming that it does what a real coach would do). We also plan to improve the animation. Rather than using tools as FinalIK[2], we plan to connect the cognitive architecture with interactive mesh techniques that can be generated from data [24]. To finish, we would like to improve the rules of the classifier by implementing a learning mechanism.

6. ACKNOWLEDGMENTS

This work was funded by the ANR INGREDIBLE project: ANR-12-CORD-001 (http://www.ingredible.fr).

7. REFERENCES

[1] J. N. Cappella. The biological origins of automated patterns of human interaction. *Journal on Communication Theory*, 1:4–35, 1991.

[2] H. G. Wallbott. Congruence, contagion, and motor mimicry: Mutualities in nonverbal exchange. *Journal on Mutalities in dialogue*, pages 82–98, 1995.

[3] W. S. Condon and W. D. Ogston. Sound film analysis of normal and pathological behavior patterns. *Journal on Nervous and Mental Disease*, 143:338–347, 1966.

[4] J.L. Lakin, V.E. Jefferis, C.M. Cheng, and T.L. Chartrand. The Chameleon effect as social glue: Evidence for the evolutionary significance of nonconsious mimicry. *Journal on Nonverbal Behavior*, 27(3):145–162, 2003.

[5] S.P. Gill. Entrainment and musicality in the human system interface. In *Ai & Society 25*, pages 567–605, 2007.

[1] http://www.ingredible.fr/?page_id=201

[2] www.root-motion.com/final-ik.html

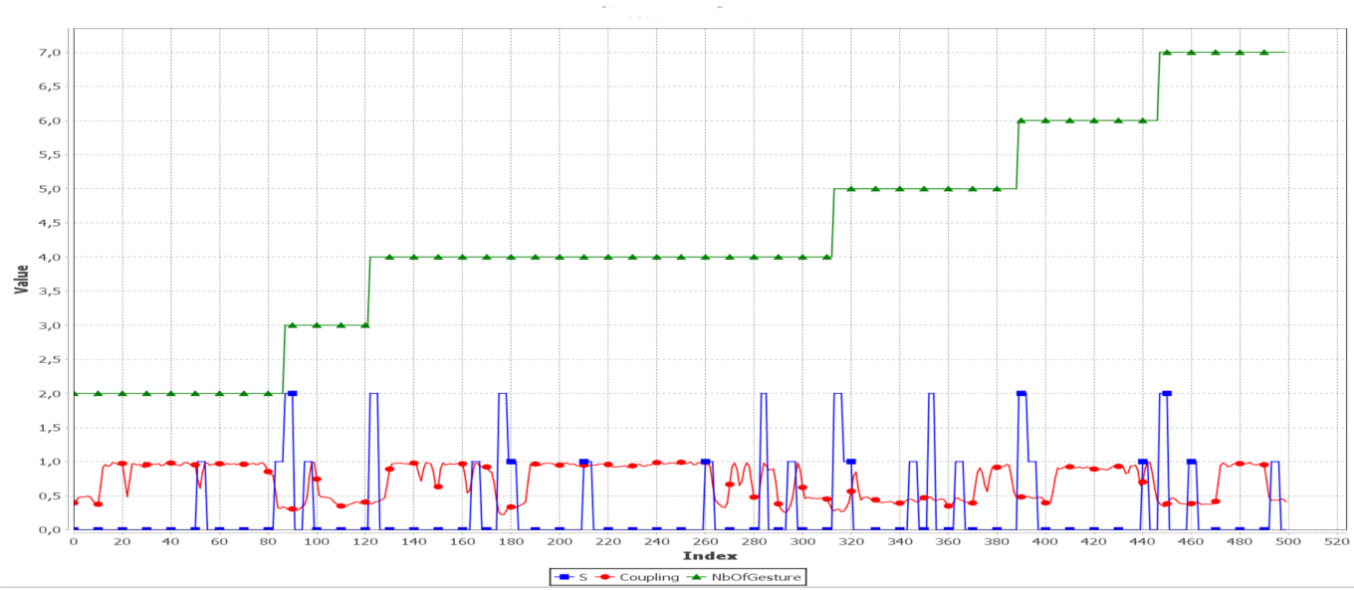

Figure 7: Evolution of the number of "known movements'

[6] S. Kopp, J. Allwood, K. Grammer, E. Ahlsen, and T. Stocksmeier. Modeling embodied feedback with virtual humans. In *Modeling Communication with Robots and Virtual Humans*, volume 4930 of *LNCS*, pages 18–37, 2008.

[7] L.-P. Morency, I. de Kook, and J. Gratch. Predicting listener backchannels: A probabilistic multimodal approach. In *8th International Conference on Intelligent Virtual Agents*, volume 5208 of *LNCS*, pages 176–190, 2008.

[8] E. Bevacqua, E. de Sevin, S. Hyniewska, and C. Pelachaud. A listener model: introducing personality traits. *Journal on Multimodal User Interfaces, special issue Interacting ECAs*, 6(1):27–38, 2012.

[9] K. R. Thorisson. Natural turn-taking needs no manual: Computational theory and model, from perception to action. *Journal on Multimodality in language and speech systems*, pages 173–207, 2002.

[10] M. ter Maat and D. Heylen. Development of multimodal interfaces: Active listening and synchrony. *Journal on Generating Simple Conversations*, 5967:92–101, 2010.

[11] F.J. Varela, E. Thompson, and E. Rosch. *The Embodied Mind*. MIT Press edition, 1991.

[12] R. Pugliese and K. Lehtonen. A framework for motion based bodily enaction with virtual characters. In *11th International Conference on Intelligent Virtual Agents*, volume 6895 of *LNAI*, pages 162–168, 2011.

[13] A. Fogel and A. Garvey. Alive communication. *Journal on Infant behavior & development*, 30(2):251–7, 2007.

[14] J.K. Burgoon, L.A. Stern, and L. Dillman. *Interpersonal Adaptation - Dyadic Interaction Patterns*. Cambridge University Press, 1995.

[15] M. Csikszentmihalyi. *Flow: The Psychology of Optimal Experience*. Harper Perennial, 1990.

[16] E. Bevacqua, I. Stanković, A. Maatallaoui, A. Nédélec, and P. De Loor. Effects of coupling in human-virtual agent body interaction. In *14th International Conference on Intelligent Virtual Agents*, volume 8637 of *LNCS*, pages 54–63, 2014.

[17] W Ross Ashby. *An introduction to cybernetics*. Chapman & Hall, London, 1956.

[18] M. Kapadia, I. Chiang, T. Thomas, N.I. Badler, and J.T. Kider. Efficient motion retrieval in large motion databases. *ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pages 19–28, 2013.

[19] D. Glowinski, M. Mortillaro, N. Dael, and A. Camurri. Towards a Minimal Representation of Affective Gestures (Extended Abstract). In *International Conference on Affective Computing and Intelligent Interaction*, pages 498–504, 2015.

[20] A.A. Samadani, S. Burton, R. Gorbet, and D. Kulic. Laban effort and shape analysis of affective hand and arm movements. *Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 343–348, 2013.

[21] C. Jost, P. De Loor, A. Nédélec, E. Bevacqua, and I. Stanković. Real-time gesture recognition based on motion quality analysis. In *7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN)*, pages 47–56, 2015.

[22] J.H. Holland. Genetic algorithms and classifier systems: Foundations and future directions. In *International Conference on Genetic Algorithms*, pages 82–89, 1987.

[23] F. Hülsmann and S. Kopp. Multi-level analysis of motor action as a basis for effective coaching in virtual reality. In *International Symposium on Computer Science in Sports*, volume 392, pages 211–214, 2016.

[24] D. Vogt, B. Lorenz, S. Grehl, and B. Jung. Behavior generation for interactive virtual humans using context-dependent interaction meshes and automated constraint extraction. *Computer Animation and Virtual Worlds, 20*, 1(26):227–235, 2015.