As published in **Communications of the ACM**, Vol 45, No 7, Jul 2001, 36-9.

The Reality of Simulated Actors

Alvy Ray Smith

March 2002

Is there something essentially human? Should actors fear for their jobs? What do actors and animators have in common? How will artists interface with their avatars? Will animators have a voice? Can a collaboration get an Oscar? I address these questions here.

Two years ago I made my first attempt at predicting when, if ever, we would be able to replace human actors with digital simulations [*Scientific American*, November 2000]. So here I also revisit and tighten those predictions.

The gist is that we must separate acting from (the appearance of) actors. My predictions then are (1) we will not replace acting, nor therefore actors, in any known way in any known timeframe, but (2) we may well replace *the appearance of* actors in my lifetime. I have in mind a parallel to the goal my colleagues and I once had, starting about 1974, to realize the first completely digital feature film. That took 20 years, far longer than originally guessed, and the result, *Toy Story* in 1994, was a cartoon. I am predicting the rest of this particular revolution, 20 years to the capability of doing a complete "live-action" motion picture, including fully realized humans. Explicitly, this will be demonstrated by the complete replacement of the appearance of a lead actor in a feature-length motion picture, including comparable amounts of screentime and spoken dialogue, and comparable numbers of closeup and medium shots. I furthermore predict that the screen representation, including voice, will be "driven" by at least one accomplished human actor, a member of the Screen Actors Guild.

Some have questioned my time estimate as too conservative, but I stick to it because the problem is harder than commonly understood, at the very minimum requiring the Moore's "law" increase in computing power of 4 orders of magnitude that 20 years promises. Using the "10× in 5" formulation—anything good about computers gets 10 times better every 5 years—we will probably need 10,000 times more computing power within a typical movie budget than currently available. Cameos of longer and longer length will appear as the techniques of human representation are mastered and Moore's law yields sufficiently cheap cycles.

Prediction (1) follows from the lack of any successful theory of consciousness. That is, there is no known way of describing how to make a machine, including ourselves, conscious. We only know that we, at least, are conscious, so it is probably possible to understand consciousness someday, but that is a statement of my personal religion and not a scientific prediction. The best current theory of consciousness is that of neuroscientist Antonio Damasio [*The Feeling of What Happens*, Harcourt Brace, 1999], but even it does not attempt to explain a fundamental aspect of the problem, the qualia problem—a tiny example of which

is the problem of how we derive "blueness" from light of certain frequencies exciting the electro-chemical system of the retinas and the brain behind them. His theory posits that consciousness and emotion are inseparable, that consciousness is in fact a feeling, and is based on his clinical observations that removal of brain parts that causes emotions to cease also causes cessation of consciousness. It is impossible to simulate acting if we cannot understand consciousness and emotion. Therefore, acting requires actors—there is no known way around them. So the remainder of this note concerns prediction (2), replacing the screen appearance of actors.

A key point is that *animators are actors*, though silent ones. When I met Frank Thomas several decades ago, one of the "grand old men" at Disney, he was acting into a mirror to inspire his animation of the character Sir Hiss in *Robin Hood* [Disney, 1973]. Today Pixar hires animators by their acting ability. Animation has *always* separated acting from the appearance of the actor. We have not tended to think of animators as actors because, until now, their screen appearance has always been a cartoon. Their screen appearance—their avatar, to borrow a term from the internet—has been an object or a comic drawing of a simple human or animal.

And *actors are animators*. A human actor can be thought of as animating his own body as his screen appearance or avatar. And doing the voice of course. The really good ones convince us that the same body, their own, is that of many different people. They seldom change gender and cannot be animals and objects. A favorite movie is *Being John Malkovich* [USA Films, 1999] which explores the possibilities available when some other actor "drives" John Malkovich's body. In one scene a woman driving his body has sex with a woman. Is that hetero- or homosexual? By the way, it is interesting to see what people choose as their internet avatars: sometimes themselves, but also animals, objects, and, surprisingly often, a human of the opposite sex.

Two major problems confront us: Animators have to be given realistic human models to animate. Actors, freed of their bodies, have to be given effective methods of driving these models, or avatars. That is, there is the problem of representing the appearance of reality in a convincing way, the "model problem," and the further problem of interfacing to such a model, the "control problem." Both will require major computational resources.

It is hard to compute a single frame of a major motion picture today. Each frame of *Toy Story* [Pixar-Disney, 1994] took an average of seven hours to compute, and each frame of *Toy Story 2* [Pixar-Disney, 1998] took about five. The best digital movies we have today—clearly and designedly cartoons—require some of the largest computations on earth, several thousand processors running around the clock for a couple of years.

My colleagues and I have long used 80 million polygons per frame as the threshold of "reality," meaning a sufficiently rich approximation that audiences cease to be concerned about its authenticity. In another 5-10 years we will see 80 megapolys as an average frame complexity. But that is only a measure of satisfactory stills.

A successful representation of a human actor must move accurately too. This is where the control problem enters. Woody in *TS* had about 100 controls for his face. Al in *TS2* had about 1000. Both are clearly cartoons. It is conceivable that a satisfactory human actor's face might require tens of thousands of controls. Presented raw to an animator/actor, this would be infeasible. So the control problem is that of presenting artists with a sufficiently rich set of controls in an intuitive way. I suspect that the control problem is quite difficult. I hope to be surprised with solutions to this problem that simply bypass the old notions of building a model and driving it directly with animation variables, and I have seen several prototype technologies that offer such shortcuts, but they have not yet been perfected.

How are traditional actors going to adapt to this new world? One obvious way is in collaboration with animators, the other kind of actor. They already do this in the voices for cartoon characters. To be clear, the actors of these characters are their animators, not the highly touted voice stars. Yet the voices—absolutely crucial to the believability of a character—do affect the presentation by the animators who are inspired by the gestures of the voice actors when creating those of the corresponding avatars. Human actors will continue to do the voices. Actor-animator collaborations will surely play an increasingly important role in the future.

Another possibility is that some animators will cross the voice boundary and come into their own as they drive, or animate, more and more realistic avatars. It has been suggested that awards be given to animator-avatar combinations as they are now given to actors driving their own bodies.

One thing is clear: Actors will not go away.