Towards Believable Autonomous Actors in Real-Time Applications

N. Courty*, F. Devillers†, S. Donikian‡, F. Lamarche§, D. Margery† and S. Menardais§

IRISA

Campus de Beaulieu

F-35042 Rennes, FRANCE

[ncourty,fdeville,donikian,flamarch,dmargery,smenarda]@irisa.fr

Abstract

Behavioural models offer the ability to simulate autonomous agents like organisms and living beings. Psychological studies have shown that the human behaviour can be described by a perception-decision-action loop, in which the decisional process should integrate several programming paradigms such as real-time, concurrency and hierarchy. Building such systems for interactive simulation requires the design of a reactive system handling dataflows with the environment, and involving task control and preemption. Since a complete mental model based on vision and image processing cannot be constructed in real time using purely geometrical information, higher levels of information are needed in a model of the virtual environment. For example, the autonomous actors of a virtual world would exploit the knowledge of the environment topology to navigate through it. Accordingly, we present in this paper our programming framework for real-time behavioural animation which is composed of a general animation and simulation platform, a framework for motion control, a behavioural modelling language, a perception control module and a scenario authoring tool. Those tools have been used for different applications such as pedestrian and car driver interaction in urban environments, or a virtual museum populated by a group of visitors.

1 Introduction

Behavioural animation consists of a high level closed control loop [31, 7, 5, 45], which offers the ability to simulate autonomous entities like organisms and living beings. Such actors are able to perceive their environment, to take decision, to communicate with each others [4] and to execute some actions, such as walking in the street or grasp-

ing an object. Realistic behaviours of autonomous actors evolving in complex and structured environments can be obtained if and only if the relationship between the actor and its surrounding environment can be simulated. Information that must be extracted or interpreted from the environment depends on the abstraction level of the reasoning performed by autonomous actors.

The issue addressed in our work concerns the specification of a general formalism for behaviour modelling based on psychological studies and compatible with real-time constraints. In the next section, a complete programming environment dedicated to behavioural animations is presented. Then, section 3 is devoted to the presentation of an example: a virtual museum. Finally, we present work in progress and conclude.

2 A programming environment for behavioural animation

2.1 Introduction

The first development to put together research works performed in the team started in 1994. The first architecture of a general animation and simulation platform (GASP) can be found in [13]. Then, we started to build both the kernel of this platform and a set of tools dedicated to the modelling of different aspects of an autonomous entity: geometry, dynamics, motion control, behaviour, artificial vision (see figure 1). The mechanical aspect is modelled with DREAM, our rigid and deformable bodies modelling system which generates numerical C++ simulation code for GASP [9]. For human motion, a bio-mechanical approach has been preferred [40]. HPTS concerns the modelling of the behavioural part of an actor [39] and is also used as an intermediate level for scenario authoring [14]. VUEMS is the acronym for Virtual Urban Environment Modelling System, and its main aim is to build a realistic virtual copy of urban environments in which behavioural simulations would be performed. Urban traffic is of a

---

*IRISA/France-Telecom R&D

†IRISA/INRIA

‡IRISA/CNRS

§IRISA/University of Rennes I

high degree of complexity, as it requires interactions on the same thoroughfare between not only cars, trucks, cyclists and pedestrians, but also public transportation systems such as trams. In the past years we have integrated all these transportation modes into GASP and applied this to different research projects [25, 35, 15, 34].
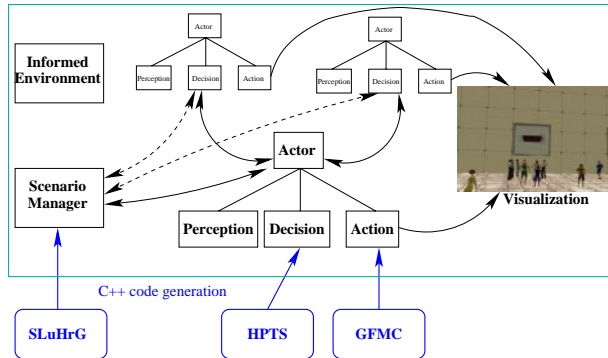

Figure 1: Architecture of an application including several autonomous characters..

In the context of national research projects, GASP has already been used with success in other research labs in France. Our objective is now to make it available to the international community as an open-source version with an API in english, including documentation and tutorials. This new version of GASP (version 3.0) is currently in beta testing and will be available at the beginning of year 2002. As GASP is a product already used by a company, we have decided to rename it as OpenMASK (Open Modular Animation and Simulation Kit). The first release of OpenMASK will be available during Imagina'02. In this section, we will first present the main characteristics of the simulation platform and then give an outline of the dedicated tools.

2.2 OpenMASK: Modular Animation and Simulation Kit

A behavioural animation is composed of a large set of dynamic entities evolving and interacting in a complex environment. To be able to model such applications, we need to implement different models: environment models, mechanical models, motion control models, behavioural models, sensor models, geometric models and scenarios.

One of the objectives of openMask (modular animation and simulation kit, and multi-threaded animation and simulation kernel) is to provide a common run-time and conception framework for the creation of virtual environments. In openMask, the building block of a virtual environment is the simulated object, which can be viewed as

the container of all the different models used for the simulation of the evolution of the environment. More formally, a simulated object can be seen as a container of a computation function, $(E_e, S, O, \theta_{n+1}) = f(E_r, I, CP, \theta_n)$, where O is a set of outputs, S a set of fired signals, E_r a set of received events, E_e a set of emitted events, I a set of inputs, CP a set of control parameters and θ_n the state of the object at simulation step n. Outputs and control parameters are object representing the public state of an object, its position for example. Inputs are connected (statically or dynamically) to the outputs and control parameters of other objects in order to establish data flow connection between objects. Signals are events sent in the environment without receivers for asynchronous communication between objects, and are received by any registered object. Therefore, objects can communicate using data-flow mechanisms and asynchronous ones at the same time, enabling implementation of different models using the most appropriate communication style for each model.

Object activation (the moment when f is called) can be done by the animation and simulation kernel either on a regular basis (giving the object a simulation frequency), only when the object receives events (if the object has no frequency), or on a regular basis but also when events are received. Therefore, the same framework can be used in a natural way for reactive objects as well as for active (living) objects.

Building an openMask application consists of composing simulated objects, by organizing them all in a simulation tree (which can be flat) and giving the configuration parameters necessary to obtain the required results. The simulation tree is used to structure relations between objects, thus creating son-father and brother relations in the simulation. These relations are for example used to establish dynamic connections between the components of a simulated entity whose behavior has been implemented using different simulated object (for example, one object would be responsible for perception, one for decision, and one for simulation of the (bio)-mechanical results of performed actions).

Once the application has been built, the last step before achieving any result is choosing the run-time kernel used. Indeed, the conceptual framework presented here is a solid foundation for many different run-time constraints, as it has been designed to enable off-line as well as online simulation, and multi-threaded execution. Indeed, the component responsible for the rendering of a simulation is a simulated object, which can therefore be included or excluded from the application. At the time of this writing, openMask has a number of different run-time kernels, for Irix and linux. These run-time kernels include a real-time kernel (only for Irix), a multi-threaded kernel (for SMP machines), a distributed kernel (using PVM) and a kernel

Figure 2: Autonomous characters visiting a museum and cooperative interaction in a virtual environment.

for PC-clusters running linux[1][33]. Discussion of the performance and tradeoffs made by each of these kernels is beyond the scope of this article, but it should be noted that the distributed and the usual kernels have been extensively used for a number of projects, ranging from behavioral animation to distributed virtual reality.

2.3 A Global Framework for Motion Control

We propose a global framework to simplify the use of motion capture for computer animation. We first focus on the process dealing with captured trajectories to calculate angular trajectories required to animate a H-ANIM compliant VRML model[2]. The motion capture files exhibit occlusions, noise and do not have any angular information. Our system [36] automatically computes the angles for each degree of freedom without noise and without missing data. It also adapts the motion to respect new skeleton morphologies and compress the final motion in order to minimize data space. The motions can be blended and synchronized together in an interactive simulation.

Our framework proposes a tool to export fully-processed and adapted motions from captured trajectories. For this purpose we provide the user with an interface that gathers all the needed functions (see figure 3). First it allows him to load a motion capture file (*c3d* files in the Vicon370 system) and a *general structure*. The user at this level can only see the markers' positions, body segments and *c3d* internal data. Once an *origin structure* is loaded, the user can run the reconstruction process in order to recover missing markers positions and avoid marker inversion. This process takes into account the morphology data of the *origin structure* in order to recover the missing points while ensuring the original skeleton links lenghts.

---

[1]Open-source versions will exclude the PC-cluster version because its development is partially funded by a private company: IWI

[2]for more information on H-ANIM please consult http://www.h-anim.org/

Recovering the missing points is performed by a 3-step process:

First step: *naïve interpolation and distance estimations.*
This step deals with a first approximation of the missing points without taking the morphology into account. In addition, it evaluates how the distances change depending on time for all couples of neighbours.

Second step: *association of uncertainty values and reconstruction ordering.*
It consists of preparing the data in order to compute the missing points. For example, it computes the reconstruction order that optimizes the interpolation quality.

Third step: *computation of the missing points.*
Depending on the number of neighbours and the distances evaluated in the first step, it recovers the positions of the missing points.

The next step is available when the *animation structure* is loaded. The final morphology is known at this step and one can adapt the motion to the synthetic figure. The user can specify constraints (like which markers positions to maintain) at this step to help the adaptation process. The resulting motion can be filtered, made cyclical and processed in order to make the main displacement follow a given direction. Before saving the motion, the user can add specific constraints such as foot contact with the ground. The saved trajectories are encoded in order to minimize the number of keyframes while tuning a precision threshold that controls the compression quality.

Once several motions are encoded, the system allows the user to replay them together while changing the position and orientation of the synthetic figure. It also ensures the synchronization of motions thanks to constraints (such as foot prints). The user can gather and blend several actions such as locomotion, vision targeting, object handing, motion replay... Each action is linked to an independent module with its own variables. In addition, for each node, the user can define priorities that can be tuned during the animation. This priority can be specified for each node and is not general to the action. Hence during locomotion one can assign a high priority to the knee and a lower priority to the hip. In this example, the high priority assumes that the knee trajectory is more important to preserve than the hip one. The actions can be interactively started and ended by the user. For example, human locomotion gives naturally high priority to the legs and low priority to the remainder of the body. While walking or running the human-like figure can also take objects that engender higher priorities to the arms. Therefore taking an object will progressively take control of the arms by

automatically increasing the corresponding action priority. A special action is always performed even if no action is specified by the user. This action consists of animating the human-like figure at rest with a set of predefined motions.

At the end of the process, the final posture is computed with the resulting motions of the active actions synchronized thanks to constraints. In the current version, our system only deals with feet sequencing constraints: for example, a single-support phase followed by a double support phase while walking or a single-support phase followed by a non-contact phase while running. As each action is linked to a captured motion the feet sequence is intrinsically defined in the motion capture file. Once the final posture is computed, the system send it to the 3d engine to animate the virtual actor.

Compared to commercial tools such as Maya and Kaydara Filmbox, our system is capable of interpolating trajectories while ensuring anatomical constraints. It also enables to adapt the motion to a different skeleton while respecting the position or orientation constraints over some particular points. Whereas Molet et al. [37, 38] used anatomical corrections for magnetic-based motion capture we propose a new method for optic-based systems that are widely used in entertainment and biomechanics. This kind of system implies new constraints such as dealing with occlusions. The main contribution of our system is to gather new methods for motion capture and adaptation in a unique tool. This tool was conceived for animators in order to improve motion design efficiency.

It is also designed for biomechanics where motion capture is one of the major measurements. This tool has been used by a biomechanics laboratory in Rennes (http://www.uhb.fr/labos/LPBEM/) in order to process very complex motions with a large number of occlusions. For example, we experimented this tool with trampoline figures (cf figure 4) that are performed in wide environments, with complex postures where parts of the body are totally hidden and that exhibit very high accelerations (up to 5 times the acceleration of gravity). The results are very promising and enable the use of optical-based motion capture for such complex motions. Without this tool it would not have been possible to analyze these motions. The tool has also been used successfully in handball throwing, running, gymnastics, and soccer.

2.4 HPTS: A Model for Behavioural Animation

2.4.1 Introduction.

Information needed to describe the behaviour of an entity, depends on the nature of this entity. No theory exists for determining either the necessary or sufficient struc-
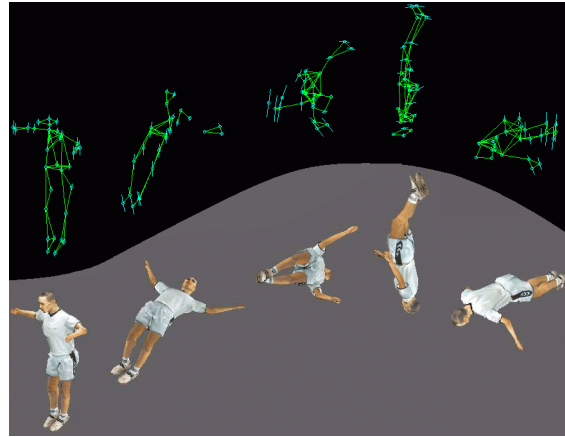

Figure 4: A trampoline motion sequence.

tures needed to support particular capabilities and certainly not to support general intelligence. As direction and inspiration towards the development of such a theory, Newell [42] posits that one way to approach sufficiency is by modelling human cognition in computational layers or bands. He suggests that these computational layers emerge from the natural hierarchy of information processing. Lord [30] introduces several paradigms about the way the brain works and controls the remainder of the human body. He explains that human behaviour is naturally hierarchical, that cognitive functions of the brain are run in parallel. Moreover cognitive functions are different in nature: some are purely reactive, while others require more time. Executions times and frequencies of the different activities are provided. Newell asserts that these levels are linked by transferring information across hierarchical levels, and that each of them operates without having any detailed knowledge of the inner workings of processes at other levels. All that is required is a transfer function to transform the information produced by one level into a form that can be used by another. Particularly important is the notion that symbolic activities occur, locally based on problem spaces constructed on a moment-to-moment basis.

Different approaches have been studied for the decision part of behavioural models in animation: sensor-effector or neural networks, behaviour rules, finite automaton approach. As human behaviour is very complex, none of the preceding models could be applied. More recently, a second generation of behavioural models has been developed to describe the human behaviour in specific tasks. The common characteristics of these new models are: reactivity, parallelism and different abstract levels of behaviours. In [6], authors describe a multi-agent development environment named DASEDIS and use it to describe the behaviour of a car driver. In [43], a tennis game application
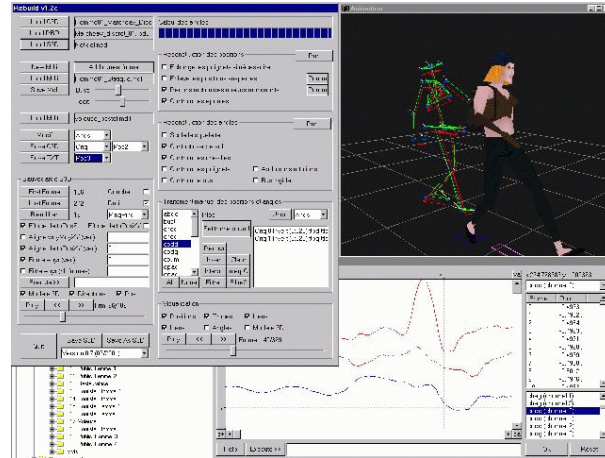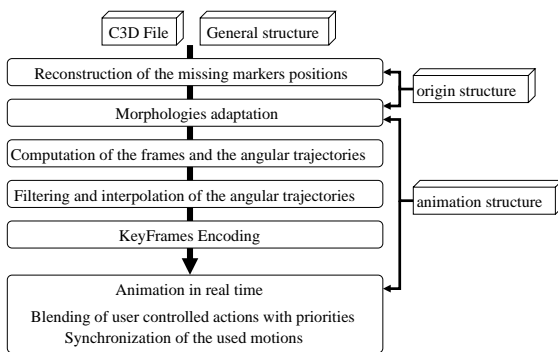
Figure 3: The framework main steps and the corresponding user interface

is shown, including the behaviour of players and referees. A stack of automata is used to describe the behaviour of each actor. In the Motivate product proposed by Motion Factory for the Game Design Market, they have also introduced Hierarchical Finite State Machines, and actions associated with the states and transitions can be described by using an object-based programming language, named Piccolo [26]. As humans are deliberative agents, purely reactive systems are not sufficient to describe their behaviour. It is necessary to integrate both cognitive and reactive aspects of behaviour. Cognitive models are rather motivated by the representation of the agent's knowledge (beliefs and intentions). Intentions enable an agent to reason about its internal state and that of others. The centre of such a deliberative agent is its own representation of the world which includes a representation of the mental state of itself and of other agents with which he is currently interacting [19]. To do this, Badler et al. [1] propose to combine Sense-Control-Action (SCA) loops with planners and PaT-Nets. SCA loops define the reflexive behaviour and are continuous systems which interconnect sensors and effectors through a network of nodes, exactly like in the sensor effector approach described above. PaT-Nets are essentially finite state automata that can be executed in parallel (for example the control of the four fingers and of the thumb for a grasping task). The planner queries the state of the database through a filtered perception to decide how to elaborate the plan and to select an action. More recently they have introduced Parameterized Action Representation (PAR) to give a description of an action, and these PARs are linked directly to PaT-Nets. It allows a user to control Autonomous Characters actions by instructions given in natural language[2]. In all these systems, the action is directly associated with each node, which doesn't allow the management of concurrency.

2.4.2 HPTS

According to Newell, our goal is to build a model which will allow some adaptative and flexible behaviour to any entity evolving in a complex environment and interacting with other entities. Interactive execution is also fundamental. This has lead us to state that paradigms required for programming a *realistic* behavioural model are: reactivity (which encompasses sporadic or asynchronous events and exceptions), modularity in the behaviour description (which allows parallelism and concurrency of sub-behaviours), data-flow (for the specification of the communication between different modules), hierarchical structuring of the behaviour (which means the possibility of preempting sub-behaviours) and time and frequency handling for execution of sub-behaviours (This provides the ability to model reaction times in perception activities). HPTS[16] which stands for Hierarchical Parallel Transition Systems, is a formalism proposed to specify the decisional part of an autonomous character. It offers a set of programming paradigms, which permit to address hierarchical concurrent behaviours. It consists of a reactive and cognitive model, which can be viewed as a multi-agent system in which agents are organized as a hierarchy of state machines. Each agent of the system can be viewed as a black-box with an In/Out data-flow, a set of control parameters and an internal state. The synchronization of the agent execution is operated by using state machines. To allow an agent to manage concurrent behaviours, sub-agents are organized inside sub-state machines. In the following, agents will be assimilated to their state machine angle, as there is not any constraint imposed on the programmer, concerning the body part of the agent. Each state machine of the system is either an atomic state machine, or a composite state machine.

Though the model may be coded directly with an im-

perative programming language like C++, we decided to build a language for the behaviour description. Figure 5 presents the syntax of the behavioural programming language which fully implements the HPTS formalism. The behavioural description language is not described in details. For a complete description of the model refers to [11], and for the management of resources and priority levels, refers to [28]. Keywords are written in bold, whereas italic typeface represents a non-terminal rule. A * stands for a $0..n$ repetition while a + stands for a $1..n$ repetition and a statement enclosed in *{ }* is optional.

```
SMACHINE Id ;
{
    PARAMS type Id {, type Id}* ; // Parameters
    VARIABLES { {type Id ;}* } // Variables
    OUT Id {, Id}* ; // Outputs
    PRIORITY = numeric expression ;
    INITIAL Id ; FINAL Id ;
    STATES // States Declaration
    {{
        Id {[Id {, Id}]} {RANDOM} {USE resource list};
        {{ /* state body */ }}
    }+}
    {TRANSITION Id {PREFERENCE Value};
    {
        ORIGIN Id ; EXTREMITY Id ;
        {DELAY float ;} {WEIGHT float ;}
        read-expr / write-expr {TIMEGATE} ;
        {{ /* transition body */ }}
    }}*
}
```

Figure 5: Syntax of the language.

The description of a state machine is done in the following way: the body of the declaration contains a list of states and a list of transitions between these states. A state is defined by its name and its activity with regard to data-flows. A state accepts an optional duration parameter which stands for the minimum and maximum amount of time spent in the state. Resources used by a state are defined by using the instruction *USE resource list*. A state machine can be parameterized; the set of parameters will be used to characterize a state machine at its creation. Variables are local to a state machine. Only variables that has been declared as outputs can be viewed by the meta state machine. A priority is attached to each state-machine and consists in a numeric expression which allow the priority to evolve during the simulation. A transition is defined by an origin, an extremity, a transition expression, two optional parameters and a transition body. The transition expression consists of two parts: a *read-expr* which includes the conditions to be fulfilled in order to fire the transition, and a *write-expr* which is a list of the generated events and basic activity primitives on the state machine. A preference value is defined for each transition.

Afterwards, C++ code for our simulation platform GASP [12] is generated. It is totally encapsulated: all transitions systems are included in their own class directly inheriting from an abstract state machine class which provides pure virtual methods for running the state machines and debugging methods. An interpreter has also been implemented, which is very useful for the behaviour specification phase as it allows to modify state-machines during the execution phase with an increase of only ten percent on the execution time.

2.4.3 Behaviour Coordination

Reproducing daily behaviours requires to be able to schedule behaviours depending on resources (body parts) and priorities (intentions or physiological parameters). A simple way is to say that behaviours which are using the same resources are mutually exclusive. This approach is not sufficient to obtain realism, as in the real life, humans are able to combine them in a much microscopic way. All day long, human being combines different behaviours, like for example reading a newspaper while drinking a coffee and smoking a cigarette. If all behaviours using common resources were mutually exclusive, an agent could not reproduce this example, except if a specific behaviour, integrating all possible combinations, is created for this purpose. This solution becomes rapidly too complex, and has motivated the recent integration of resources and priority levels into HPTS [28].

In the contrary of some previous approach, it is not necessary to specify exhaustively all behaviours that are mutually exclusive; this is done implicitly just by attaching resources to nodes, preference values to transitions and a priority function to each state machine, and by using a scheduling algorithm at run-time [29]. Each state of a state machine can use a set of resources which can be considered as semaphores. Thus, resources are used for mutual exclusion between behaviours. Entering a node implies that resources are marked as taken and exiting it implies that thoses resources are released. In order to control the execution of parallel state machines and to offer an automatic adaptation between different behaviours, it is necessary to add notions of priorities and preferences. The degree of preference is a coefficient associated to a transition and corresponds to the proclivity of the state machine to use this transition when the associated transition is true. This coefficient allows to describe a behaviour with different ways of realization; possible adaptation depending on resources availability or need can be described. A priority function is associated to each state machine. This function returns a value representing the importance of a behaviour in a given context. This function can be used to control a behaviour during the running phase. As it is user defined, it can be correlated with the internal state of the character (psychological parameters,

intentions) or with external stimuli. By combining those two notions, it is possible to create a scheduling method which globally favours the realization of most important behaviours while automatically adapting the execution of running ones. The scheduling system allows to describe independently all behaviours with their different possibilities of adaptation. During running phase, the adaptation of all other running behaviours is automatic. Moreover, consistency is ensured because the scheduler can only exploit consistent propositions of transition for each behaviour depending on the others.

In the example of fig. 6, the behaviour uses the following set of resources: Hl (left hand), Hr (right hand), rHl (reserve left hand), rHr (reserve right hand), M (mouth) and E (eyes). Resources rHl/rHr are used to handle releasing of resources Hl/Hr. The scheduler can only act on the next transition of a state machine. Hands are resources that often need more than one transition to be freed, for instance, putting down an object to free the hand resource. Then a state which only use resource Hr/Hl corresponds to a behaviour of freeing a hand resource. Note that once behaviours are described through state machines, they are controlled through their priority. This property allow to handle every type of executive behaviour without need of information about their internal structure in term of resources or possible adaptations.
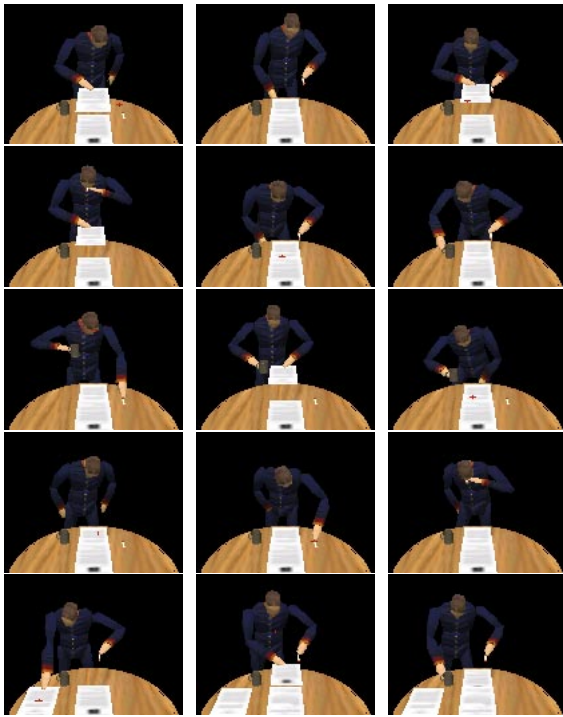

Figure 7: Behavioural Coordination Example.

2.5   Perception Control

Perception can be achieved through a direct access to the world data-base or a subset of this data-base (*e.g.*, [47, 49]) or through synthetic vision (*e.g.* [3, 22, 48, 27]). Once data related to the environment are collected, the decision module provides the actor with a set of actions to be executed. The animation of the avatar may be handled considering various techniques from full kinematic approaches to motion warping and other. These three modules have to be integrated in a more complex system that provides a task to be achieved by the actor.

We consider the case where the task given as input to the autonomous actor is specified as a visual task. Though every task cannot be specified this way, most of interesting task for a humanoïd can be taken into account: gazing at an object, following moving objects or trajectory, tracking, etc. We have proposed a complete framework that allows to directly control the humanoïd kinematic chain considering only visual informations, bringing a close link between the perception and action steps[8]. The decision step is then in charge of giving visual tasks to be achieved. This approach is based on the visual servoing framework. This method has been first introduced in the robotics field [23, 17] and considered more recently in the computer animation context [21, 32]. Along with realization of the specified task, this approach allows the introduction of constraints in, for example, the joint trajectory. It is then possible to simply handle the joint limits problem that is fundamental in the humanoïd animation problem.
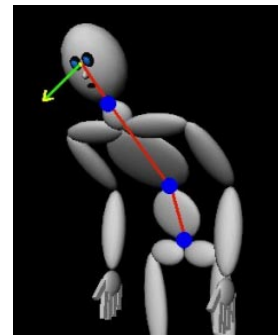

Figure 8: An example of articular chain involved in the perception process.

We choosed to link the visual information with the articular joints of the humanoïd by using a method similar to inverse kinematic. The position/orientation of the eyes are given by the visual task, and is linked with the articular joints by the inverse of the Jacobian of the chain (cf figure 8). Hence, the eyes constitute a "special" node of the inverse kinematic process. One of the main advantage of such a technique is to allow a closed loop control,
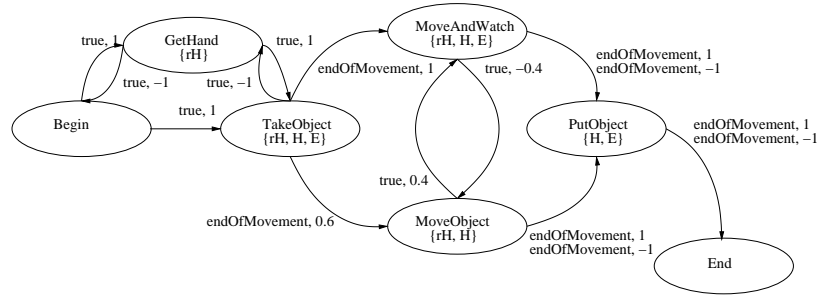
Figure 6: Moving object behaviour.

which means that a new command is computed at each frame wrt. a visual constraint, allowing reactivity of the humanoïd to modifications of the environment.
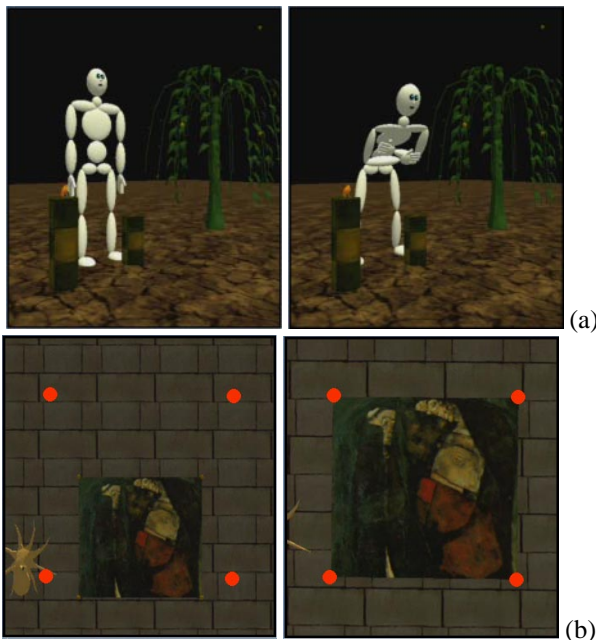

Figure 9: Servoing on four points: (a) extern view of the humanoïd (b) image "perceived"

In the following example, we consider a focusing task hard to handle with basic animation techniques. The goal of the visual task is to see a painting centered in the image. The specification of the task is done as following: four points matching with the corners of the painting are chosen. We choose their desired locations such as the painting appears centered in the image. The interaction matrix is then defined as the stacking of the four Jacobians relative to each of the four points. It is a $8 * 6$ matrix of rank six, which creates a rigid link between the humanoïd's eyes and the painting. The resulting animation is presented in figure 9 (the painting is not in the field of view of fig-

ure 9(a)). In the images rendered from the humanoïd's point of view, we drawed the four desired positions of the points (in red). At the end of the animation, they match the four corners of the painting. Thus, the specification for this task has been done in two steps: specify the four corners of the painting, and place their desired positions in the image. From a behavioral point of view, we can assume that the desired position of the four corners may be an information contained in the painting (*i.e.*, looking the painting would mean get the positions of the four corners). This approach has also been used with success to control a camera including constraints given by the cinematographic language[32].

2.6 Informed Environments

Using 3D modelling systems allow the generation of realistic geometrical models in which walk through is possible in real time. As this modelling operation is still a long, complex and costly task, a lot of work has been done to partially automate the rebuilding process. All these techniques are very useful for the visual realism of virtual urban environments but they are not sufficient due to the lack of life of these digital mock-ups. Walking through these virtual city models do not provide a real life feeling as they are uninhabited. In order to populate these virtual environments, it is necessary to specify the behaviour of autonomous characters with the ability to perceive their surrounding space and act on it. An autonomous actor whose main action is obstacle avoidance in an unstructured environment does not need other knowledge than the geometrical one. In order to simulate more sophisticated behaviours, other kind of information must be manipulated. The simplest behaviour, for a pedestrian walking in a street, consists in minimizing possible interactions, which mean avoiding static and dynamic obstacles. But, even in this simple walking activity, one needs to know the nature of objects he will interact with. For example, a public phone is considered as an obstacle to avoid for most people, but some of them will be interested by its func-

tionality and will use it. For the crossing of a street, one activity consists in reading the signals, which mean that it is necessary to associate semantic information to geometric objects in the scene, and to update it during the simulation. N. Farenc[18] has suggested using an informed environment, dedicated to urban life simulation, which is based on a hierarchical breakdown of an urban scene into environmental entities providing geometrical information as well as semantic notions. S. Raupp Musse [41] has used this informed environment to animate human crowds by using a hierarchical control: a virtual human agent belongs to a group that belongs to a crowd, and an agent applies the general behaviours defined at the group level. The knowledge on the virtual environment used by the crowd is composed of a set of obstacles (bounding box information of each obstacle to be avoided), a list of interest points (locations that the crowd should pass through and their associated regions) and a list of action points (regions where agents can perform actions).

In the realm of behavioural psychology, there have been a few studies on visual perception, mainly based on Gibson's theory of affordances [20]. The theory of affordances is based on what an object of the environment *affords* to an animal. Gibson claims that the direct perception of these affordances is possible. Affordances are relations between space, time and action, which work for the organism. What the world is to the organism depends on what the organism is doing and might do next. For computer scientists, Gibson theory is attractive because it assigns to each object some behavioural semantic, i.e. what the human being is likely to do with a given object. By associating symbolical information to objects, Widyanto experienced the Gibson's theory of "affordances" [50], while M. Kallmann [24] introduces smart objects, in which all interesting features of objects are defined during the modelling phase.

In accordance with Gibson's ecological theory, components of the virtual urban environment should be informed. To produce more realistic behaviours, we have specified a city model *affordant* to autonomous actors[46]. In the field of psycho-sociology, M. Relieu [44] has defined the notion of positive and negative affordances for a pedestrian. A positive affordance specifies the fact that the extrapolated trajectory of the corresponding pedestrian is not supposed to intersect the planned trajectory of the other one, while a negative affordance points out that they may collide. M. Relieu says also that a mobile entity uses the urban discrimination [3] to focus his attention, to select pertinent information for his actions inside the current region, while he maintains a secondary task to observe what is happening in regions close to its circula-

tion area. Thanks to this knowledge, autonomous virtual actors can behave like pedestrians or car drivers in a complex city environment. A city modeller, named VUEMS (Virtual Urban Environment Modelling System), has been designed, using this model of urban environment, and enables complex urban environments for behavioural animation, and their 3D geometric representation, to be automatically produced. The scene produced by VUEMS is loaded and is then available for use by all autonomous entities. First, sensors can determine visible objects in their environment and then the behavioural module can have access to the information on these visible objects. The behavioural model of pedestrians, that has been developped, includes social and driving rules of interaction (minimize the interaction and choose in priority the left side to overtake), as explained in [47].

2.7 Scenario Authoring

The scenario component of a behavioural simulation carries out the responsibility for orchestrating the activities of semi-autonomous agents that populate the virtual environment. In common practice, these agents are programmed as independent entities that perceive the surrounding environment and under normal circumstances behave as autonomous agents. However, in most experiments and training runs [25], we want to create a predictable experience. This is accomplished through direction of objects behaviours. To facilitate coordination of activities, objects have to be built with interfaces through which they can receive instructions to modify their behaviours. The scenario manager controls the evolution of the simulation by strategically placing objects in the environment and guiding the behaviours of objects to create desired situations. We have specified a scenario language [14, 10], which permits to describe scenarios in a hierarchical manner and to schedule them at the simulation time. This language has its own syntax and is composed of a set of specific instructions. The language contains classic instructions such as *if, switch, repeat until, random choice, wait* and which are executed inside a timestep. It contains also more specific instructions (*waitfor, eachtime, aslongas, every*) that will spend more than one timestep to be finished and that can run in parallel during the execution of the scenario they belong to. All those instructions can be composed in a hierarchical manner. To manage actors, the language offers also specific instructions to specify the interface of an actor, to reserve and free actors. By using priorities, a scenario can take an actor from another one, which will be informed by a message. Concerning messages, each scenario can subscribe to messages that are of interest for itself.

As the language is built upon C++, it is always possible

---

[3] A street is composed of connected lanes devoted to distinct mobile entities such as cars and pedestrians

to include C++ code everywhere inside scenarios by using four specific instructions: *include, declaration, implementation* and *destructor*. A scenario can be decomposed into sub-scenarios and each scenario is corresponding to a set of tasks or actions, ordered on a global temporal referential. Tasks in a scenario can modify characteristics of actors, create them or ask an actor to perform a specific action. Internal representation is based on the use of Allen's logic and of rewriting rules to produce Hierarchical Parallel Transition Systems (cf figure 10). A scenario can start at a predefined time given by the author but can also be started when a situation occurs (conditional scenario). Some of those scheduling informations are stored in a dynamic execution graph. A scheduler uses it to start or terminate scenarios. To detect situations, triggers and sensing functions are managed by the simulation observation.
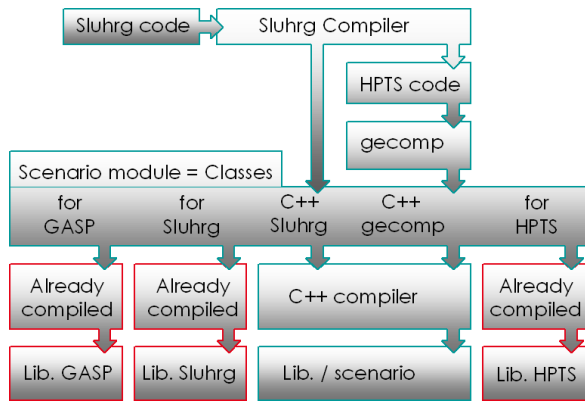

Figure 10: The global architecture of SLURGH.

3 Application to a Virtual Museum

The virtual Museum is an application which has been developed for a museum in Paris (Cité des Sciences et de l'Industrie), and has been inaugurated in June 2001 as part of the new permanent exhibition on images. This application integrates a lot of our recent research developments. It consists of a museum visited by a group of autonomous characters and it is also inhabited by a flock of suribirds (birds with the behaviour of surikats). The humanoid motion control integrates three kinds of technics: a bio-mechanical model for locomotion, inverse kinematics algorithm for grasping or climbing and motion capture for specific gestures. Different motions can be applied simultaneously to separate parts of the human body. As for Urban Environment, an informed environment has been specified to model the structure of the building and the interior architecture (cf map of the museum in figure 11).

Figure 11: The Map of the Museum.

Thanks to an initial scenario, a set of specific characters has been modelled:

- an attendant,

- a thief who want to steal paintings and sculptures,

- a photographer who damages the colors of a painting when he takes a picture of it,

- a mother and her child who have conflicting desires,

- a guide who present the exhibition to a group of visitors;

- a woman who is in charge of restoring damaged paintings.

The purpose of this application is to propose the real visitor to observe the life evolving in a virtual world and to see how the modification of behavioural parameters of one or more autonomous characters can affect this virtual life[4]. For example, the mother may have to choose between following the explanation given by the guide and looking at her child. If the child obey to her mother, there is no problem for the mother to follow the visit. In the contrary, her behaviour will depend on her interest for the visit and on the look she wants to have on the activity of her child. Concerning the attendant, it is possible to choose the level of interest (no interest, weak, medium, high, very high) that he will have to survey the thief, the photographer and the child. It is also possible to choose its speed and its weariness. Behavioural parameters can also be modified for the other main characters. It is also possible to specify a global level of disorder inside the museum: this will for example modify the response of visitors to damages caused by the thief and the photographer.

Figure 12 shows some snapshots taken from the application. On the top left image, the thief is in action, while on the top right image the guide is moving inside the museum. On the bottom left image, we can see a group of visitors listening to the explanation given by the guide (including the mother and her child). The bottom right

---

[4]This can easily be done due to the preemption mechanism, and the dynamicity and parameterization of state machines.

Figure 12: Images of the Virtual Museum Application.

image is an overview of one part of the museum, including around twenty autonomous characters. As this application should run all the day, sevend days a week, when the attendant ask the thief or the photographer to leave the museum, a women is asked to restore paintings and sculptures are put back on their pedestal. After a moment, the thief or the photographer are asked to come back into the museum by the entrance.

4 Conclusion

Our main objective is real-time simulations of several entities evolving in realistic informed environments. Many studies have been performed by psychologists to analyse the human behaviour. The behavioural model allows us to describe, in a same way, different kinds of living beings, and to simulate them in the same virtual environment, while most of behavioural models are presently re-stricted to the animation of one model in a specific environment. The use of Hierarchical Parallel Transition Systems allows us to take into account several programming paradigms important to describe *realistic* behaviours. Because of the integration of our behavioural model in a simulation platform, we have also the ability to deal with real time during the specification and the execution phases. Another important point is that our behavioural model has been built to generate dynamic entities which are both autonomous and controllable by a scenario, allowing us to use the same model in different contexts and moreover with different levels of control.

Concerning behaviour coordination, our scheduler is currently only able to handle a fixed number of resources declared at compilation time. An extension would be to allow resource declaration at runtime in order to handle external resources. Another extension will be to connect this work to a higher level of reasoning, in order to

determine automatically which behaviour should be activated or inhibited. Concerning scenario authoring, work in progress concerns scenarios in natural languages to be able to offer an authoring tool to scenarists of interactive drama. This requires to take into account spatio-temporal semantics of natural language and the theory of drama as expressed by structuralists. A first demonstration of this work, entitled *Lifetime at the grocery*, will be shown at the Industry and Innovation Village of Imagina.02.

Video sequences can be found on the following web page: http://www.irisa.fr/prive/donikian.

References

[1] N. Badler, B. Reich, and B. Webber. Towards personalities for animated agents with reactive and planning behaviors. *Lecture Notes in Artificial Intelligence, Creating Personalities for synthetic actors*, (1195):43–57, 1997.

[2] R. Bindiganavale, W. Schuler, J. Allbeck, N. Badler, A. Joshi, and M. Palmer. Dynamically altering agent behaviors using natural language instructions. In C. Sierra, M. Gini, and J. Rosenschein, editors, *International Conference on Autonomous Agents*, pages 293–300, Barcelona, Spain, June 2000. ACM Press.

[3] B. Blumberg. Go with the flow: Synthetic vision for autonomous animated creatures. In *AAAI Conference on Autonomous Agents*, 1997.

[4] B. Blumberg and T. Galyean. Multi-level direction of autonomous creatures for real-time virtual environments. In *Siggraph*, pages 47–54, Los Angeles, California, U.S.A., Aug. 1995. ACM.

[5] D. Brogan, R. Metoyer, and J. Hodgins. Dynamically simulated characters in virtual environments. *IEEE Computer Graphics and Applications*, pages 58–69, 1998.

[6] B. Burmeister, J. Doormann, and G. Matylis. Agent-oriented traffic simulation. *Transactions of the Society for Computer Simulation International*, 14(2), June 1997.

[7] E. Cerezo, A. Pina, and F. Seron. Motion and behaviour modelling: state of art and new trends. *The Visual Computer*, 15:124–146, 1999.

[8] N. Courty, E. Marchand, and B. Arnaldi. Through-the-eyes control of a virtual humanoïd. In N. M.-T. e. H.-S. Ko, editor, *IEEE Int. Conf. on Computer Animation 2001*, Seoul, Korea, 2001.

[9] R. Cozot and B. Arnaldi. A language for multibody systems modelling. In *European Simulation Symposium*, Erlangen-Nuremberg, Oct. 1995.

[10] F. Devillers. *Langage de scénario pour des acteurs semi-autonomes*. PhD thesis, Université de Rennes I, Sept. 2001.

[11] S. Donikian. HPTS: a behaviour modelling language for autonomous agents. In *Fifth International Conference on Autonomous Agents*, Montreal, Canada, May 2001. ACM Press.

[12] S. Donikian, A. Chauffaut, R. Kulpa, and T. Duval. Gasp: from modular programming to distributed execution. In *Computer Animation'98*, pages 79–87, Philadelphie, USA, June 1998. IEEE.

[13] S. Donikian and R. Cozot. General animation and simulation platform. In D. Terzopoulos and D. Thalmann, editors, *Computer Animation and Simulation'95*, pages 197–209. Springer-Verlag, 1995.

[14] S. Donikian, F. Devillers, and G. Moreau. The kernel of a scenario language for animation and simulation. In *Eurographics Workshop on Animation and Simulation*, Milano, Italia, Sept. 1999. Springer Verlag.

[15] S. Donikian, G. Moreau, and G. Thomas. Multimodal driving simulation in realistic urban environments. In S. Tzafestas and G. Schmidt, editors, *Progress in System and Robot Analysis and Control Design*, pages 321–332. Lecture Notes in Control and Information Sciences (LNCIS 243), 1999.

[16] S. Donikian and E. Rutten. Reactivity, concurrency, data-flow and hierarchical preemption for behavioural animation. In E. B. R.C. Veltkamp, editor, *Programming Paradigms in Graphics'95*, Eurographics Collection. Springer-Verlag, 1995.

[17] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE trans. on Robotics and Automation*, 8(3):313–326, June 1992.

[18] N. Farenc, R. Boulic, and D. Thalmann. An informed environment dedicated to the simulation of virtual humans in urban context. In P. Brunet and R. Scopigno, editors, *EUROGRAPHICS'99*, pages 309–318. Blackwell, Sept. 1999.

[19] J. Funge, X. Tu, and D. Terzopoulos. Cognitive modeling: Knowledge, reasoning and planning for intelligent characters. In *SIGGRAPH'99*, pages 29–38, Los Angeles, Aug. 1999.

[20] J. Gibson. *The ecological approach to visual perception*. NJ: Lawrence Erlbaum Associates, Inc, Hillsdale, 1986.

[21] M. Gleicher and A. Witkin. Through-the-lens camera control. In *Proc. of SIGGRAPH 92, in Computer Graphics Proceedings*, pages 331–340, Juillet 1992.

[22] D. H. Noser, O.Renault and N. Magnenat-Thalmann. Navigation for digital actors based on synthetic vision, memory and learning. *Computer & Graphics*, 19(1):7–19, 1995.

[23] S. Hutchinson, G. Hager, and P. Corke. A tutorial on visual servo control. *IEEE trans. on Robotics and Automation*, 12(5):651–670, October 1996.

[24] M. Kallmann and D. Thalmann. Modeling objects for interaction tasks. In *Eurographics Workshop on Animation and Simulation*, Lisbon, Portugal, Sept. 1998. Springer-Verlag.

[25] J. Kearney, P. Willemsen, S. Donikian, and F. Devillers. Scenario languages for driving simulations. In *DSC'99*, Paris, France, July 1999.

[26] Y. Koga, G. Annesley, C. Becker, M. Svihura, and D. Zhu. On intelligent digital actors. In *IMAGINA'98*, 1998.

[27] J. Kuffner and J. Latombe. fast synthetic vision, memory, and learning models for virtual humans. In I. C. Society, editor, *Computer Animation '99*, pages 118–127, Genève, Suisse, mai 1999.

[28] F. Lamarche and S. Donikian. The orchestration of behaviours using resources and priority levels. In M. Cani, N. Magnenat-Thalmann, and D. Thalmann, editors, *Computer Animation and Simulation 2001*, pages 171–182, Manchester, UK, Sept. 2001. Springer-Verlag.

[29] F. Lamarche and S. Donikian. Automatic orchestration of behaviours through the management of resources and priority levels. In W. L. Johnson and C. Castelfranchi, editors, *Autonomous Agents & Multi-Agent Systems (AAMAS 2002)*, Bologna, Italy, July 2002. ACM.

[30] R. G. Lord and P. E. Levy. Moving from cognition to action : A control theory perspective. *Applied Psychology : an international review*, 43 (3):335–398, 1994.

[31] P. Maes, T. Darrell, B. Blumberg, and A. Pentland. The alive system: Full-body interaction with autonomous agents. In *Computer Animation'95*, pages 11–18, Geneva, Switzerland, Apr. 1995. IEEE.

[32] E. Marchand and N. Courty. Image-based virtual camera motion strategies. In *Graphics Interface Conference, GI2000*, pages 69–76, Quebec, Montréal, May 2000.

[33] D. Margery. *Environnement logiciel temps-réel distribué pour la simulation sur réseau de PC*. PhD thesis, Université de Rennes 1, 2001.

[34] D. Maurel and S. Donikian. *Intelligent Vehicle Technologies*, chapter ACC Systems - Overview and examples. Arnold, 2000.

[35] D. Maurel, M. Parent, and S. Donikian. Influence of acc in stop&go mode on traffic flow. In *SAE Future Transportation Technology Conference*, Costa Mesa, California, USA, Aug. 1999.

[36] S. Menardais and F. Multon. Amélioration des trajectoires acquises par des systèmes optiques pour l'animation de personnages synthétiques. *Revue Internationale de CFAO*, pages 99–113, 2000.

[37] T. Molet, R. Boulic, and D. Thalmann. A real time anatomical converter for human motion capture. In *Eurographics Workshop on Computer Animation and Simulation*, pages 79–94, Sept. 1996.

[38] T. Molet, R. Boulic, and D. Thalmann. Human motion capture driven by orientation measurements. *Presence*, 8(2):187–203, 1999.

[39] G. Moreau and S. Donikian. From psychological and real-time interaction requirements to behavioural simulation. In *Eurographics Workshop on Computer Animation and Simulation*, Lisbon, Portugal, Sept. 1998.

[40] F. Multon, L. France, M. Cani-Gascuel, and G. Debunne. Computer animation of human walking : a survey. *Journal of Visualization and Computer Animation*, 10:39–54, 1999.

[41] S. R. Musse. *Human Crowd Modelling with Various Levels of Behaviour Control*. PhD thesis, EPFL, Lausanne, Suisse, Jan. 2000.

[42] A. Newell. *Unified Theories of Cognition*. Harvard University Press, 1990.

[43] H. Noser and D. Thalmann. Sensor based synthetic actors in a tennis game simulation. In *Computer Graphics International'97*, pages 189–198, Hasselt, Belgium, June 1997. IEEE Computer Society Press.

[44] M. Relieu and L. Quéré. *Les Risques urbains : Acteurs, systèmes de prévention*, chapter Mobilité, perception et sécurité dans les espaces publics urbains. anthropos : Collection VILLES, 1998.

[45] D. Thalmann and H. Noser. Towards autonomous, perceptive, and intelligent virtual actors. In *Artificial Intelligence Today*, volume 1600 of *Lecture Notes in Artificial Intelligence*, pages 457–472. Springer, 1999.

[46] G. Thomas and S. Donikian. Modelling virtual cities dedicated to behavioural animation. In M. Gross and F. Hopgood, editors, *EUROGRAPHICS'2000*, volume 19:3, Interlaken, Switzerland, Aug. 2000. Blackwell Publishers.

[47] G. Thomas and S. Donikian. Virtual humans animation in informed urban environments. In *Computer Animation*, pages 129–136, Philadelphia, PA, USA, May 2000. IEEE Computer Society Press.

[48] X. Tu and D. Terzopoulos. Artificial fishes: Physics, locomotion, perception, behavior. In *Computer Graphics (SIGGRAPH'94 Proceedings)*, pages 43–50, Orlando, Florida, July 1994.

[49] B. Webber and N. Badler. Animation through reactions, transition nets and plans. In *International Workshop on Human Interface Technology*, Aizu, Japan, October 1995.

[50] T. Widyanto, A. Marriott, and M. West. Applying a visual perception system to a behavioral animation system. In *Eurographics Workshop on Animation and Simulation*, pages 89–98, Vienna, Austria, 1991.