Informatik-Fachberichte 258

A. Reuter (Hrsg.)

GI – 20. Jahrestagung II

Informatik auf dem Weg zum Anwender

Stuttgart, 8.-12. Oktober 1990
Proceedings

Springer-Verlag

Berlin Heidelberg New York London
Paris Tokyo Hong Kong Barcelona

Intelligence for Animation

by Prof. Dr. Dr. Mihai Nadin

Intelligent animation evolved from the traditional art of animation obsessed with the "illusion of life" (cf. Thomas and Johnston, 1984) via computer graphics. Indeed, as some of the most successful computer graphics animators stated, computer animation resulted from the application of traditional animation techniques to initially 2D and later 3D computer-supported animation (cf. Kitching, 1973; Burtnyk and Wein, 1977; Booth and MacKay, 1982; Sturman, 1984, Lassetter, 1987). For as long as the computer was used only as a tool, but not as a new medium for thinking or as a medium with its own characteristics, the issue was that of mimicking animation techniques. Once Disneyland opened a branch in the computer world, our TV screens became full of flying logos and characters of dubious aesthetic quality. Later, the question raised in the practice of using computers for animation did not concern better animation, appropriateness (within the medium and in contradistinction to other media, film in particular) but of productivity: How can we generate faster, and if possible cheaper, computer replicas of the good old Disney times? This subject is of no concern to me.

The relevant questions of intelligent animation were actually made possible not by the commercial interest in animated videoclips but by the interest in visualization. I believe I am the first to have taught a class in visualization (1986) and this at the time prior to the first grant applications under this new buzzword (which replaced computer science and AI on the list of almost sure candidates for funding). The first question I had to answer was what is "visualization," since students could not find the word in any dictionary. Actually, I conducted a class in *visual intelligence*, concerned with the cognitive aspects of images. We came to realize through class discussions that animation is a subclass of modelling defined by the choice of a simulated world. One can model the behavior of an object or system along a timeline, or along *time* and *world* coordinates (physical space, fictional reality, design space, personal or psychological environment, etc.). We focused not on technique and mimicking Disney in software, but on the characteristics of

intelligence pertinent to expressing and understanding movement, change over time, autonomous behavior in a world populated by other moving entities subjected to change and subject to the perception of change. Although this will not constitute the subject here, we came to realize that a good modelling of chaos phenomena would probably constitute an appropriate frame for the "virtual reality" of animation: Throw any object to be animated in the modeled world of chaos, and allow for any behavior pertinent to the inherent characteristics of these objects. I still have a desire to test this thought.

But back to the subject. A simple conclusion is that automatic synthesis of visual forms (endowed or not with functionality) and intelligent motion control constitute the nucleus of any modeling system. In the meanwhile, physically based modelling confirmed this conclusion. The synthesis of shapes (corresponding to some geometry knowledge base), together with a physics knowledge base (pertinent to the action of forces, to speed and acceleration, to energy levels, entropy) in charge of the simulation of physical laws constitute

effective modeling environments. David Haumann (1988), with his Dynaflex, shared with me some of the pains of simulating mechanical properties of joint and muscle action, as well dynamic deformation. It became again obvious that unless we know more about the nature of cognition, we will use many CPU cycles to mimic (in this case in an automatic routine) what it takes to move even simple shapes. I will only mention that the entire motion of the object (in particular, a character) as specified through high level controls is *pre-computed*. The so-called visual appeal (how pleasant, how veridical, i.e., realistic, how appropriate) is controlled through the selection of physical constraints. In other words, we change the physics to facilitate the aesthetics, assuming aesthetic sensitivity is provided. But we knew this from the days when Disney invented Micky Mouse and many other characters, if indeed not earlier.

It should by now be clear that my interest is in a cognitive theory of animation that precedes the burning of CPU cycles, not in a post explanatory theory that tells us what was accomplished. We already know that each physical model (of a rigid world, of a deformable world, of an isotropic universe of properties, i.e., that are the same independent of direction, etc.) is partial. Accordingly, the intelligence necessary to support animation is one of selection of goals (what do we want to move, i.e., what kind of change over time is of interest?) and of specifications of objects (usually incomplete, but not randomly incomplete, rather, incomplete according to some controlable criteria). Intentionality (aesthetic, scientific, functional) brings into the picture the appropriate physics or other knowledge domain, the appropriate chemistry principles, or biology, or abstract mathematics, or aesthetics (of art or of designed objects). The realism of the world is apparently only physical. There is a chemical level of realism, a biological level, and a mathematical level (corresponding to the level of abstract thinking!), and evidently an aesthetic level.

Accordingly, an animation system should allow for a choice (or several choices) of the constraints, as well as for the appropriate mechanisms of control, which are qualitatively different in these different domains of constraint.

Objects from no matter what kind of world (fictional, biographical, physical, etc.) are abstracted in images; images are currently synthesized from shapes (the universe of primitives that Moholy-Nagy and Le Corbusier described). These shapes are outputted from the graphics pipeline based on transformation, clipping, and rendering. To add to geometry physical, biological, or aesthetic interpretations, one needs, before any kind of hardware, the appropriate symbolic mathematics or any other appropriate computable formalism. Inverse dynamics, optimization, and especially simulated annealing are currently used, but never together. In order to come close to showing how things change over time, we need a higher level of mathematics, which is probably closer to chaos theory or even a different mathematics (of qualities). Nobody really knows how, for instance, people estimate age when looking at someone's face, but many try to animate faces and the process of aging. There are attractors that can be defined, and based on these, the

intelligent system can deal with aging (a change over time), not only of human beings, but also of materials, shapes, interactions. In other words, animation becomes a form of knowledge, an objective that Disney animators never had, and those imitating them never realized. As a particular form of computational knowledge, animation knowledge becomes a medium for testing hypotheses, for exploring new designs, and for learning.

Considering some simple examples of "movement" (cf. figures 1-5) we will be able to understand that the subject of intelligence for animation is represented by how we know about the world, how we express goals in this world, how we can change the state of the world, and what kind of knowledge we need to plan our strategies. Movement representation and perspective are related (Fig. 1).
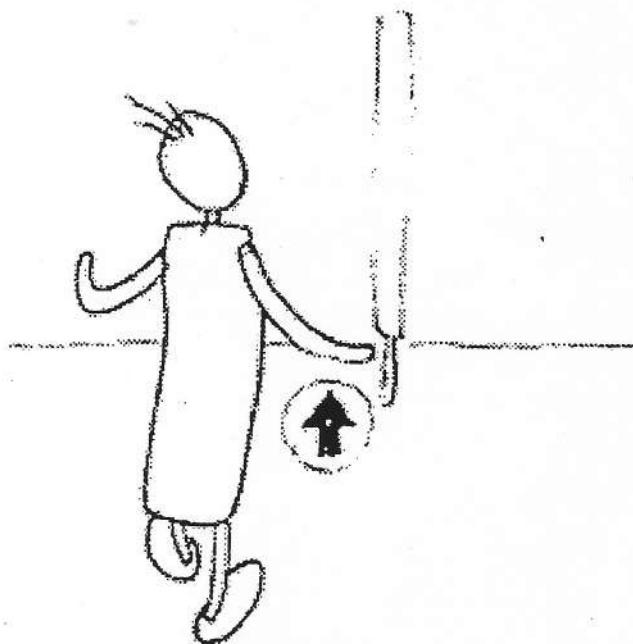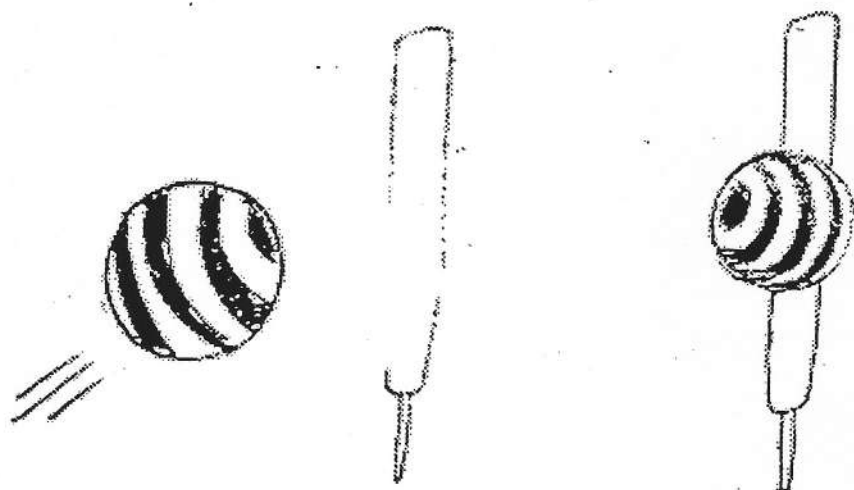
fig 1

fig 2

From a cognitive perspective, it is essential to understand how Euclidean based perspective techniques participate in our interpretation of such concepts as direction, proximity, coherence. Movement in a three dimensional space often involves not only a change of coordinates, but also of relative position (Fig. 2)

The ball moved closer to the paddle, and also rotated. The cognitive process of understanding the layers of movement, as they are visually represented, is based on perception and interpretation of visual "cues." Collision detection (and prevention) can be relatively easily automated; the constraint imposed is that two different entities should never occupy the same position in space. But the notion that a ball flying in the direction of the paddle might be "deflected" around the paddle is not inherent in the physics of the movement (cf. Fig. 3)
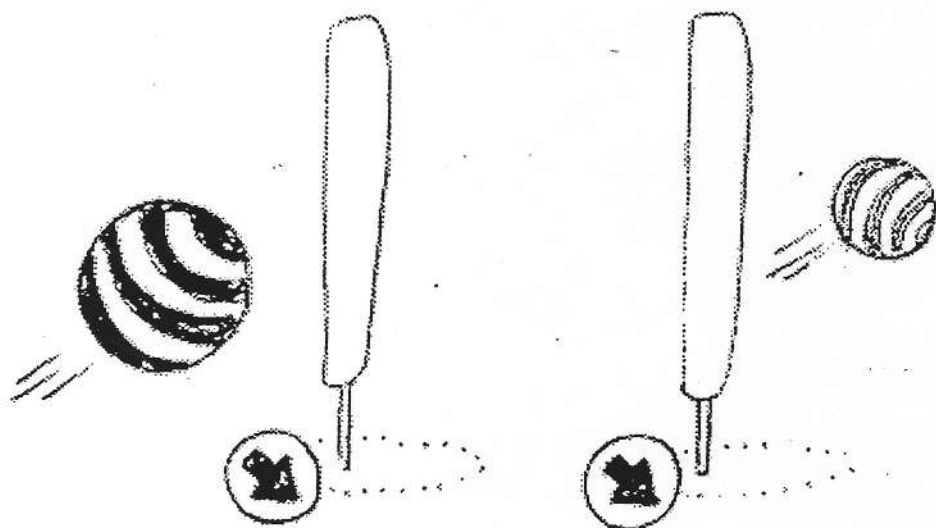
fig 3

Objects moving towards each other are in a different situation. Endowing objects with knowledge regarding the relative position of other objects is relevant only if this knowledge can be refreshed at a rate high enough in order not to jerk the movement. but this sould imply that each object "knows" in which direction and how fast the other objects are moving well in advance of the movement. Endowing objects with alternate scripts is probably a better cognitive decision (cf. Fig. 4).
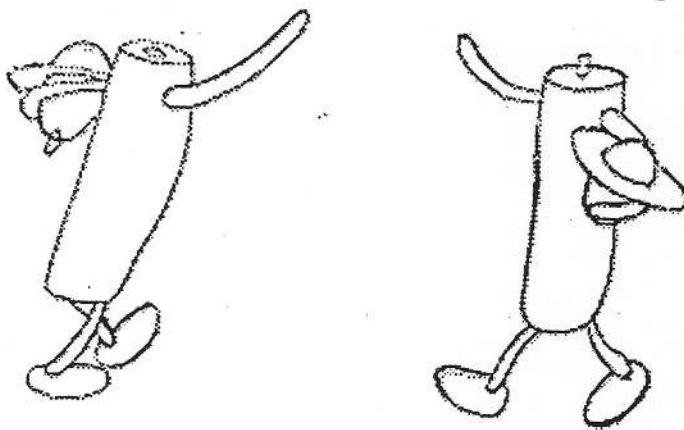
fig 4

Finally, in the hierarchy of each complex object, relations different from those that can be propagated as constraints are often expressive in animation (Fig. 5).

Despite their differences, in each of these examples, an initial state is changed into a goal state. In the jargon of domain independent planners (cf. Fikes and Nilsson, 1971), a plan is a set of ordered actions to be carried out by an agent, according to fixed preconditions (the physics of the world, or the chemistry,

biology, aesthetics, etc.). Conjunctive goals, complex control structures, and time related constraints in plan actions can also be pursued. Once the element *uncertainty* is introduced, the planner no longer searches through the space of partial plans (until it finds one that satifies the conditions in the goal state), but assigns some uncertainty value to the goal state and searches for closest matches. In other words, the planner does not complete a plan for the entire duration, but generates short plans and assumes interweaving of planning and execution. Since the response within the system to each action is uncertain, later moves depend on the new state of
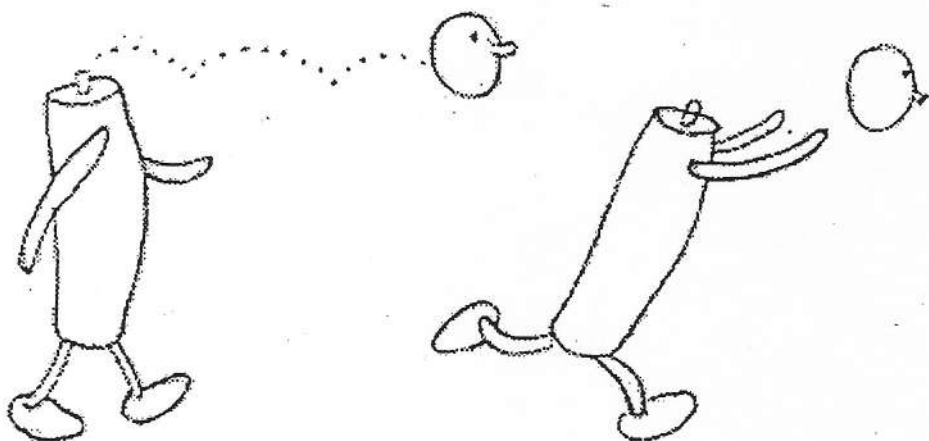
fig 5

the system. The layer of tracking (indexing) the changes and refreshing relationships is part of the control component (cf. Fig. 6).

The knowledge base contains generic plans and heuristics pertinent to the objects animated. More exactly, the heuristics specifies how the plans are to be modified based on self-evaluative controls (for which neural networks are trained). Adjustments from the user reflect not only

observations of the results, but also of the record kept by the system of options discarded in the process of animating. The knowledge representation framework supports acquisition of families of solutions (structural similarity vs. atypical situations, e.g., all objects in the physical environment are subject to gravity, i.e., they fall down; in some cases, an object can behave
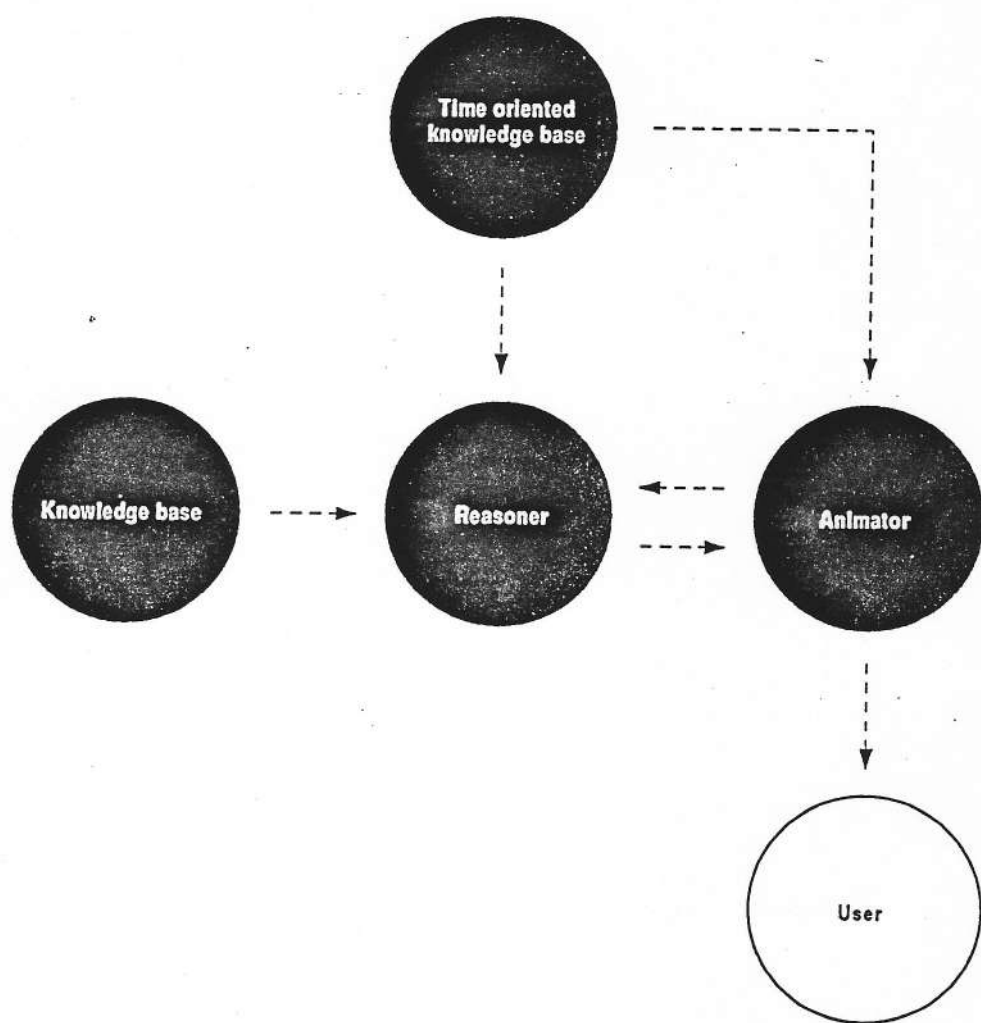
fig 6

atypically, such as in animations of objects that are not subjected to gravity). Temporal reasoning is supported by the knowledge base that keeps records of all changes over time; data retrieval from this knowledge base, together with inference and control mechanisms allow the user to access either a fine level of decision ("granularity") of the planning mechanism, or a coarse level.

The Animator accumulates and analyzes data (current and past) pertinent to the movement/change of any object or of clusters of homogenous or non-homogenous objects. It is obvious that the Animator and the Reasoner are run-time processors in charge of managing internal and external data. Whenever animation is initiated, it can be tuned either by dials, or by visual input. Nevertheless, the interface reflects the visual nature of the activity.

Once input is analyzed by the Animator, data is passed to the Reasoner, which updates the time-oriented knowledge base and builds a sequence of movements. The possible range of plan alterations

result from the knowledge base. The response of the system is mapped by the heuristic knowledge in direct relation to the hierarchic structure of the movement/change plans and the procedural specifications of the hierarchy among various possible plans. Thus a frame hierarchy is established (it can be implemented as augmented transition networks). An object-oriented language, reflecting compositional relationships (mainly aesthetic or morphologic) allow us to model structural features with frames organized in hierarchies (PART-OF and ISA). The automation of the construction of the knowledge base and of the index of protocol knowledge will require that we limit class frames. Heuristic and procedural knowledge are grouped by contexts in order to facilitate effective indexing (and thus run-time efficiency). However, it is too early to define what kind of heuristic classification is appropriate, or even if we need several different forms. The following implementation diagram is suggestive of the complexity of the enterprise.
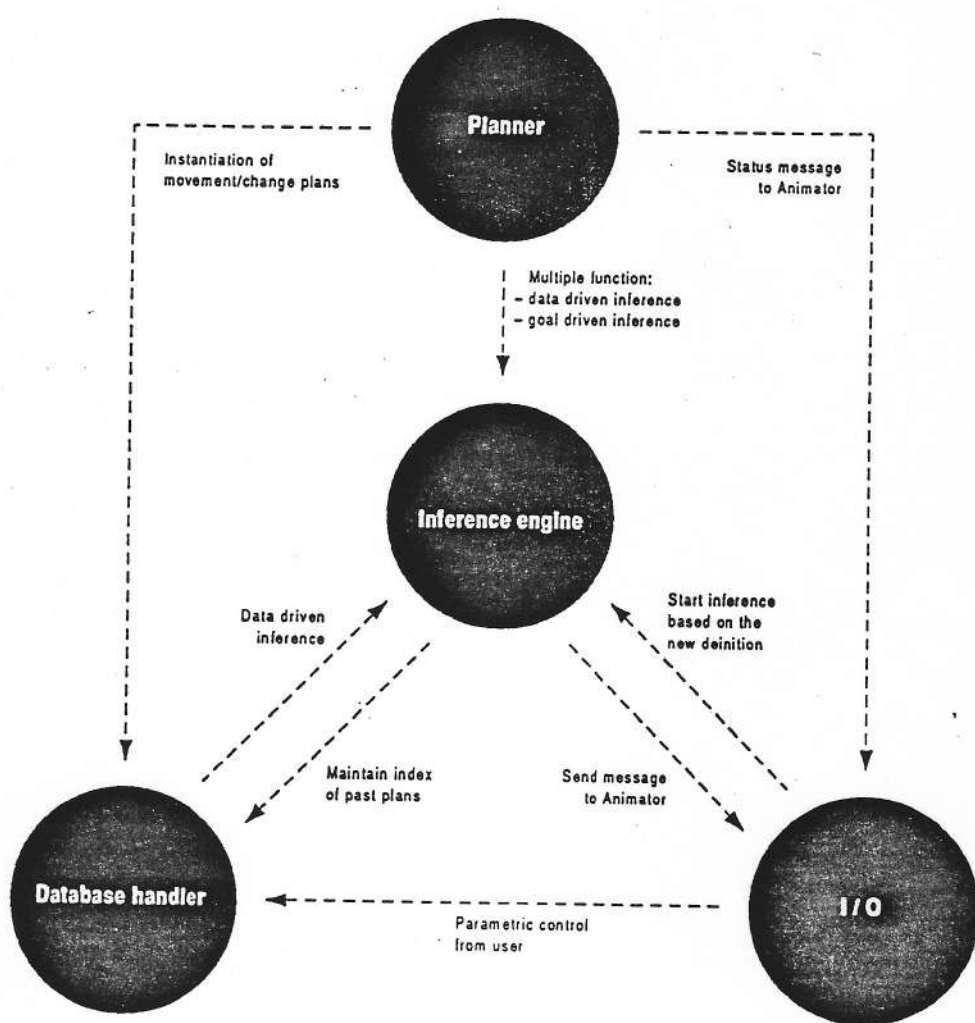
fig 7

The absence of a backtracking procedure is by no means accidental. Indeed, animation is a planning activity that requires starting from the top level again. Since uncertainty makes it difficult to derive plans from so-called first principles, the system has to explicitly maintain an index (history) of previous plans of movement or change. Real time interactive physical simulation requires motion tracking. While some accept simulations, in which conditions control the behavior, others insist that interactivity is a better way since the state of the system, together with the input values, better describe real world than video game environments. James Blinn (1989) correctly noticed that "We've seen the progression from key frame animation, specifying positions, to physically-based modelling, which is specifying accelerations and forces and whatnot." The next level should be that of using the defined cognitive characteristics of animation in a system with learning capabilities.

References

Kitching, A. Computer animation—some new ANTICS, in *Broadcast, Kinematography, Sound Television Journals*, 55(12) December 1973. pp. 54-64.

Burtnyk, N. and M. Wein. Computer Animation, in *Encyclopedia of Computer Science and Technology*, vol. 5, pp. 397-496. New York: Marcel Dekker, 1977

Lassetter, John. Principles of Traditional Animation Applied to 3D Computer Animation, in *Computer Graphics*, 21(4), 1987

Fikes, R.E. and N. Nilsson. STRIPS: A new approach to the application of theorem proving to problem solving, in *Artificial Intelligence*, 2 (3/4), 1971. pp. 189-208

Blinn, James. cf. Physically-Based Modeling: Past, Present, and Future, in *Computer Graphics*, 23(5), 1989. p. 205

Nadin, M. and M. Novak. Mind: a Design Machine—Conceptual Framework, in Intelligent CAD Systems I. (P.J.W. ten Hagen and T. Tomiyama, Eds.). New York/Berlin: Springer Verlag. 1987

Nadin, M. Interactive Diagrams for Design Genetics (Implementing Intelligent Processors), in Intelligent CAD Systems II (V. Ackman, P.J.W. ten Hagen, P.J. Veerkamp, Eds.). New York/Berlin: Springer Verlag. 1989

Nadin, Mihai. Visualization, course notes.