A Brief Memory of Stephen Hawking.

Alan Yuille

Bloomberg Distinguished Professor of Cognitive Science and Computer Science at Johns Hopkins University.

Krieger School of Arts and Sciences magazine. To appear 2019.

I heard about Stephen Hawking's death at midnight in Baltimore almost as soon it was announced. At once I had a flashback to Stephen turning in circles with his electric wheelchair with a broad grin on his face.

I did my PhD with Stephen on Quantum Gravity in the late 1970's. At that time Stephen was not yet world famous, although I remember TV producers, an astronaut, and a Hollywood actress visiting. But his reputation among Physicists was exceptionally high. "He is already working at Einstein's level and if he succeeds in quantizing gravity he will go down in history as being greater than Einstein", one Professor told me. "He's cleaned up all the problems in General Relativity and now he is coming into our field of Quantum Mechanics and he will clean it up also", said another who had done Nobel prize quality work.

Up to that point, Stephen's career had been a ten year run of spectacular success. As a graduate student he had started by working on Einstein's theory of General Relativity (GR), which is arguably one of the greatest intellectual achievements of mankind. Unlike almost every other theory of physics it was constructed by Einstein with almost no guidance from experiments. GR is not only aesthetically attractive – sometimes it is called the beautiful theory – but it also seeks to capture the large scale structure of the universe itself. Indeed the beauty of GR, together with its audacity of trying to "solve" the universe, seduced me as a sixteen year old schoolboy and converted my daydreams from becoming a professional sports player to becoming a second Einstein (leading to dramatic improvements in my academic work to the surprise of my teachers). But until the early 1960's the conceptual and mathematical complexity of GR meant that its predictions were not clearly understood. It involved mind-boggling concepts like the warping of space and time and even Einstein made conceptual mistakes about its implications including doubting that gravitational waves (recently discovered) were true predictions of the theory. In mathematical terms the theory was highly non-linear, unlike most of the classical laws of physics, making it very hard to solve and make predictions. Indeed until the early 1960's, when Stephen's research started, it could only be solved in a very limited number of highly restricted settings and it was unclear what the theory predicted in more realistic and unrestricted situations.

Stephen's research on GR helped change all that. His advisor Dennis Sciama introduced him to new ways of thinking about GR which had recently been developed by the mathematician Roger Penrose (whose interest in GR had been inspired and encouraged by Sciama). This led to the, so called, golden age of GR where Stephen, Roger Penrose and a few collaborators found way to make predictions from GR in unrestricted situations. These results were astounding. They showed that black holes were extremely likely to occur, predicted most of their properties, and calculated what would happen when black holes collided. But with typical audacity, in his PhD thesis Stephen turned these results on their head and showed that GR predicted that the universe had to start with a singularity, or big bang. Only a few physicists in the world at that time had the physical intuition and the mathematical skills to make and understand these predictions. Stephen's work in this area alone would have earned him a Nobel prize if experiments had been able to confirm his predictions. Indeed the recent construction of gravitational wave detectors (co-led by Stephen's friend Kip Thorne) raises the hope that experimental verification will come soon.

But then Stephen trumped this work in 1975 by his incredible prediction of black hole radiation, now known as Hawking radiation. The story goes that when Stephen presented his result at a conference one of the organizers said it was complete nonsense and closed the session. The organizer's reaction was not completely unreasonable since Stephen's result combined three completely separate and apparently unrelated areas of physics: General Relativity, Quantum Mechanics (QM), and Thermodynamics. This suggests an, as yet unrealized, deep connection between these fundamental theories. There is still no experimental verification of this result, and it is unclear how and when it will come, but the circumstantial evidence for it is so strong that most physicists think it must be true. The formula for Hawking radiation will appear on Stephen's tombstone.

When I met Stephen two years later he had left GR behind, finished analyzing Hawking radiation, and was setting his sights on developing a theory which unified Quantum Mechanics with General Relativity. This was one of the last remaining big challenges in fundamental physics and a major step on the road to the holy grail of theory of everything which would unify all of physics. Quantizing a theory meant taking a "classical" theory like GR (one that worked at the scales of everyday life and larger) and making it consistent with the strange properties of physics that occur at very small scales Quantum Mechanics had been developed in the 1920's and had perhaps even more conceptual complexity than GR, such as the uncertainty principle which states that it is impossible to know the position and speed of a particle at the same time. By this time, three of the four fundamental forces (i.e. theories) of physics had been successfully quantized, with GR being the only exception. But quantizing GR was very challenging. Firstly, quantizing any theory was difficult and the first success, Quantum Electrodynamics (QED) which quantized Electromagnetism (EM), took over twenty years starting with work by Dirac (who was Sciama's PhD advisor and hence Stephen's academic grandfather). The difficulties were both conceptual and mathematically technical. It was hard to provide a quantized theory that gave well behaved mathematical predictions, so almost all attempts could be rejected on mathematical grounds without even bothering to check that they agreed with experiment. Secondly, QED required quantizing a linear theory (EM) which was much easier than quantizing a non-linear theory like GR. Physicists had learnt how to quantize some nonlinear theories, the so-called gauge theories, but this had also taken almost twenty years and these methods couldn't be applied to GR. Thirdly, quantum effects are typically very weak except at tiny scales and classical GR is already a weak force (compared to EM). So perhaps quantizing gravity would result, after a lot of hard work, in very small corrections to GR which would be almost impossible to observe experimentally?

The challenges of quantum gravity made me pause when I was offered the chance of doing a PhD supervised by Stephen. But his reputation – when could I ever work with somebody like him? – and his audacity and cheerfulness persuaded me. I first saw his famous grin when I asked him if there was any obvious way to test Quantum Gravity and he answered "none!" He was developing an approach, called the Euclidean approach to Quantum Gravity, which was based on an intuitive reformulation of quantum mechanics developed by the flamboyant physicist Richard Feynman (also building on ideas from Dirac, who had been Sciama's PhD advisor). Moreover, Stephen conjectured that quantum gravity effects might be strong because tiny black holes, caused by quantum effects, might unleash large amounts of Hawking radiation (small black holes convert matter to energy far more efficiently than nuclear fusion and so could result in the ultimate weapon).

Conversely, to truly understand Hawking radiation, and how it might lead to an enormous explosion, as the black hole radiation increased as its size became smaller, you also needed to understand Quantum Gravity.

Stephen's research group was small and rather like being in an extended family. We hung around him during most of the day and wife Jane would visit sometimes with children Robert and Lucy (Timothy was born during that time). A postdoc Don Page and later a student Nick Warner lived in Stephen's house and accompanied him to work and back to home later. We took turns to help feed him and perform other duties. I still recall with shame that I once forget to tell Stephen that I was going to London on a day when I supposed to be looking after him. We occasionally went to films with him like "Some Like it Hot" and "The Black Hole". I recall helping him to play dungeons and dragons one afternoon when he was bored. I baby sat his children once and he also held New Year's Eve parties. But despite his friendliness and informality I could never bring myself to call him Stephen to his face, perhaps because I'd gone to traditional English schools where first names were only used for very close friends.

"Stephen believes the best work is done in the tea room" I was told when I started my PhD. At morning and afternoon tea his group would sit together in the DAMPT tea room and have informal discussions including physics and almost everything else. At lunch we'd all go a block away to the Graduate Center (known as the Grad Pad) or occasionally to a nearby pub. Free ranging discussions would continue over lunch and I recall learning about Chomsky's theory of linguistics which, with typical physics arrogance, was described as being "basically trivial" (though not by Stephen).

At that time Stephen could talk but with increasing difficulty. I struggled to understand him, but never really mastered it. During my first year he stopped giving his own talks because it became clear that few people in the audience could understand his pronunciation (they were the ones who laughed at his jokes). So he prepared his talks and sat on the podium as a student presented them. Being painfully shy at the time, I was glad that I was never asked to do this. Giving my own first talk in front of Stephen was challenging enough though he gave me helpful advice beforehand and congratulated me afterwards. Indeed I remember him as being very supportive and encouraging despite my frequent periods of lack of progress and his sense of humor was almost always present. In particular, I remember him joking that maybe attempting to quantize gravity was the first sign that a physicist was going crazy (this is based on an English schoolboy joke, the second sign of madness would be thinking you had succeeded). His disability was almost never discussed and he seemed determined never to let it stop him from doing anything. He was so cheerful and upbeat that, whenever he was present, the thought of feeling sorry for him never crossed my mind.

Eddie Redmayne captured many aspects of him perfectly in the "Theory of Everything". When I saw the first scene with Redmayne in the wheelchair I couldn't have distinguished him from Stephen. The poses, the muscle movements, the personality, and the grin all seemed perfect. But for me, the film felt strange because the other people looked nothing like their real life counterparts – what was Stephen doing surrounded by these fakes? Moreover, although most of their personalities seemed accurate a few, like Dennis Sciama, were very different from how I remembered them.

One of the advantages of this family atmosphere was that it was natural to hang around and listen when Stephen talked with his many distinguished visitors. These included physics "greats" like Dirac, Gell-Mann, and Schwinger (who had succeeded in quantizing EM with Feynman and Tomonaga) and world class mathematicians like Atiyah, Singer, Milnor, Penrose, and a young Shin-Tung Yau. Unusually in those days we also had visitors from China and some of the leading Russia scientists. Unfortunately Stephen's closest counterpart in Russia, Yakov Zeldovich, was not allowed to visit western countries but he and Stephen were friendly at a conference in East Germany which was then behind the Iron Curtain. I particularly recall Dirac because he was the most senior and gave a direct link to the invention of quantum mechanics and to Einstein. Stephen and Dirac seemed friendly (and were academically related via Sciama) but Dirac spoke very little. I heard that Sciama had once excitedly gone to his then advisor Dirac saying he had a new theory of the universe and asking if he could explain it to him, but Dirac replied with the single word "no". I was luckier because I got two words from Dirac "thank you" when I picked up food at a cafeteria and stepped aside to make way for him.

I rarely saw Stephen after finishing my PhD because I went directly to the US to work on Quantum Gravity as a postdoc and shortly afterwards left physics and switched to artificial intelligence. During my PhD I'd had long periods of self-doubt. My PhD thesis was reasonable, though not stellar, but its contributions were mainly mathematical and didn't seem to make much progress towards quantizing gravity. Stephen's Euclidean approach relied on a mathematic trick, called analytic continuation, and it was unclear to me that it could be justified ("you can't do that" argued a US quantum gravity expect, "I just did", replied Stephen). I wasn't sure whether my doubts arose because of the difficulty of the problem or were due to the limits of my intelligence. Perhaps I wasn't smart enough to quantize gravity? During my first year in the US I learnt about alternative attempts to quantize gravity from the US experts, but I wasn't convinced by any of them. I started speculating that perhaps nobody was smart enough to quantize gravity and, even if somebody succeeded, how would we know if no experiments could be done? A few physics theories had been developed guided by mathematical beauty and consistency (first articulated by Dirac) and Stephen's Hawking radiation was a prime example, but these remained exceptions and even Dirac's own work on Quantum Mechanics had considerable experimental guidance.

These doubts led me to switch subject and move to MIT to work on Artificial Intelligence (AI) and theoretical neuroscience. I preferred to work in a new area which was just starting rather than in an area like physics which seemed overworked. In a new field like AI there were many wide open problems waiting to be explored. By contrast those problems remaining in physics seemed fiendishly difficult. AI was also attractive to me because understanding the brain is surely as important as understanding the universe. Moreover, like physics it was not purely ivory tower, but instead had the potential to make a huge impact on the world (much current technology is based on 19^{th} and 20^{th} century physics). Finally, AI also involved mathematics while I liked to do.

Stephen has long been interested in AI and in his last years he has spoken out about its potential dangers saying it could "either the best, or the worst thing, ever to happen to humanity". In 2015 he was a co-signer of the "open letter on artificial intelligence" which called for research on the societal impacts of AI. As an active researcher in AI, I think the short-term risks of AI are often over-exaggerated but I agree with Stuart Russell, an AI expert who also signed the letter, that "it is not too early to start thinking about these things". On the positive side I see the huge benefits AI can bring to society, like its potential for early diagnosis and treatment of diseases like cancer, its ability to help the visually disabled and the elderly, and its use for autonomous space exploration. I agree that it poses risks for employment in some industries though my one-time squash opponent Uday Karmarkar, who has studied the history of industrialization and technology on jobs and wages, says that historically employment patterns change enormously over time and new technology typically leads to new jobs being created. But this time it may be different and, in the long term, we may need to take seriously drastic changes like moving to a "leisure economy" where people only need to work 15-hour weeks, as speculated by the economist Keynes in 1930. Another major risk is that, in Stephen's words "AI might bring "new ways for the few to oppress the many" if its achievements and benefits are not shared across society. In the worst case AI will enable robots armies, surveillance techniques and psychological warfare (e.g., by exploiting online information) which might realize the horrors of Orwell's dystopian novel 1984. I'm much less worried, at least in the short term, about Stephen's concern that "someone will design AI that improves and replicates itself, (which) will be a new form of life that outperforms humans". The gap between human intelligence and AI is truly enormous. AI systems remain fragile "idiot savants" which work brilliantly for a limited number of tasks in a small number of highly restricted domains but lack the flexibility, adaptability, of general purpose of human intelligence. The human brain is almost unimaginably complex and we are a long way from solving it despite huge achievements in neuroscience. It will require several breakthroughs, some of which may be harder than quantizing gravity or developing a physics theory of everything, in order to truly understand human intelligence and design AI systems with human-like capabilities. I note that Stephen's willingness to publically address the risks of AI reflects his broad interests and concerns about social issues, similar to the way he recently publically argued that the National Health Service in Britain is being seriously weakened by shortage of government funding.

Sadly I never had the chance to discuss these issues with Stephen though I doubt that my opinions would have influenced him in the slightest. When I switched to AI, in the early 80's, he was interested in hearing about it on the few occasions that we met. In those days, of course, AI was very primitive and not advanced enough to threaten anything. But perhaps Stephen was already worrying about its potential. In 1979 he was appointed to be the Lucasian Professor of Mathematics at Cambridge (his distinguished predecessors included Isaac Newton and Dirac) and gave his famous inaugural lecture where he speculated that Physics might be over in thirty years. He ended it by joking that even if Physics wasn't over, then maybe physicists would be because they all be replaced by computers! His worries here have so far been premature and arguably AI has enabled physicists to do better research.

My meeting with him after the early 80's were infrequent. The most memorable was when he received an Honorary Fellowship at Harvard University. I was a junior professor at Harvard at the time ("have you got a backup plan" Stephen joked, knowing of Harvard's reputation for treating its junior professors). My friend Jim Clark suggested that we nominate Stephen for an Honorary Fellowship at Harvard. So we drafted a letter and asked leading professors at Harvard to sign it (all were delighted). We heard nothing for months until a friend of mine, Consuelo Correa, phoned saying that Stephen was getting a Harvard Fellowship live on TV! Jim and I got a note to Stephen, dodging Helmut Kohl's bodyguards (another new honorary Fellow), and Stephen kindly invited us to join him at the Fellowship lunch where we heard Ella Fitzgerald sing (who was also receiving a Fellowship). Another occasion, many years later, was at Caltech where his rock star status meant that his talk in a huge auditorium was sold out. Luckily a friend of mine was able to exploit her press credentials and smuggled me in as a member of the press. This was the last time I saw Stephen. His condition had deteriorated and he didn't flash his usual grin of recognition, but made warm eye contact instead. He started his talk by mentioning his first junior school, a little known place called Byron House which has since closed, and which by a strange coincidence was also my first school. I'd forgotten that at one tea time discussion at Cambridge we'd discovered that three of Stephen's graduate students (myself, Chris Pope, and Nick Warner) had all been at schools attended by Stephen. As Stephen remarked, this coincidence was even stranger because none of the schools were well-known or well-regarded. Even more curiously, James Lighthill, who held the Lucasian professorship in between Dirac and Stephen and in 1973 wrote a famous report critical of AI, also went to Byron House.

After leaving physics for AI I followed physics at a distance as a type of spectator sport. In the first few years. I sometimes worried that I'd made a bad decision. Maybe there would be an enormous breakthrough in Quantum Gravity, I'd miss all the fun, and my friends would be celebrated in history? Perhaps Euclidean Quantum Gravity would succeed or, if not, perhaps alternatives like String Theory and M-theory would turn out to be correct? I promised myself that if they did I'd take a few months holiday and try to learn them. But so far I have not been persuaded to try. From a distance it seems that the search for Quantum Gravity and a theory of everything is much harder than anticipated when the quest began. Lack of experimental data is a major hindrance. The most recent experimental findings on gravitational waves and the Higgs boson confirm theories that existed long before I started my PhD. Another major finding, the fact that the universe is expanding much faster than previously believed, was not predicted by any of the theories. There has been enormous progress in understanding the technical and mathematical properties of the physical theories, but I'm too far removed to be able to say anything knowledgeable about this. I recall a lunch at Berkeley where a mathematician friend strongly defended String Theory against the criticisms of a geneticist, but his defense was on the grounds of its mathematical progress and not as a scientific theory that had been verified.

For myself, I'm concerned that popular accounts of this area do not seem to distinguish between results which are known to be correct (i.e. carefully verified by experiment like GR and QED), results which are speculative but have strong circumstantial evidence (e.g., Hawking radiation), and other that are much more speculative. In this last category, I'd include the multiverse (except in a trivial sense), higher-order dimensions, inflation, and M-theory. I'm not criticizing physicists from working on these topics because they are much more qualified than me to decide which future directions of physics are best to explore. But, as a reader of news stories about physics, I think it is unfortunate that speculative work is often presented as if it is on as firm footing as established physics theories like GR and QED. I sometimes find myself directing friends to Peter Woit's website "Not Even Wrong" to get a healthy dose of skepticism about the state of fundamental physics.

My following of physics has involved watching Stephen's expanding role as a popularizer of science and as a scientific celebrity. His scientific writing was always extremely clear because of his ability to emphasize the key points and use the best, and simplest, words to describe ideas. This may be partly due to his disability. As a student I was privately frustrated when I occasionally helped Stephen with dictation because progress was very slow with every word taking an effort. But the result was exceptional clarity. This clarity of thought has translated brilliantly to his popular scientific writings and television presentations. It is hard for me to judge, since I know the materials the books are based on, but many people have told me how much they liked his books and how much excitement about physics they convey. I still worry, in his later books, about the lack of distinction between how much of the physics is known for certain and how much is more speculative. But this concern can be raised about much of the recent popular literature on physics and reflects to a large extent the current state of fundamental physics.

I think of fundamental science as a heroic endeavor somewhat similar to exploration. Historically most explorations ended in failure and those few which succeeded took many years and much hardship (read Antonio Pigafetta's diary about the first voyage to circumnavigate the world, where only 18 people survived out of more than 570). Scientific progress, and understanding of the universe, requires that we keep exploring and honor those who make it their life's work. It is too early to tell whether Stephen's work on the wavefunction of the universe and the no-boundary conditions of the early universe will be validated as his early work on GR and Hawking radiation almost certainly will be. The lack of experimental data, because the cost of doing them exceeds our financial resources and technological skills, requires that physicists develop them driven by criteria like mathematical beauty as Dirac advocated. This gives uncertain and possibly misleading guidance but it has also succeeded for physicists like Einstein, Dirac, and also Stephen. Moreover, lacking experimental guidance, it is hard to see what other strategies to explore.

Stephen himself, of course, can be considered to be one of the most heroic people of all time. The cover picture of him on the Brief History of Time, as a disabled man in a wheelchair against a background of stars, captures his audacity and willingness, despite enormous physical handicaps, to address the most existential questions of all. Namely how can a breed of monkey on a minor planet in an unfashionable part of an undistinguished galaxy understand the fundamental laws of the universe?