Published in Creative-Science 2010 (CS'10) Kuala Lumpur, Malaysia. 19th July 2010
http://www.creative-science.org/

# Science Fiction for Scientists!! An Introduction to SF Prototypes and Brain Machines

Brian David JOHNSON
*Intel Corporation, Portland, USA*
brian.david.johnson@intel.com

**Abstract.** This paper describes the process of Science Fiction (SF) prototyping as a design tool. An overview of the SF prototyping process is given, followed by a brisk overview of the relationship between science fictiona dn science fact as see in fiction, movies and comic books. Finally we finish up by discussion a specific SF prototype, the research and writings that fed into the process and the outcomes from the application and use. This paper should be read in conjunction with the short story *Brain Machines* (also included in this edition).

**Keywords.** ***Science Fiction Protopying, SF Prototyping, Expereince Design, human computer interaction, usage model planning, scenario based design, Quantum controllers, robotics, autonomous agents, intelligent environments, prototyping, creative science, free will.***

## Introduction - What is a SF Prototype

A Science Fiction (SF) prototype uses science fiction based explicitly on science fact as a design tool in the development of technology. Through traditional research and development we begin to define and understand *what* a technology is (Figure 1). This is the typical work that is going on in the industrial lab and universities all over the world. Usually this work continues iterating itself until the technology is refined to such a point that it can be productized to incorporated into an existing product.

 

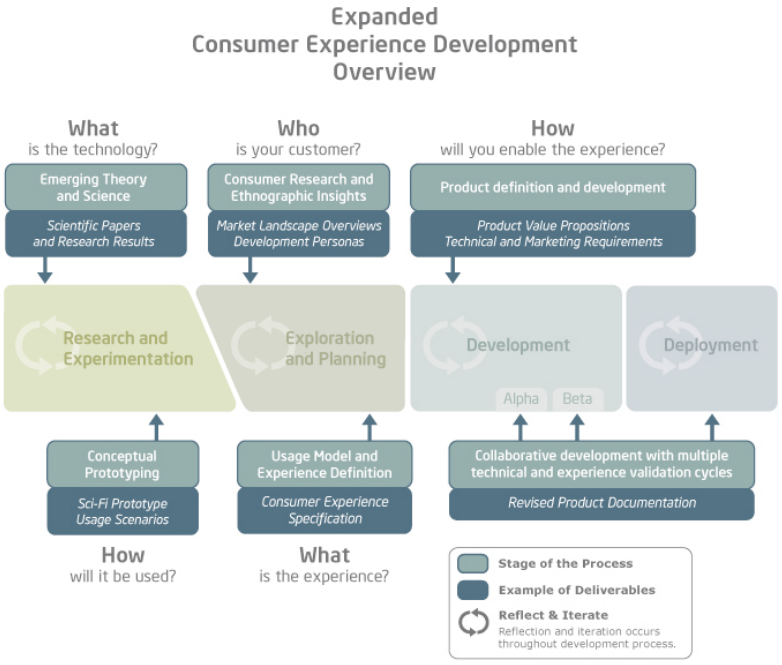


**Figure 1.** SF Prototyping process applied to a traditional product development process

The SF prototyping process adds a step this process. The SF prototype asks *how* this technology will be used. (Figure 1) It provides a virtual reality in which the implications, problems and benefits of the technology can be explored. This exploration could uncover both best case and worst case scenarios but it can also explore the subtleties of how people will use and interact with the technology.

The ultimate goal of the SF prototype is to provide a new perspective on the technology that is feeding into its development. I would also argue, as discussed later in this paper, that the SF prototype can actually benefit its feeder technology by providing specific experiments and scenarios for the application of the technology, thus illuminating possible benefits or pitfalls. By identifying these elements the development of the actual technology can be broadened and sometime accelerated.

Before we get into the specifics of the SF prototype process and the example of *Brain Machines*, let's take a quick trip through history and explore the long and symbiotic relationship between science fact and science fiction.

## 1 Science Fiction and Science Fact: A Brief History

Science fiction and science fact have a long and fruitful history together. Generations of scientists have been inspired by the awe and wonder of thrilling tales and technological adventures that fill science fiction's history. American science fiction

luminary Huge Gernsback was one of the first to overtly link the world of science fiction and the world of science fact. In the pages of his early 20th Century magazine, *The Electric Experimenter*, Gernsback places science fiction stories next to articles about science fact. Gernsback believed that, "A real electrical experimenter worthy of the name must have imagination and a vision for the future"[1] Gernsback was such a visionary author and editor that he has been honored by the Science fiction community. The Hugo Award, named after Gernsback, is one of the highest honors that can be bestowed upon a science fiction author.

Nearing the middle of the 20th Century Gernsback tradition of combining science face and science fiction continued when editor John W. Campbell encourage a then relatively unknown science fiction author names Isaac Asimov to incorporate logic and the scientific method into his stories. Until that time most science fiction stories were more romantic or fanciful, focusing on the drama or mystery of the story. Mary Shelley's Frankenstein is often sighted as the classic example of this kind of story. Shelley's gripping novel is truly pioneering in its subject matter and power but the story is more metaphysical than scientific. The novel wrestles with the themes of creation and the science of man that has lost control of its creation. However, Shelley does not concern her narrative with the details of the monster's creation. In fact, Dr. Frankenstein's lab, re-imagined countless times in movies and comic books is barely described by Shelley at all in any detail or scientific way. "In a solitary chamber, or rather cell, at the top of the house, and separated from all the other apartments by a gallery and a staircase, I kept my workshop of filthy creation…the dissecting room and the slaughterhouse furnished many of my materials"[2]

With all of this said, Shelley does not delve into the science of her story because the exploration of the science is not her end goal. It is a matter of intent. Frankenstein is a masterpiece about the responsibility of science and ultimately what it means to be human. It is not about the nuts and bolts of how one might create human life by sewing together the corpses of criminals and using lightening to give them consciousness.

Editor Campbell and write Asimov set about to create a new kind of science fiction story. As story that was based upon logic and the scientific method. From this idea came Asimov's hugely popular robot stories and the three laws of Robotic that fueled the logic behind them.

1. *Robbie* the first robot story appeared in 1939, although it was based on this idea of using an underlying logic to fuel the narrative, Asimov's Three Laws were not fully formed until 1942 and his story *Runaround*. The laws are as follows: A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey any orders given to it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.[3]

Asimov's Three Laws was a pivotal point in the combination of science fiction and science fact. For the fist time the author was using logic as the basic machinery to drive the narrative. Asimov's robot stories are about scientists and robots as they explore the implications of the Three Laws in sometimes desperate situations. The stories are not overly metaphysical; they explore in a plain and comfortable writing

style what would happen if a series of robots had to obey these three laws in the real world. In this way Asimov's stories can be seen as some of the first SF prototypes.

As we near the end of the 20th Century the link between fact and fiction is quite strong. Many authors such as Vernor Vinge (A First Upon the Deep, Rainbow's End, True Names), Greg Bear (Mars) and Cory Doctorow (Down and Out in the Magic Kingdom, Makers, Little Brother) readily point out that their fiction is not only based upon emerging science but they are in fact looking to use their fiction as a means to not only affect that science but also how that science is perceived and used in the real world.

Gregory Benford scientist and science fiction writer went one step beyond this in his book *Beyond Human-Living with Robots and Cyborgs*, co-authored with Elisabeth Malartre, saying that "Science has often followed cultural anticipation, not led it. Fiction and film have meditated upon upcoming social issues of robots and cyborgs for centuries."[4]

Benford and Malartre point out that often science fiction has allowed society and the scientific world a means to explore the cultural implications of new technologies before it is invented. This unique insight provides us a key component of the SF prototype. Can we use science fiction as a means for understanding and exploring science fiction before it is invented? Can we use science fiction as a tool for the development of science fact? The framework of the SF prototype allows us to accomplish just this goal.

The history science fiction and science fact goes far beyond the printed page. Science fiction movies rank among the highest grossing films of all time. At one science convention alone I met two roboticists who told me under their breath that the real reason they got into science in the first place was because they wanted to build the *Star Wars* droids C3P0 and R2D2.

Comic books too have a long love affair with science fiction. Science fiction themes rampant in mainstay super heroes like Superman, Spiderman and X-Men. Superman is an alien from the planet Krypton. Peter Parker has been turned into Spiderman from a radioactive spider bite. The X-Men delve into these of genetic engineering and scientific experimentation run amok. Many standalone comic collections like the *British 2000A.D*. and the American *Heavy Metal* deliver enough space travel, aliens and futuristic weaponry to satisfy any science fiction fan.

In 1956 DC Comics editor Julius Schwartz was given the task of rejuvenating the DC universe. As a science fiction fan Schwartz provided scientific explanations for the super powers of its heroes like the Flash, going so far as to explain that the Flash could run across water because he was running so fast that he never broke the surface tension. Linking science and to these thrilling adventures is just as significant as Asimov's Three Laws or Benford and Malartre cultural insights. In fact the logical structure and plotting of comic books may provide one of the most potent formats for the exploration of scientific themes that we have for SF prototyping.

The long standing relationship between science fiction and science fact is clear. The fruits of their collaboration surround us both in the technology we use and the stories and fictions we enjoy. SF Prototypes endeavor to harness the power of this relationship and use it as a development tool. Ultimately it is a matter of *intent*. Mary

Shelley's *Frankenstein intent* was to explore metaphysical and philosophical themes and present them to her readers. SF Prototypes begin with the *intent* of using emerging science and research to inspire science fictions that will explore the implications of the technology. SF prototyping is a tool for the development of science because its *intent* is to explore these areas. It is designed to provide a new perspective and means to innovation. It is in the include of the SF prototype in the scientist method and development process that we can achieve a new perspective that would then reveal new information or streamline the development process where the traditional scientific method could not. In 2009 Julian Bleecker captured the essence of this idea in his essay *Design Fiction: A short essay on design, science, fact and fiction design* when he wrote, "Productively confusing science fact and science fiction may be the only way for the science of fact to reach beyond itself and achieve more than incremental forms of innovation."[5]

## 2 Brain Machines, Free Will and Domestic Robots

*Brain Machines,* the story that follows this introduction, was the second in the series of Dr. Simon Egerton robot stories. *Brain Machines* was also a further development of the SF prototype. These conceptual tools took current scientific writings and used a fictional world to examine various implications of the theory as well as the situations they might bring about.

The scientific theories at play in *Brain Machines* come from two recent works. The first is a chapter from Michael Brooks' exceptional book *13 Things That Don't Make Sense*. Chapter eleven is entitled, *Free Will – Your decisions are not your own*. In it Brooks does a brisk work of moving though a history of free will experimentation and the latest advances in neuroscience research. Ultimately he shows that science is proving that humans really don't have free will but that "for all practical proposes, it makes sense to retain the illusion. Human consciousness, our sense of self and intention, may be nothing more than a by-product of being enormously complex machines that are our big-brained bodies, but it is a useful one, enabling us to deal with a complex environment."[6]

The second work is a paper from Italian astrophysicist Paola A. Zizzi called *I, Quantum Robot: Quantum Mind control on a Quantum Computer*. In the paper Zizzi explores using quantum metathought and metalanguage as a way to control robots or computers that could become self aware. Simply put, metathought is "the mental process of thinking about our own thought…the process of thinking about thinking." Zizzi uses metalanguage to keep a robot from attaining free will. "With opportune boundary conditions, an apparently self-aware quantum robot reaches a level of thought. In this case the robot can still be controlled by a metalanuage which prevents him to reach the level of metathought."[7] The goal of Zizzi's theory is to keep a robot from attaining free will.

The third and final work that feeds into Brain Machines is the paper *Instability and Irrationality: Destructive and Constructive Services within Intelligent Environments* by Simon Egerton, Victor Callaghan and Graham Clarke. This work explores the role of multiple personalities in an artificial intelligence (AI) and both the positive and negative affects of instability and irrationality on the system. The paper asks, "Does

chance have a role in intelligent environments? … chance and non-deterministic behaviour can play a fundamental and important role in intelligent environments…Underpinning our ideas is the view that intelligent environments may be seen as a complex system of interacting services...such complex systems can produce unexpected interactions that cause unplanned and often undesirable instabilities. However, not all instabilities are undesirable and in the second half of this paper, we present a conceptual notion that views system instability as a form of irrationality and propose a quantum control model for service agents within smart environments. We conjecture that irrational control models enable the service agents to perform better than if they were using traditional, rational, control models." [8]

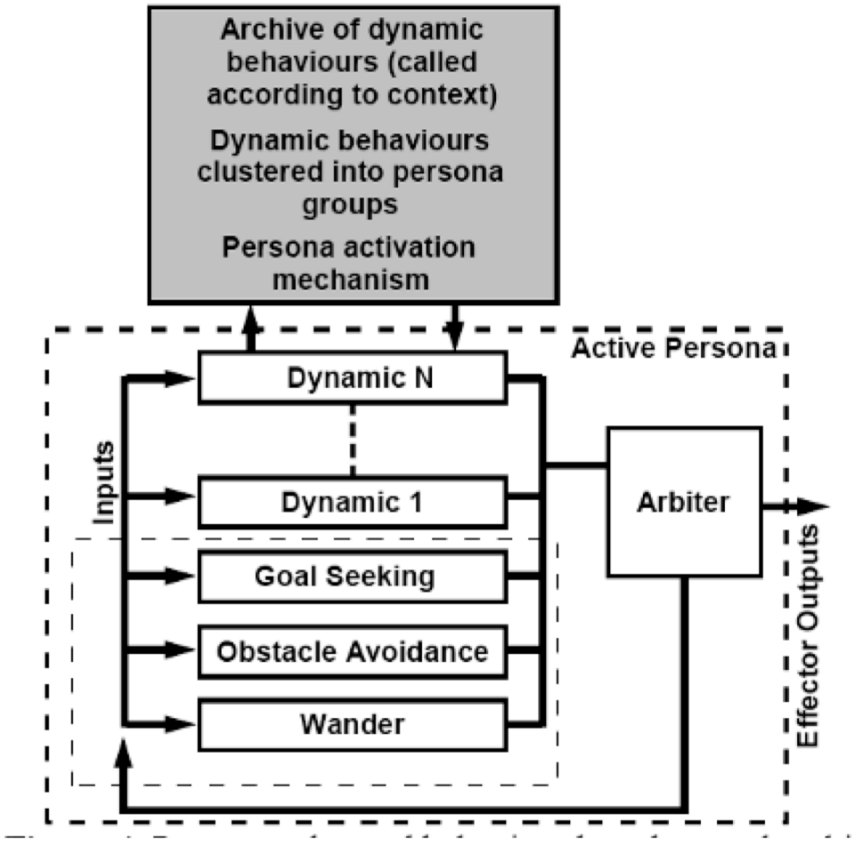


**Figure 2.** Persona enhanced behaviour based control architecture

The paper first establishes the idea that an AI can have multiple personalities. Much like the human brain that possesses multiple *personas* (e.g. student, parent, child, worker) so too can an AI break itself into these collections of behaviors and actions. The authors argue that by segmenting these personas it allows the AI to adapt and operate in complex environments (Figure 2).

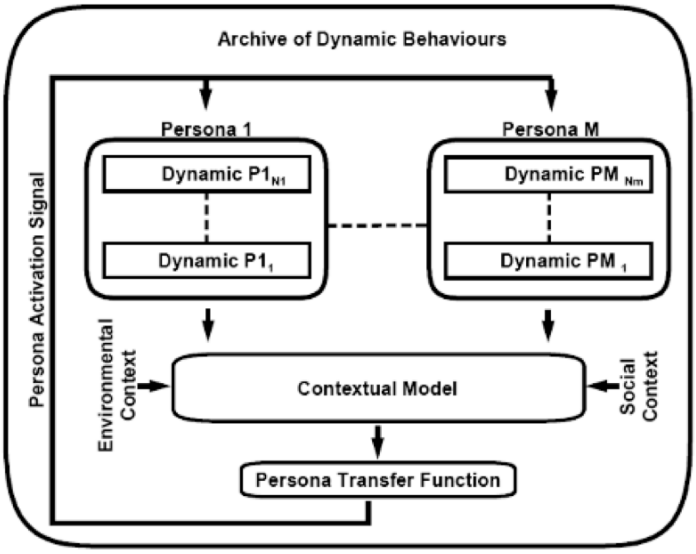


**Figure 3.** Persistent dynamic behaviours clustered into personas. An appropriate transfer function decides which of the personas is currently active based on content

The second point explored by the authors is that once we have established these multiple personas them we will need a contextual module to allow the AI to switch between these personas as needed (Figure 3). Information is fed to this contextual model from both environmental contexts (aka what is going on around the AI) and social content (aka What is acceptable for the AI to do and react to in its given environment or situation). This contextual model allows the AI to affectively switch between personas via the Persona Transfer Function depending up the subtle changes in its surrounding, thus allowing it to operate more affectively.

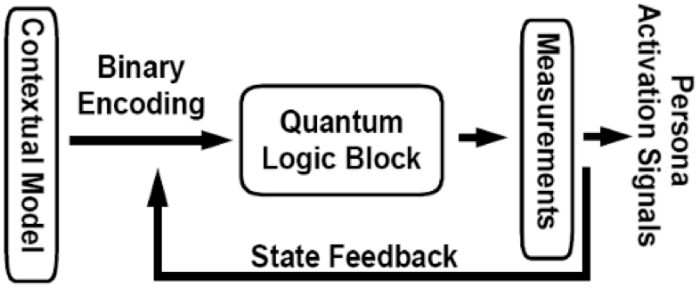


**Figure 4.** A quantum based transfer function, contextual inputs are derived from the activation levels present within each persona, the binary outputs determine which behaviour is currently active

The final bit of the paper is where it really gets good. Not that we have established that an AI can have multiple persona and that this AI has a means of switch to these various personas depending upon its environment now they introduce the element of irrationality and instability. This is achieved via the Quantum Logic Block (Figure 4). The Quantum Logic Block allows the AI to make multiple decisions at the same time, much like the human brain. These multiple decisions can product, on a small level,

seemingly irrational or unstable results. However these results are measured and fed back into the AI and the contextual model, allowing it to learn from both *good* decisions and *bad* decisions as well. This innovative approach to AI, means that the system can learn faster and adjust to its environmnet quicker than a traditional, linear AI approach. It also means that the AI could appear to have free will.

## 3 Free Will – Good for Humans…Bad for Robots

The SF prototype *Brian Machines* is all about free will and free will is a tricky thing. For humans it seems it is at the center of what makes us human. It is fundamental for us to function in our cultures, societies and governments. "Free will goes to the center of our sense of self, our autonomy as human beings. Strip us of it, and we are nothing more than animals." [9]

But when you apply the notion of free will to machines you get a dramatically different affect; fear. Free will in a robot, computer or a non-human is seen as dangerous. A quantum robot with free will is a threat to humanity. Quantum robots "might even become self-aware, conscious and have *free will*. This will be the sign that the technological singularity has been reached. Such a singularity might be very dangerous if quantum robots decide to act against human beings and take advantage of them" [10]

Obviously when it comes to free will there are two overriding assumptions. The first is that humans must possess free will or at worst they must maintain their delusion of free will to operate in complex societies and environments. "In the illusion of free will, it seems we have been equipped with a neurological slight of hand that, while contrarational, helps us deal with a complex social and physical environment."[11]

The second assumption is that machines, robots and computers must never develop free will. A machine that can think and act for itself strikes fear into the hearts of many scientists and science fiction fans alike. The big worry is that when machines get smarter than humans then they will take over our role as top dog here on Earth. I'd argue this is more a reflection of human tendencies rather than the ultimate goals of robots. Humans like being in charge and we assume that everybody must want to take it form us.

The American scientific icon Isaac Asimov had an interesting take on this back in 1977 when he was writing his non-fiction robot series for American Airlines magazine. "But if computers become more intelligent than human beings, might they not replace us? Well shouldn't they? They may be as kind as they are intelligent and just let us dwindle by attrition. They might keep some of us as pets, or on reservations. Then too, consider what we're doing to ourselves right now – to all living things and to the very planet we live on. Maybe it is *time* we were replaced. Maybe the real danger is that computers won't be developed to the point of replacing us fast enough."[12]

The Egerton robot science fiction prototypes really are experiments in extremes. They look for the worse case scenario, go right for the heart of the debate, and search for the nastiest problems that might arise in science and culture.

*Brain Machines* examines the worst legal and psychological affect of a society coming to grips with the terror of non-determinism. It also puts forward the idea that

free will in robots may have positive affects. If humans must retain our delusion that we have free will to survive in a complex environment, then why not apply the same principle to artificial intelligence? If thousands of years of human evolution have taught us anything it's that adaptability is crucial for survival. Why wouldn't a non-deterministic approach to robots and artificial intelligence not increase its chance of survival in a complex environment?

## 4 Conclusion

This paper is both an introduction to the SF prototyping process as well as an exploration of the research writings that fed into *Brain Machines*, an example SF prototype that has been included with this edition. These two pieces when read together will give you overview and specific example of the SF prototyping process.

## Acknowledgment

This work would have not been possible if it hadn't been for the exceptional research and development work of Victor Callaghan, Simon Egerton and Graham Clarke.

## References

[1] M. Ashley, *Time Machines: The Story of the Science-Fiction Pulp Magazines from the beginning to 1950*, Liverpool University Press, (2000), 30–33.
[2] M. Shelley, Frankenstein
[3] I. Asimov. Runaround. *Astounding Stories*, (1942)
[4] G. Benford and E. Malartre, *Beyond Human-Living with Robots and Cyborgs*, (2007)
[5] J. Bleecker,, *Design Fiction: A short essay on design, science, fact and fiction*, (2009)
[6] M. Brooks, *13 Things That Don't Make Sense, Vintage*, (2009)
[7] P.A. Zizzi, I, Quantum Robot: Quantum Mind Control on a Quantum Computer, *Journal of Physics* Conf. Ser., 67,. 012045, (2008)
[8] S. Egerton, V. Callaghan and G. Clarke Egerton, Callaghan, Instability and Irrationality: Destructive and Constructive Services within Intelligent Environments
[9] M. Brooks, *13 Things That Don't Make Sense*, Vintage, (2009), p160
[10] P.A. Zizzi, I, Quantum Robot: Quantum Mind Control on a Quantum Computer, *Journal of Physics* Conf. Ser., 67,. 012045, (2008), p2
[11] M. Brooks, *13 Things That Don't Make Sense*, Vintage, (2009)
[12] I. Asimov. *Our Intelligent Tools.* American Airlines, (1977)