Simulation theory and anticipation for interactive virtual character in an uncertain world

Cédric Buche, Anne Jeannin, Pierre de Loor

Simulation theory and anticipation for interactive virtual character in an uncertain world

Journal:	*Computer Animation and Virtual Worlds*
Manuscript ID:	CAVW-11-0014
Wiley - Manuscript type:	Special Issue Paper
Date Submitted by the Author:	04-Mar-2011
Complete List of Authors:	BUCHE, Cédric; UEB - LISyC - ENIB, CERV Jeannin, Anne; UEB-LISyC-UBO, Computer Science De Loor, Pierre; UEB-LISyC-ENIB, Computer Science
Keywords:	anticipation, real-time interaction, decision making, behavioral model, virtual juggler, virtual character

Simulation theory and anticipation for interactive virtual character in an uncertain world

Cédric Buche
ENIB/UEB/CERV
buche@enib.fr

Anne Jeannin-Girardon
ENIB/UEB/CERV
anne.jeannin@innermess.fr

Pierre De Loor
ENIB/UEB/CERV
deloor@enib.fr

This paper deals with simulations of real-time interactive character behavior. The underlying idea is to take into account principles from cognitive science, in particular, the human ability to anticipate and simulate the world behavior. For that purpose, we propose a conceptual framework where the entity possesses an autonomous world of simulation within simulation,

in which it can simulate itself (with its own model of behavior) and the environment (with an abstract representation, which can be learnt, of the other entities behaviors). This principle is illustrated by the development of an artificial juggler, which predicts the motion of balls in the air and uses its predictions to coordinate its own behavior while juggling. Thanks to this model it is possible to add a human user to launch balls that the virtual juggler can catch whilst juggling.

Keywords: anticipation, real-time interaction, decision making, behavioral model, virtual juggler, virtual character, computer animation

1 Introduction

This study is focused on the real-time interaction between a virtual character, or agent, and a dynamic open world. In this world, real users are able to disturb, at any time, the behavior of the virtual character. In this case, using a precise representation of the behavior of the world is impossible. However, it is a very important challenge to develop such a kind of behavior in order to address complex sensorimotor interactions with humans for video games, virtual theater, sport or any application implying improvisation, adaptation or co-evolution between human and virtual creatures. Despite the availability of numerous propositions for interactive behavior in computer animation (see section 2), our goal is to use ideas and concepts from cognitive science to enhance credibility about interactions. To be more precise, focus is on the simulation theory, the human's anticipation ability and capacity to learn the world with which it is interacting. The result is that interactive characters can improve, in real-time, their behavior adaptation ability. This paper is organized as follows: an overview on interactive animation of virtual characters, and on main challenges in this field, is presented in section 2. It points out that, usually, the dynamics of the environment is pre-given and steady. Section 3 gives three concepts from cognitive sciences considered as important in human ability during interactions within an uncertain and variable environment. These concepts are anticipation, simulation and attention. Then, section 4 proposes a conceptual framework based on 3 parts: i) general knowledge about the environment which can be learnt during interactions, ii) simulation world, which allows the anticipation of the current interaction and the definition of the object of attention of the character and iii) control of the virtual agent in interaction with its own world, but based on the prediction issued from the simulation. An illustration of this model is provided in section 5 through implementation of an interactive juggling game. It shows the ability of the virtual juggler to adapt its reaction to various disturbances, to play with other virtual jugglers and also with a human player.

2 Interactive characters

Numerous investigations have been aimed at simulating the behavior of virtual characters in real-time. Several approaches dealt with the development of algorithms dedicated to the synthesis of the gesture quality [1, 2]. But, none of them took into account interaction abilities of the character. At the opposite, some models developed in robotics are interaction oriented and rely on cognitive science, but the problem of animation realism is not addressed [3]. In-between hybrids architectures can describe high level real-time reasoning, thanks to state machines, planning algorithms and synchronization mechanisms [4, 5]. Some other ones are rule-based [6], but, generally, the management of interactions introduces a bottleneck in term of the capabilities to take into account all possibles scenarios. In the domain of animated and conversational agent, interaction is more generally addressed. For instance, JACK is an architecture able to manage the dialog between two agents [7], REA [8] allows the inclusion of the user's gaze and provides algorithms to link voice to gestures. GRETA [9] communicates with complex emotions and MAX

[10] recognizes the hand gestures thanks to the treatment of data issued from a motion capture glove. [11] identify some subtle interactively contingent phenomenas during human interaction which lead to a *social resonance*. For instance, [12] presents a system for authoring interactive characters. ELCKERLYC [13] is an adaptation of SAIBA which is able to anticipate the behavior of a user to change the animation from a set of precomputed possibilities. Because it relies on anticipation, it is close to our work but limited by the use a predefined animations. Finally, close to our applicative example, [14] propose an architecture for the hand coordination of a virtual juggler. However, as these authors focused on important technical issues, some essential features of human interaction, addressed in cognitive science, were neither considered, nor made explicit by these numerous approaches. These features would be able, in the long run, to enhance credibility of the dynamics of interactive behaviors. In a first step, they can improve the adaptability of a virtual character to different types of disturbances issued from a poorly known world because of it's variability and its opening on humans.

3 Three notions from cognitive sciences

Cognitive science is a wide domain, enriched by many points of view. Here, focus is only on the three key concepts addressed in this study.

1. *Anticipation*: animals and humans use their memories of the past so as to anticipate the consequences of their actions and the behavior of those around them. Some philosophers put the anticipation at the basis of cognition [15, 16]. The phenomena of anticipation are held parallel to the reasoning and they allow active correction of the action [17].

2. *Simulation*: this concept is close to anticipation, but it explains how anticipation is per-

formed. Some psychologists and neuroscientists claim that the brain is a simulator for action in the environment [18, 19, 17, 20]. With simulation theories, anticipation is not a disembodied abstract and rational reasoning, but rather an active process based on the imagination of interaction with an imaginary world: it is an explicit internal simulation.

3. *Attention*: sensory anticipation includes the use of predictive environmental models to orient the entities' perceptions more effectively, especially in order to process expected event rather than to take into account the whole environment [18, 21]

4 Conceptual Framework

Our models are part of a conceptual framework described in Figure 1. It takes into account notions like anticipation and explicit internal simulation. To take a decision and to control its interaction into a *virtual world* (at the bottom of the Figure), an autonomous agent uses predictions provided by a simulation (the *imaginary world* in the middle of Figure 1), performed from approximate knowledge, *i.e.* this simulation is not the result of an analytic calculus from accurate physical features of the environment. These features are approximated in an *abstract world* (at the top of the Figure) and hence, some variations in the future of the *virtual world* are possible. Hence, the agent needs to perpetually correct its control through comparison of approximated anticipations against real perceptions (when they exist). The result is a possibility of error of estimation and then of failure during an interaction. Moreover, these failures are not arbitrary because they realize a natural feature: an approximation during anticipation. For instance, the more surprising a disturbance is, the less efficient the behavior is. In addition, section 5 shows that these approximations can be used to perform active perception by the virtual juggler, and thus reflect the concept of attention (see section 3).

Finally, the *abstract world* is a sum of approximative knowledges about the dynamical features of the world. Theses knowledges are learnt during interactions. Thus, the agent can adapt its worldview through experience. For that, different techniques from machine learning can be used (reinforcement, lazy learning, etc). This idea was used to define the behavior of virtual sheepdogs able to anticipate and to learn the decision making of virtual sheeps by the use of fuzzy cognitive maps [22]. Now, we will show that the conceptual framework presented here can be applied in a sensorimotor interaction context with humans.

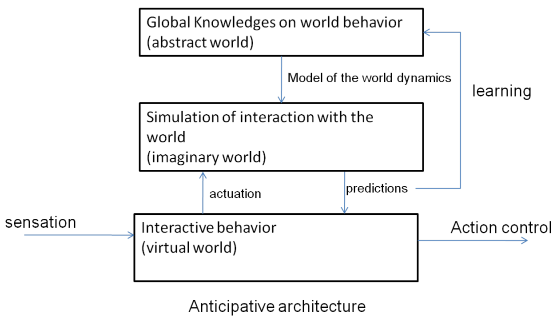

Figure 1: Conceptual framework for anticipative agents.

5 Example : interactive Juggler

The problem of virtual juggler was discussed in [23, 14]. But, in these approaches, neither the modeling of approximative anticipation nor the theory of simulation was taken into account. More generally, the relationships between cognitive sciences and character's behavior were not addressed. Here, we will show that the proposed conceptual framework can account for not only adaptation, but also plausible errors, through more or less predictable interactions, especially, with a real human character. An illustration of its application is presented in Figure 2. This application is called JABU: Juggler with Anticipatory Behavior in virtual Universe (see Figure 3).

The virtual world of the juggler has physical properties (inertia, gravity, wind, etc.) through the use of the ODE[1] physics engine. Of course, these quantities are not explicit in the model of

[1] Open Dynamic Engine, http://www.ode.org/

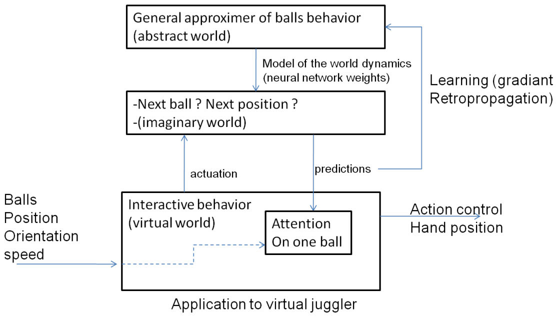

Figure 2: Instantiation of the framework for a virtual juggler.

Figure 3: The JABU application.

control. This control is adjusted through an attentional process focused on the *next* (anticipated) ball (actually one ball by hand). The approximate position of the balls is made by their simulation in the imaginary world of juggling. The function approximation properties of this imaginary world come from different neural networks. The abstract world corresponds to the weights of the arcs of these networks. Since they are universal approximators, we will see later that they allow real-time adaptation of the juggler gestures to different types of disturbances (this is also illustrated by the video associated with this article). The implementation of these principles is described hereafter.

5.1 Decision-making process

The hands have independent functions: this means that there are neither complex juggling moves nor tricks, but simply a succession of ball catches and throws, where each movement is independent of the others. As soon as a ball comes at the same height as the hands, it must be caught and thrown back. Hence, at this stage, the goal is not to get a realistic dynamics of gesture; no complex arm control model is used. Nevertheless, the time taken for a hand to move is not negligible and makes the juggler at risk of delayed move, which means missing the ball; this is also amplified by prediction errors. As mentioned above, the precise reproduction of the movement is not our priority and the hand's movement time is an empirically adjustable variable which reflects the delay between the decision being made and the action being carried out. In the following section, to facilitate the readability while keeping things brief, any reference to some hand activity means that the theoretical model was implemented for the anticipatory decision-making applied to our juggler.

The different phases of juggling are as follows. The juggler begins by looking for a ball in the air. Once the ball has been spotted, the hand has to be at an estimated reception point (prediction T1). Then, this reception point can be refined. In order to do so, the hand must estimate and correct the anticipated trajectory of the target ball (prediction T2) which is the *object of attention*. Each hand will therefore be able to catch or miss the target ball. If the ball is caught, the juggler will be able to throw it in the air. Whatever the future of the first ball (caught or missed), the juggler's hand once again starts looking for the next flying ball.

5.2 Predictions

Within the context of juggling, information must be gathered quickly in order to maintain the juggling dynamics. The use of perceptron-type neural networks (NNs) to make predictions about the trajectory is adequate. Indeed, NNs are quickly executed, and online learning occurs both quickly and effectively. Furthermore, NNs correspond to the need to manipulate (both spatial and temporal) digital data. It is, of course, also possible to use deterministic equation models of movement to make predictions. However, such precise predictions would be extremely noise-sensitive (disruption of the environment as the ball falls) and would not account for the use of approximations and readjustments in real-time which seem to be the basis of the anticipatory mechanisms that we aim to respect [17].

5.2.1 Prioritizing the balls (T1)

NN T1 provides the estimated temporal and spatial data for each ball at the moment it is thrown (see Figure 4). These data are used to categorize the balls and attribute them priorities so as to trigger the attentional process on the priority ball. The data required to calculate these estimations are the current speed of the ball and the height h at which the ball has to be caught (see Table 1).

Inputs	Outputs	Parameter	Objectives
Vx	Δt	h	Temporal classification
Vy	Δx		Vague
Vz	Δy		spatial prediction

Table 1: Inputs/Outputs of NN T1.Vx,Vy,Vz are the ball speeds along the three spatial axes, Δt is the time at which the ball is supposed to reach the point Δx and Δy at the height of the hand.

5.2.2 Refining the prediction of the target ball (T2)

NN T2 refines the spatial prediction about the place where a ball will fall while it is falling down

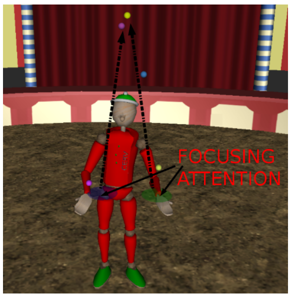

Figure 4: T1 estimates the position at which it will cross the hand plane (represented with circles) and how long it will take.

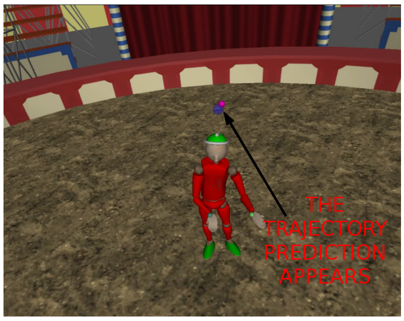

Figure 5: At any given time as the ball falls, T2 makes a more accurate estimation of its position in Δt seconds (represented by blurred ball).

(see Figure 5 and Table 2). Information can be obtained at different temporal levels (according to Δt).

Inputs	Outputs	Parameter	Objectives
Vx	Δx	Δt	Refined spatial predictions
Vy	Δy		
Vz	Δz		

Table 2: Inputs/Outputs of NN T2.

5.3 Interaction between virtual jugglers and with the human

The general features of this proposition allow interactions between several jugglers. To do that, the only change is the direction of the ball launched by each juggler (see Figure 6).

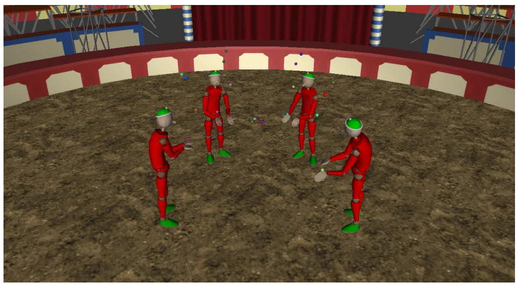

Figure 6: Multiple-jugglers.

The juggler can also catch a new ball thrown by a human user (Figure 7). This is useful for evaluating the believability of the virtual juggler (real-time decision-making, online adaptation, etc.). Introducing a human user also requires the introduction of a new type of prediction (T3). One should note that T3 is similar to the prediction T1, except that the ball is not thrown by the virtual juggler. The human user interacts with the virtual juggler by using a Wiimote (remote game controller from the Nintento Wii console). This peripheral device measures the movements of the human user's hand.

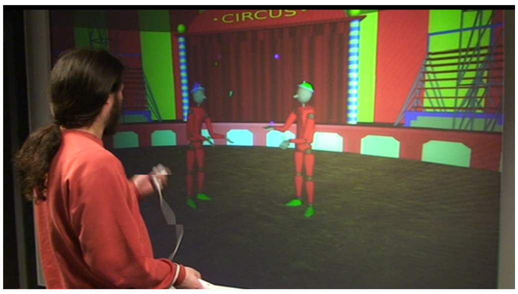

Figure 7: A human can juggle with virtual jugglers using the Wiimote.

5.4 Learning

The abstract world is represented by the functions encoded in the NN T1, T2 and T3. They are updated in line from the observation of several variables. In its example base, NN T1 has access to throws made by the juggler itself (low speed along x and y axes) whereas NN T3 records the balls thrown at a distance by a third person (much greater speeds). Each hidden layer has 19 neurons, which leads to 3x19x19x3 multilayer perceptrons. Learning is, thus, conducted with a maximum of 100 iterations using the FANN[2]. The parameters to be determined for the NN of T1 and the NN of T2 are h and Δt, respectively. In the example under study, $h = 2.5cm$ and $\Delta t = 0.1s$.

[2]Fast Artificial Neural Network (FANN) library available at http://leenissen.dk/fann/

5.5 Disturbing the environment online

This section will deal with the evaluation of the anticipatory mechanism with its qualities and impact on decision-making and the final result: the juggler animation. The generalization abilities of NN allow the in line adaptation of the juggler's motion to disturbances. For the tests, the initial conditions are varied over a given time. Moreover, 42 balls are thrown towards the virtual juggler (one ball every 0.75 seconds). The purpose is to observe the number of balls missed by the juggler (*i.e.* which fall below its knees and which it is unable to catch). Two other experiments consist in disturbing the juggler to validate its robustness to variability in the environment. At first, jerks are introduced in the projectile trajectories because they become maces rather than balls (see Figure 8). In this case-study, through the prediction by NN T1 is less accurate, NN T2 is able to correct it properly, and the juggler continues to juggle when balls are *transformed* in maces.

Second, gravity in the virtual environment is varied, and wind is added (see Figure 9). The juggler is not informed of these changes.

Figure 8: Juggling with maces.

Figure 10 shows the result for gravity variations. In abscissa, the different values of gravity in m/s^2. In ordinate, the number of balls which are dropped is an average over 10 tests of one minute each for each gravity value. One can observe that juggling is possible for gravitational values between 6 and 15 (normal gravity: 9.81). In cases of extremely low gravity, few balls are recorded as dropped, as they have not time to fall to the ground during the short simulation time. Figure 11 shows results for wind variations. The acceleration according to wind speed (in m/s^2, with direction indicated by positivity or negativity) is in abscissa. The number of dropped balls is in ordinate. The average values are taken for 5 simulations for each wind value. About juggling, the range of speeds in which the juggler continues to juggle correctly is much smaller (between $-0.2m/s^2$ and $+0.2m/s^2$).

Figure 9: Disturbing the environment conditions in line (wind, gravity).

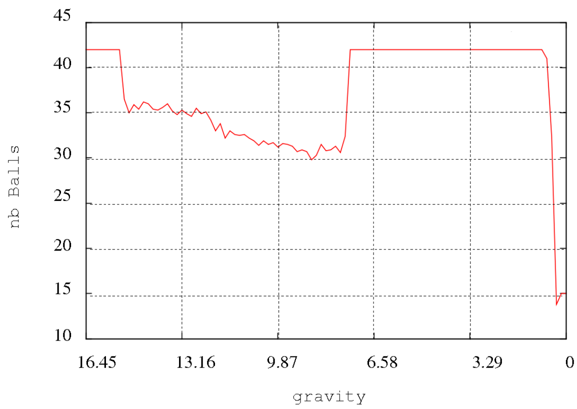

Figure 10: Average number of dropped balls according to gravity.

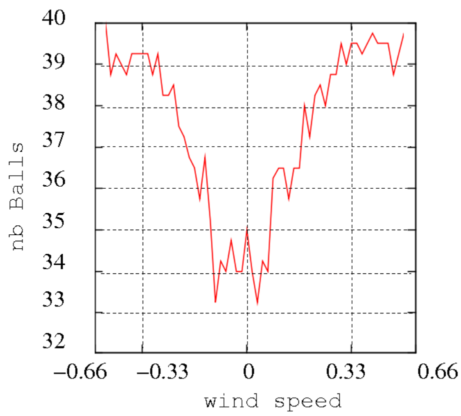

Figure 11: Average number of dropped balls according to wind.

6 Conclusion

This study was based on the assumption that the behavioral believability of a virtual entity can be increased by integrating an anticipatory ability enabling the prediction of the behavior by other entities and their impact upon the environment.

This led us to develop a conceptual framework taking into account some results from cognitive science. Its relevance was tested on the case-study of juggling: a virtual juggler anticipates the trajectory of balls without calculating them accurately. Indeed the juggler hypothesizes within an open and uncertain environment with variable properties, that is to say, that are unknown from an analytical standpoint. Universal approximators obtained through learning are used. One problem is that this type of approximator is well adapted to trajectories prediction but is certainly worst to address more complex behavior like the anticipation of human activity for instance. In such a case, it is important to address other predictive model without losing the general features of our proposition. For example, in [22], we propose an algorithm to learn a fuzzy cognitive map. Such kind of models are able to take into account behavior including decision choice and memory. Of course, using such model implies to define the link between the perception of the character and the set of symbols which can represent the behavior in the imaginary world.

Of course, this study address neither the quality of gestures, nor the comparison with real data from juggling. To do that, we have in perspective the improvement of this proposition with realistic models of gesture by integrating works like [24]. For the moment, the purpose was to show that it is possible to exhibit plausible failures in the task when taking into account simulation and anticipation.

We are currently orienting our investigations towards the addition of different juggling strategies. The imaginary world of a simulation within a simulation could be used to test many different possibilities. The results of such simulations would help to provide strategies which are better adapted to the virtual world. In addition, we would also like to work on a new kind of prediction dealing with the behavior of the human interacting with the juggler.

References

[1] M. Kipp, M. Neff, K. H. Kipp, and I. Albrecht. Towards natural gesture synthesis: Evaluating gesture units in a data-driven approach to gesture synthesis. In *Proceedings of the 7th international conference on Intelligent Virtual Agents*, pages 15–28, Berlin, Heidelberg, 2007. Springer-Verlag.

[2] H. van Welbergen, B. J. H. Basten, A. Egges, Z.M. Ruttkay, and M. H. Overmars. Real time character animation: A trade-off between naturalness and control. In Mark Pauly and Guenther Greiner, editors, *Eurographics - State-of-the-Art-Report*, pages 45–72, Munich, 2009. Eurographics Association.

[3] K. Prepin and A. Revel. Human-machine interaction as a model of machine-machine interaction: how to make machines interact as humans do. *Advanced Robotics*, 21(15), December 2007.

[4] N.I. Badler and B.L. Webber. Planning and parallel transition networks: Animation's new frontiers, 1995.

[5] F. Lamarche and S. Donikian. Automatic orchestration of behaviours through the management of resources and priority levels. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems: part 3*, AAMAS '02, pages 1309–1316, New York, NY, USA, 2002. ACM.

[6] J. E. Laird. Extending the soar cognitive architecture. In *Proceeding of the 2008 conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference*, pages 224–235. IOS Press, 2008.

[7] Justine Cassell, Catherine Pelachaud, Norman Badler, Mark Steedman, Brett Achorn, Brett Douville, Scott Prevost, and Matthew Stone. Animated conversation: Rule-based generation of facial expression, gesture and

spoken intonation for multiple conversational agents. pages 413–420, 1994.

[8] J. Cassell, T. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H. Vilhjálmsson, and H. Yan. Embodiment in conversational interfaces: Rea. In *CHI'99 Conference*, pages 520–527. ACM Press, 1999.

[9] Radoslaw Niewiadomski, Elisabetta Bevacqua, Maurizio Mancini, and Catherine Pelachaud. Greta: an interactive expressive eca system. In *International joint conference on autonomous agents and multiagent systems : AAMAS (2)*, pages 1399–1400, 2009.

[10] Stefan Kopp and Ipke Wachsmuth. Synthesizing multimodal utterances for conversational agents: Research articles. volume 15, pages 39–52, Chichester, UK, March 2004. John Wiley and Sons Ltd.

[11] Stefan Kopp. Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. *Speech Communication*, 52(6):587 – 597, 2010.

[12] A. Bryan Loyall, W. Scott Neal Reilly, Joseph Bates, and Peter Weyhrauch. System for authoring highly interactive, personality-rich interactive characters. In *SCA '04: Proceedings of the 2004 ACM SIGGRAPH/Eurographics*, pages 59–68, 2004.

[13] Herwin Welbergen van, Dennis Reidsma, and Job Zwiers. A demonstration of continuous interaction with elckerlyc. In *MOG 2010 : 3rd Workshop on Multimodal Output Generation*, volume WP 10-02, 2010.

[14] F. Multon, S. Ménardais, and B. Arnaldi. Human motion coordination: a juggler as an example. *The Visual Computer*, 17(2):91–105, 2001.

[15] H. Bergson. *Matière et mémoire*. 1896.

[16] R. Rosen. *Anticipatory systems*. Pergamon Press, 1985.

[17] A. Berthoz. *Le sens du mouvement*. Odile Jacob, Paris, 1997.

[18] C.H.M. Brunia. Neural aspects of anticipatory behavior. *Acta Psychologica*, 101:213–242, 1999.

[19] G. Hesslow. Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences*, 6(6):242–247, 2002.

[20] G. Rizzolatti, L. Fadiga, V. Gallese, and L. Fogassi. Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, 3(2):131–141, 1996.

[21] D.J. Simons and C.F. Chabris. Gorillas in our midst: sustained inattentional blindness for dynamic events. *perception*, 28:1059–1074, 1999.

[22] C. Buche, P. Chevaillier, A. Nédélec, M. Parenthoën, and J. Tisseau. Fuzzy cognitive maps for the simulation of individual adaptive behaviors. *Computer Animation and Virtual Worlds (CAVW)*, 21(6):573–587, 2010.

[23] F. Julliard and S. Gibet. Reactiva'motion project: Motion synthesis based on a reactive representation. In *GW '99: Proceedings of the International Gesture Workshop on Gesture-Based Communication in Human-Computer Interaction*, pages 265–268, London, UK, 1999. Springer-Verlag.

[24] Matthieu Aubry, Pierre De Loor, and Sylvie Gibet. Enhancing robustness to extrapolate synergies learned from motion capture. In *CASA 2010, 23rd International Conference on Computer Animation and Social Agents*, June 2010.