Strategic Thinking for Leadership in Science and Technology

Dr. Paul J. Werbos*

1.Introduction

The benefits of R&D both to society and to our basic understanding vary by orders of magnitude from project to project and from program to program, depending on the level of strategic thinking, vision, follow-through and mental energy which we put into them. R&D opportunities exist which, with a high probability, could solve all of the major life-or-death challenges facing the human species (except perhaps the challenge of population growth), but at present levels of R&D effectiveness, it seems more likely that the opportunities will not be captured, and that the consequences will be severe at all levels of life. Furthermore, there are serious risks of R&D opening up the wrong kinds of possibilities which could hurt more than they help. This is why I have accepted the challenge from William Bainbridge, the editor of this book, to try to summarize all I have learned about effective strategic thinking for science and technology – even though this cannot be done in a simple linear scholarly way, and even though no one on earth can claim to have a complete answer to this challenge. The inherent complexity of leadership in science and technology is second only to the inherent complexity of leading entire nations; it requires deep appreciation of many contradictory approaches, which must be weaved together in a complex way, and which no one on earth has fully integrated in logical order.

To try to bring some order to this chapter, I will begin by discussing fundamental concepts of strategic thinking and optimal policy, ranging from old but important basics through to recent developments involving intelligent systems. Next I will discuss my own view of the most important and fundamental challenges before us, and where we stand in addressing them.

2. Foundations and Methods

2.1. The Basics: from Von Neumann to Pareto to Hitch to Climate Policy 101

Of course it is easy to formulate a "strategy" or "vision" for R&D if you don't care what the outcomes are. Just convene a meeting of powerful "stakeholders," smile a lot, get them to free associate out loud, and then put together a collage of the most harmless, watered down words they emitted, making sure they are all sure that the check is in the mail to them personally. This happens often enough in the world, and is a major source of programs which are very effective in enriching middlemen/middlewomen in the short term, and of producing absolutely nothing in the end.

But what if we want results?

---

* The views herein are personal views, and not the official views of any organization. However, as work done by a US government employee, it may be freely copied and used, so long as this footnote is retained and there is use or reference to the entire chapter.

If we want results, strategy formulation for R&D is no different in principle from strategy formulation in other areas. First, we have to have some idea of what kinds of results we actually want. In principle, we need a kind of global metric, to define what are better results and what are worse results. For many years, economists inspired by utilitarian philosophers like John Stuart Mill, argued that we need to define a *utility function* $U(\underline{X})$, which gives an overall rating or score to any possible outcome $\underline{X}$. The vector X is simply a collection of all the n variables, X_1 through X_n, which we view as important in evaluating the outcomes. At this stage, they do not have to be *measured* variables; the challenge, to begin with, is to define what we really want, whether we can measure it or not.

If we have one idea of what we want on Mondays, and a different idea on Tuesdays, we usually end up fighting ourselves. We end up in a situation which is worse according to *both systems of preference* than some other situation we could have gotten to if we were more consistent. Thus if we want results, we will try to be consistent about what we want, at least at any one time, as we are formulating a strategy. We need a utility function. This is a key principle which I will discuss more. Of course, we also need to be very adaptive as well in S&T policy, as new information comes in, changing our options and what we expect to get from them.

For many years, economists tried to relax the requirement to have a utility function. They argued that all we need are “indifference curves,” which can be considered to be a kind of “ordinal utility function.” We only need to compare different possible outcomes, not to give numerical ratings. But soon after World War II, Von Neumann and Morgenstern (1953) showed that this is not good enough, if we live in a world of uncertainty. For all its initial limitations, Von Neumann’s work on cardinal utility theory is an essential prerequisite to real competence in strategy thinking for science and technology.

In the 1960s, Hitch and others (1967) did an excellent job of explaining and translating these basic facts of life to the defense community, with support from Robert McNamara and John Kennedy. Most important, they gave a simple example from economics about what happens if we have two different national measures of value, which we could call X and Y, used in different agencies like army and air force. Suppose we face a menu of a hundred possible projects, some of which do more for X, some of which do more for Y. If one agency values X twice as much as Y, it will choose one set of projects for what it controls. Its utility function, whether explicit or not, is 2X+Y. If another agency values Y twice as much as X, its utility function will be 2Y+X. But what if both agencies could agree on a single utility function, X+wY, for some w? (Or what if the Secretary of Defense could impose a common w on both agencies?) Usually, there will exist many values of w which would result in better outcomes BOTH for X AND for Y. This is an example of the principle of Pareto optimality, of aiming for win-win outcomes, which is essential to effective strategy making in many realms. There are many tricky issues in actually implementing this principle, as Hitch discusses, but it is fundamental and universal.

Even many experts in “systems theory” have some important misconceptions about Pareto optimality. When there are M different actors, all making different choices based on different utility functions, we end up in a “game.” If each actor takes the actions of other actors for granted, and tries to maximize his/her own utility function, the collective outcome is called a "Nash equilibrium." Many engineers call this a "Nash solution," and imagine it is a good thing to implement or attain. But in fact, the Nash equilibrium is usually far inferior to the best that the actors could do, if they cooperated with each other. A major focus in the work of Von Neumann, both in mathematics and in US S&T policy, was to try to get past Nash to Pareto, or at least to something that won't blow up and kill us all, quite literally (Poundstone 1993). In market design – a major tool of S&T policy – one of the major goals is to design rules or games in which the Nash equilibrium does get closer to a Pareto optimum.

Rules for managing proposal review and R&D allocations involve similar issues – at all levels, from international negotiations, to Congress, down to individual review panels. In areas like climate change, it is especially important to try to design systems which offer an efficient, rational tradeoff between variables like greenhouse gas emission, dependency on fossil oil and cost or GNP – any one of which poses risks worth trillions of dollars and relevant to human survival as such if we do not work harder for such efficiency. Failure to aim for such an efficient tradeoff was perhaps the main reason why climate legislation failed to get anywhere in the Senate in 2009.

How would one apply Hitch's analysis to agencies like NSF or NIH? Crudely speaking, NSF leads the nation in advancing one variable, the development of fundamental understanding of science and principles of technology. But NSF's choices can have a big impact on other sectors – for example when there is a possibility of a breakthrough in energy leveraging special capabilities available only at NSF. NSF and NIH and other agencies can all serve common national values, even though they have different missions and capabilities, by always trying to monitor BOTH the impacts on their core mission and their leveraging of the resulting capabilities. This is one way to interpret or follow through on NSF's two review criteria – "intellectual merit" (seen as benefit to basic understanding) and "broader impacts." Interagency partnerships are another useful mechanism for trying to overcome inconsistency effects and achieve government-wide efficiency, so long as they do not trample on the diversity of skills and ideas which is also essential.

Real follow-through on this principle is not easy, but is important. At agencies like NSF, it is usually hard to align judgments about broader impacts with serious global strategic thinking, rather than casual politically correct blessing from narrow specialists more interested in their own field. Even at the level of Congress, it requires special energy to rise above the tendency to pacify constituents concerned about big issues with empty gestures whose main effect is to enrich a favored stakeholder. Concrete strategic dialogue about substantive global issues, engaging the entire organization, may be more useful here than traditional approaches. It is sometimes useful to have joint initiatives (NSF 2002) where strategic thinkers from one area have a dialogue with technical experts from other areas, on the same panel, and questions of broader benefits are discussed in a concrete strategic way. Of course, it is also important to avoid degrading core missions of agencies like NSF, by encouraging them to fund what other agencies could handle better, or allowing their stakeholders to corrupt the process. There are many such balancing acts here, where managers who tend towards one extreme philosophy or its opposite can easily wreak havoc.

2.2. Probability, Risk and Uncertainty

R&D funding has often been compared to wildcat drilling. If one tenth of the projects hit pay dirt, that can pay for all the rest of the activity. If no one takes risks, one can achieve certain results – a certainty of achieving nothing really useful. When researchers propose to apply widely-known noncontroversial methods to address well-known research challenges or "hot topics," this is sometimes called "low risk research." But in actuality, it is much riskier than "high risk research" because it is far less likely to produce anything which hasn't already been done by someone else. Likewise, "long term" research sometimes produces notable benefits very quickly, when it is really a matter of trying out a new and controversial approach grounded in new principles.

Several consequences follow from these facts of life.

First, true competence in strategic thinking for R&D at all levels requires that we go beyond the basics (section 2.1), and develop a deep appreciation of probabilities and risk and decision trees as a way of life. ("Real options" are part of that.) In the 1960's, Howard Raiffa (1968) of the Harvard Business School initiated the field of "decision analysis" through decision trees, which translated the core concepts of Von Neumann into workable principles for management. His work, or the equivalent, is an essential prerequisite here.

Raiffa pointed out that we as humans cannot get away from the need to assess and continually update *subjective probabilities* – the probabilities of key events conditional on everything we know at the time. Can't make (rational) decisions without it.

Sometimes, in managing a review panel, I ask: "Try to project what the consequences will be, both direct and indirect, if we fund this particular proposal here and now. And project what *difference* it would make, versus what happens if we do not. What do we risk losing, in the end, if we do not fund this? Everything you know is on the table. For example, if you know that another agency is already funding it, that would affect your expectations. How large could those net benefits be to science and to society, versus those of competing proposals?" On rare occasions, a panelist will object that he certainly does not know, and cannot possibly answer the question. While I strongly encourage and support humility, I stress that we have no choice. In spending taxpayer's money, someone has to make a judgment on this. We are stuck with making the most informed, wisest subjective estimate we can, and we depend very heavily on sources of information like panels which at least clarify the nature of the uncertainty – most of all, the probability of a really large net benefit. (Of course, the real protocol goes through a few stages – sensitizing people to the two larger review criteria, raising the issue for specific proposals, asking what we might lose in concrete terms, and only in the end debating about larger scale tradeoffs issues.)

Raiffa also showed, through experiments, that world-class experts not trained in probability estimation tend to a really terrible job of it. He asked experts like those at Harvard and high-up jobs elsewhere to take a test on "guesstimating" various numbers. For example, what were car sales in the US in 1970 (no checking internet!)? What is a number so high that there is only a 5% chance sales could have been that much? That it would be impossible they could be so much? On average, with such experts –

"impossible" or "0.1% probability" actually happens 1/3 of the time, while "implausible" or "1%" occurs about half the time. This probably results from experts *applying* certain tacit assumptions or conventional wisdoms, without really thinking about the probability that the assumptions themselves might be wrong.

This in turn has many implications. It fits with the classic belief that the more traditional types of review panels, which gravitate towards median opinion, tend to be far more conservative and risk-averse than rationality calls for. There has been a lot of high-level coaching in recent years, urging people to work hard to think out of the box more, be more open to new ideas and new paradigms, to question their assumptions, and to look for more radical, transformative research. That does appear warranted – though, again, it's not entirely easy to really follow through. The most important lesson for many of us is to discipline our own thinking more than untrained people normally do, and truly suspend judgment more than people naturally do. In managing panels, it is important to ask questions about the reasons why people have the expectations they do, to probe the limits of uncertainty, especially in enforcing full consideration of minority opinions. To some extent, it also calls for hedging portfolios, to account both for diversities in subjective probability distributions and in awareness of where the results might be used.

Raiffa also tries to teach a subjective appreciation of the importance of "buying information." This is also a crucial part of the research process. It is very sad, for example, when people debating the cost of technology A versus the cost of technology B do not pay enough attention to the R&D needed to buy the information on how cheap we might make EITHER of the technologies be.

Disagreements about probabilities can be turned at times into vehicles for agreement at the policy level. For example, if one advocate says that fuel A has a 90% probability of being the best and cheapest, while another advocate disagrees and says that B has the best chances, both might be persuaded to agree to a "level playing field" in which equal incentives go to both and the market will decide. But at the R&D level, it is not so easy, since judgments must be made about what technologies offer big enough impact.

Another key dimension of risk is technology risk versus market risk. At an aerospace conference a few years back, many advocates of return to the moon or of space solar power argued that we should use 1960's launch technology as the foundation of that effort. "It has lower risk." But the cost of 1960's launch technology is so high that there would be a huge "risk" (near certainty) that the market simply would not accept electricity from space solar power based on that technology. The overall risk is far less in exploring the decision tree of how to reduce costs, at least down to what the market could bear, even though that requires a lot of "wildcat drilling" and strategic thinking. Many in the aerospace industry did feel their chances would be better using 1960's (Ares) technology for return to the moon – but even there, the "market" (Congress and White House) simply balked in 2009 and 2010 when the real price tag turned out to be much higher than promised. Logsdon (2006) discusses some aspects of how this error was made. Relations between myopic lobbyists and key appropriators were of course a central factor; in other words, a lack of strategic thinking at that level was a special problem. Ignoring market risk, misleading the public and taking a myopic view could possibly result in the end of these efforts, sooner or later, if there is not a quick and effective reversal of course.

Whenever we focus on difficult tangible long-term goals, like achieving a sustainable global energy system or economically sustainable (profitable) human settlement of space, the decision tree has many levels. We often visualize it as a kind of pipeline stretching from high-potential highly uncertain possibilities through to technologies ready for immediate deployment. Mankins (2002), Ramirez and Sauser (2009) and Conrow (2009) provide practical guidelines for keeping track of this complexity.

At the final stage of this pipeline, we often have no practical use for any option but the lowest-cost option, versus other options which do the same thing. For example, if we compare sustainable technologies for producing kilowatt-hours of steady, baseload electricity in the same places, high cost options are not of immediate use. However, when technologies are not fully worked out, and the future potential to reduce costs is not yet known, we need to focus on the probability and strategy for getting the costs lower. To do this right, we first need to be able to assess the biases and uncertainties of the diverse sources of cost estimates for technologies at different stages in the pipeline. Merrow (1993) and Flyvberg (2006) offer excellent introductions to this topic. In the end, it is essential to develop an understanding of the specific underlying principles which offer the most serious hope of real breakthroughs, in order to distinguish new directions which have serious hope from those which are basically just repackaged hope, empty promises from big names, and lobbying from unworthy money-hungry coalitions.

2.3 Dynamics, Intelligence and Where We Are Today

However great their heuristic value, decision trees are simply not a practical way to describe the full complexity of S&T leadership. They are not powerful enough even for much simpler problems, such as how to manage a hybrid car for minimum fuel use and pollution (Prokhorov 2008).

The S&T challenge is closer to a class of problems which have been studied very actively in engineering and in neural networks in recent years: the challenge of maximizing expected utility, U, over multiple time periods, in the face of a complex, nonlinear, stochastic environment. Great progress has been made (Si et al 2004, Werbos 2009a), which is relevant in two ways. First, it is relevant to how we do strategic thinking. Second, it is an important part of what we should be thinking about in S&T policy. All of this progress builds on the earlier work of Von Neumann and Raiffa, but takes it further into the realm of complexity and adaptation.

In reviewing this area (Werbos 2009a), I have concluded that the most important Grand Challenge for science and technology here is to understand and replicate that level of ability which exists in the brain of the smallest mouse in two key areas: its ability to learn to "predict" (and reconstruct) its environment, and its ability to learn to maximize something like its probability of survival over multiple time intervals. The mouse brain can learn to predict and interpret much huger flows of data than any engineering system today, and cope with tasks which are also very difficult for engineered systems.

This leads to some interesting paradoxes. How is it that a tiny artificial neural network, used in areas like mortgage risk assessment, can do so much better than full-fledged human brains on the same task? (This is only one of many examples, and not my main focus.) The main reason is that we can make the artificial network really pay attention. If we made fuller use of our human brains and minds, we could do much better. The brain of a human is more powerful than that of a mouse, but, just as simple decision trees can help us in using our brains better, so too can this further extension.

In the early days of optimization by linear programming (LP), LP people would often come into a factory and show them how they could reduce costs a huge amount by shifting over to using LP. But when they came back a few years later, they learned that "the humans watched the LP, learned its tricks, and then did as well or better themselves." In the same way, we can learn some tricks and even some useful mathematical tools from recent progress in intelligent systems.

Another new computational tool, crucial to optimizing large systems, is a method sometimes called "backpropagation" (Werbos 1994) and sometimes called "reverse differentiation" (Werbos 2005). Whenever we focus on a single outcome at a time, this method allows us to work quickly through the dynamics of any large system, in order to pinpoint the inputs or parameters which have the biggest impact on that outcome. There is an old poem "For want of a nail… the kingdom was lost." In every major domain of R&D which I have waded into, there do turn out to be several crucial "nails," usually neglected by the highest level powers that be, which drive the outcome of the overall effort. I have often wondered whether my subjective understanding of backpropagation has been important in helping me find a lot of those "nails."

Perhaps the most central concept in the new literature is the concept of "value function." This is not the old utility function, U(X), which describes our long-term or intrinsic values. It is a kind of secondary or strategic utility function, sometimes called J(X) and sometimes called V(X), which describes how well we are doing right now. In effect, the new mathematics tells us how to adapt a short-term performance measure J, which tells us how well we are really doing in getting to our true long-term goals. Performance measures are already quite common in all kinds of management, but they often undermine the long-term goals instead of reflecting what they really require of us. The same new body of tools also offers ways to calculate $\underline{\lambda}(X)$, an entire vector of value signals which economists would call "shadow prices."

As a practical matter, then, we will be forever dependent on trying to do a better job of mobilizing individual human brains and minds to really pay attention to high-level strategic goals here, to define value measures which reflect an ever-changing structure of subgoals flowing from the larger goals, and overcoming the innumerable obstacles to effective follow through.

In some areas, management consultants have said "if you can't measure it, you can't manage it." In the complex arena of new S&T, experience shows that "if you can't understand it, you can't manage it or measure it without screwing it up." Mechanical systems which try hide from this basic fact of life, such as 25-point scoring systems to decide what to fund, are a very popular way to screw up. Kotter (1996) and Morton (1970) have described how important it is – and what can be done – to overcome the usual bureaucratic entropy and illusions which afflicts these areas. In brief – "with vision the people perish" (Proverbs 29-18).

Classical artificial intelligence also offers some important insights here (Simon 1996). Decades ago, Newell, Shaw and Simon – the real founders of artificial intelligence – worked hard to build connections between algorithms for problem solving and analysis of human problem solving, including scientific creativity. In building systems to answer difficult questions or prove theorems, they found that the distinction between "forward induction" and "backward induction" is extremely important. In forward induction, people start from what they know now, think about possible things they could do to go a step or two forward, and try to learn to evaluate which step ahead would have the greatest value. In backwards induction, people define a goal, and try to work backwards from the goal, to find a path to get there – first at a very abstract level ("highly chunked") and then in detail.

They found that forwards induction can work well in simple, well-defined problems. But for very challenging, complex problems fraught with uncertainty, only backwards induction offers a good chance of getting to the goal. This is extremely important to S&T, where the challenges are as complex, challenging and fraught with uncertainty as any in human life. The most important large challenges and opportunities before us all require more backwards induction in our thinking, but the skills and organizations in the world today are simply not well-suited to support that kind of thinking on a steady, reliable, intensive basis.

Some followers of classical AI have proposed that S&T and world economies should be managed via a hierarchy of explicit goals and subgoals very similar in spirit to what failed in the Soviet Union, in the face of great uncertainty. Modern neural network approaches, based on adaptation and value functions, are far more compatible with the use of market mechanisms or market-like mechanisms, which are essential in managing complexity, new information and unexpected change. Nevertheless, to explain the ability of mammal brains to solve difficult challenges, it is still essential to include a *variation* of backwards induction. It is essential that there be serious attention paid to "goal states" (and subgoals, and a kind of hierarchy of time scales), and that value functions be adapted in a way which fully reflects the best and worst possibilities for the future.
The mathematics of this are not so simple (Werbos 2009a), but the lesson for S&T policy is clear. Human society needs to provide more effective support for human brains to work out maps of where we could go, and of what we need to worry about, in mapping out possible futures, and guiding projects which can help us get there. It helps to ask people to think about these questions at the time of a review panel, but the strategic issues are far too complex for this to be a complete answer.

Going beyond what we can learn from intelligence at the mouse brain level, we as humans also have capabilities for empathy, for sharing experience, and for some kind of collective intelligence or deep dialogue (Werbos 2009a). It may be premature to make strong claims about these capabilities before we even understand the brain of the mouse in a scientific way, but they too are important for us to try to use and develop as best we can.

Many mathematicians have written at times about a classic study of creativity (Hadamard 1954) which is relevant to strategic thinking in general. At some level, successful strategic thinking and creativity are almost the same thing. Both involve charting a path to new paradigms and breakthroughs. Effective visual thinking, deep intuition and conscious mapping of the concrete landscape of possibilities was crucial to the success of all the great mathematicians studied by Hadamard. But Von Neumann was a unique case, where he used words and vision together. I would speculate that he did not just merge words and visual mathematical thinking in a fuzzy collage; rather, he implemented something like a decision tree, or the true scientific method we all pay lip service to, as a kind of metalevel to guide his image-based thinking more consciously than the others did. His unique contributions to US technology in World War II and after may be partly due to his unique learned habit of pushing questions and logic all the way to a logical conclusion, without being slowed down by the usual dampers and fears which inhibit most of us, until a logical strategic understanding was reached – at which point he, like the rest of us, did have to be careful and strategic in how to follow through past the social minefields which have always been a fact of life.

The cultivation of human creativity, human potential and human sanity (getting past alienated and irresponsible ideals of professionalism) is also important to effectiveness in S&T – but beyond the scope of this chapter.

2.4 Some Qualitative Observations

Years ago, when we co-organized a joint summer school on neural network engineering in Mexico, the Mexicans chose a curious symbol to represent the conference. It was a tree – the "sian kaan," the sacred tree of the Mayans, sustaining life by creating a link between the world of the clouds and the world of the mud. As I thought about, it realized that this was a good way to symbolize the challenges inside NSF at the time – the challenge of building better bridges between high-level visionaries occasionally lost in the clouds, and the people down in the mud of specific disciplines, closer to reality but also somewhat stuck and limited in movement. But traditional notions of bureaucratic efficiency and management and social systems design can easily create a situation where everyone is stuck in the mud.

In a recent major national conference on science policy in Asia, the Science Minister began by using that same analogy of a tree. There are so many aggressive and energetic tigers down there (people creatively fighting for funds for their own disciplines), but maybe we need a different kind of animal, at the top of the tree, to keep those tigers under control. Maybe like a cow, he said. In fact, there are funding systems in the world which operate a lot like putting a fuzzy-eyed cow at the top of the tree, having them send their droppings down to where the tigers are, and watching what happens. There are even a few idealists, lost in the clouds, who imagine that they can create dialogue between tigers who come from hungry lands by putting them together in the same room with just one piece of meat and no kind of protection mechanism. (Perhaps this experience is why the Minister proposed a cow…?)

Before that conference in Asia, a friendly group from Chengdu pulled me away from my previous plans, and carried me in a beat-up van along the road to Chongqing, stopping at an archeological dig where they exhibited a different kind of tree. This tree was a central symbol used by the ancient Shu kingdom, which appears to have left written records much older than any writing from further north in China. On their tree, they had dragons over the tigers, but eagles or fish hawks at the very top. I won't say more about the meaning of the dragon in China, since that is a very complex subject, but I will talk about the eagle, a very fast and clear-eyed creature.

Many of the fundamental questions of science and technology have a way of moving from one discipline to another (with a need to remember what was learned form the previous discipline) if one really chases after them in an energetic, strategic way. I think of the big questions as something like a rabbit, which moves quickly from field to field. NSF emphasis on cross-disciplinary research and partnership is one tool to try to make it possible to chase these rabbits more effectively, but we still need the eagles. Not just at the top of the government tree, but wherever a suitable habitat can be created. A highly creative and effective "eagle" in the private sector recommends Morton (1970).

It is useful to think of the entire system of R&D and technology as kind of gigantic neural network or food chain or garden. It is crucial to prune excessively aggressive invasive species, while maintaining a high degree of diversity, richness, and connectivity, while being realistic about the need for some sparse connectivity in each individual node.

3. Grand Challenges For the Coming Century

In 1999, the Foundation for the Future organized a workshop, Humanity Three Thousand, bringing together dozens of futurists from different backgrounds and schools of thought. It concluded with a debate on whether humans will become extinct by the year 3000, or firmly on the path to extinction. When all the competing trends are considered, the audience was divided roughly 50-50 on this question.

So long as there is considerable reasonable doubt about the survival of the human species, one might argue that the top of our hierarchy of values should be survival of the species. Certainly some of us need to think very hard about how to minimize the probability of human extinction.

The right hand side of Figure 1 depicts a refined version of this challenge.

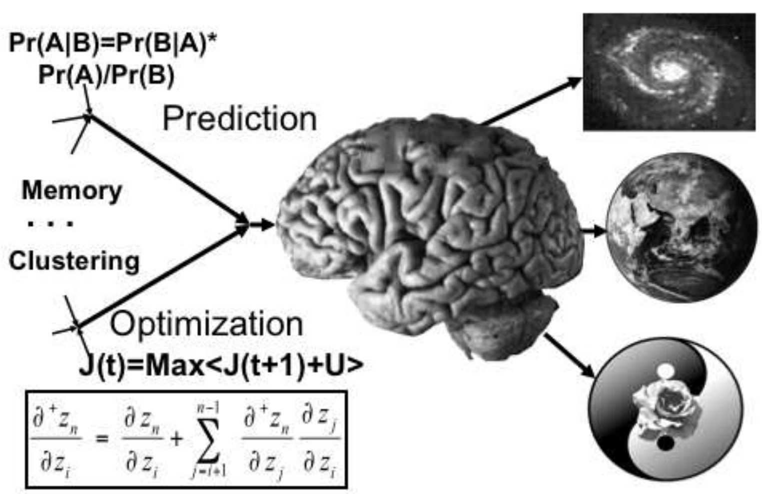

Figure 1. Iconic representation of four grand challenges

The image of the earth represents the grand challenge of achieving sustainability on earth. The core problem here is that we are still continuing relentless population growth, and relying on natural resources such as fossil oil which simply will not last forever. Some of would define "the sustainability problem" as a situation where we must change or die. The grand challenge here is to do what we have to do to survive. S&T is a crucial part of how we can address this challenge, and minimize the probability that humans become extinct on earth. Of course, it is good that NSF includes sustainability in this sense as one of the key pillars of research – including technologies and other measures to enhance sustainability, and also including strategic thinking on the subject.

One of the key subgoals, necessary to global sustainability, is the achievement of a sustainable global energy system. This is a complex challenge for rational strategic thinking in itself, described in great detail in Werbos (2009b). Other key subgoals are also part of intense new international dialogues (Millennium Project 2010), which provide an important starting point but need to be expanded and enhanced. Economic growth can be an important tool or subgoal to help us achieve sustainability, but many of us would view money as a tool for achieving survival, not vice-versa. All of the points discussed here, and more, are crucial to the energy example, but it would require another whole chapter at least to describe how they play out.

While at DOE and NSF, since 1979, I have seen again and again how most authoritative analyses of energy policy by economists or engineers are way off base, because each discipline on its own only provides a partial view. To see the true picture, one needs a broader vision, seeing through "both the eyes" at the same time, and integrating the images. Beyond that, for really useful policy guidance, one must also integrate sober understanding of administrative realities (as summarized here) and legislative realities. But even if one integrates all four into an effective strategy, the fifth element – hard core national and global politics – has no clean solution, and will probably continue to tax our rationality for as long as humans live on earth. It helps to remember that if mere mice can stay sane, and look to their future hopes, even with foxes ever ready to eat them, so too can we.

Furthermore -- the earth is not the entire universe. Looking to the far future, many of us believe that the rest of the universe, and even the rest of the solar system, is much bigger than the earth. Many of us believe that there is some hope that humans could achieve a permanent sustainable presence beyond the earth. There are uncertainties and debates about this – but certainly it is a grand challenge to maximize the probability of achieving this. This provides another example where rational strategic thinking can be used (Werbos 2009c).

Beyond earth and outer space, there are also issues of inner space. Some would say that there are deep and fundamental issues about the quality of life, and not just survival as such. If we all survive as unconscious robots, what have we accomplished? The third icon, the rose on a yin-yang, is intended to symbolize the grand challenge of humans living up to their higher inner potential and development. There are huge challenges in creating the kind of dialogue between cultures and nations which would let us make quantum improvements in this area. Today's efforts at improved K-12 education in the rich part of the world are relevant to this challenge, as a glass of water is relevant to an ocean. Perhaps this is the greatest, most unmet challenge of all, and the hardest to do full justice to in any of the world today.

Normal science and technology can make many contributions, large and small, direct and indirect, to these three grand challenges. But fundamental science is also facing three very fundamental, unifying challenges, as discrete in their own way as the challenge of putting a man on the moon, requiring similar strategic thinking:

1. The challenging of understanding and replicating the highest level of intelligence we can find in the brain of the smallest mouse (Werbos 2009a).

2. The challenge of finding the true “theory of everything,” the ultimate laws of physics.

3. The challenge of putting the phenomenon of life (and related self-organization) on a solid, universal mathematical footing.

Of course, Figure 1 only depicts the first of these challenges. All three require more creative and strategic thinking than they have received so far. All three require a more crossdisciplinary approach, grounded in basic unifying mathematical principles, which is not the same thing as having ten thousand equations per journal issue. Not only tools but insights from engineering and computer science offer crucial missing pieces with all three. I have often felt very deep regret that we couldn’t somehow get our acts together for a more direct, frontal strategic effort on all three, making better use of all our capabilities.

Between the three – I would crudely view challenge two as “the revolution of the past” (started by Newton), challenge one as the opportunity/revolution of the present (just now emerging as a tangible mathematical opportunity) (Werbos (2009a), and challenge three as the potential mathematical revolution of the future. All three merit major strategic efforts here and now, but of a different character. All three cry out for more “eagles” and systems to get them really moving. (By contrast, a lot of other important challenges, like some of the Grand Challenges from the National Academy of Engineering, could do well enough under the management of “dragons.”)

Of the three challenges, one may argue that challenge two is the one which does have sufficient coherent support already. For example, the Physics Division of NSF certainly thinks about the “theory of everything” challenge. They have issued a list of “great questions” which provide important, focused subgoals. But greater consideration of empirical results from engineering [(Werbos 2008) and of fundamental mathematics may allow more aggressive exploration of the unknown, and new possibilities for breakthroughs, in my view.

Challenge one cries out strongly for new thinking and new funding vehicles as well. The Cognitive Optimization and Prediction topic (NSF 2007) recently funded by the Office of Emerging Frontier Research Initiation of NSF addresses what is possible, but it was only a one-year effort and was unable to fund more than a handful of the breakthrough areas here. Figure one depicts five major subgoals or areas for research in that sphere: (1) the goal of understanding and replicating the brain’s ability to learn to predict and reconstruct reality, to “predict anything”; (2) the goal of understanding and replicating its ability to learn an optimal strategy of action, in tasks such as maximizing the probability of survival in challenging environments where absolute guarantees of survival are impossible; and (3) using these tools and understanding to make as much real contribution as possible to the three larger challenges I discussed at the start of this section, and learning what we can from those testbeds. Each of these subgoals, in turn, requires strategic planning and a diversity of new research projects to make it attainable, drawing on mathematical principles illustrated in the figure. Research related to the optimization challenge has begun to emerge more clearly in recent years (Si et al 2004, Werbos 2009a). Research related to the prediction challenge has perhaps gone backwards in the past few years, but new strategic thinking – currently in discussion at conferences (as posted at www.werbos.com/Mind.htm) but not yet published – might be able to turn this around.

Recent work on xenobiology (Ward 2005), on creation of artificial cells, on the genome as a learning system ("metagenetics"), older work on the quantum mechanics of autocatalytic systems in infinite solution and on self-organization in general, provides an important start for addressing challenge 3, but requires a lot more fundamental thinking even to clarify and integrate the roadmap. There are a lot of mathematical principles and ideas which cut across the three challenges as well.

All of these six grand challenges – three for science and three for society – are basically "hopes." A sane and healthy brain puts more of its energy into hopes than into fears. A major challenge to science management is to unleash human creativity, not only through providing direction and encouraging more strategic thinking, but also through encouraging more hope and less fear. But some fears are necessary, and also demand strategic thinking. There are powerful vested interests at work which tend to downplay the serious uncertainties and threats which can occur through the misuse of science. Among those which I find most worrisome are the risk of military computers so intelligent that they risk the scenarios in the "Terminator movies," the risk of abuse of electrical brain stimulation, the risk of nuclear technology worsening the life-threatening problem of nuclear proliferation, and the risk that we may start unleashing new forces more energetic than nuclear fusion on the surface of the earth (if we are smart and creative enough to figure out how). Organizations like the Lifeboat Foundation have provided some venue for strategic thinking regarding such threats, but we have a long way to go to do justice to them. The most important lesson from such threats is that we need to give higher priority to **basic scientific understanding**, as a strategic priority, so that we have some hope of knowing what we are doing before we blunder into problems we do not understand. Strategic thinking and funding to back it up are what we need.

References

Conrow, Edward H, 2009, *Estimating Technology Readiness Level Coefficients*, presented at AIAA Space 2009 Conference and Exposition, paper number AIAA 2009-6727. Reston, Virginia: AIAA.

Flyvberg, Bert 2006, From Novel Prize to project management: Getting risks right, *Project Management Journal*, August 2006.

Hadamard, Jaques, 1954. *The Psychology of Invention in the Mathematical Field*. Dover.

Hitch, Charles Johnson, 1967, *Economics of defense in the nuclear age*. Ann Arbor, Michigan: University of Michigan, 1967, http://www.lib.umich.edu.

Kotter, John P., Leading Change, Harvard Business Press, 1996

Logsdon, John P., 2006, A failure of national leadership: Why no replacement for the space shuttle. In Steven J. Dick and Roger D. Launius, eds, *Critical Issues in the History of Spaceflight*. Washington DC: NASA. NASA report/book number NASA SP-2006-4702.

Mankins, J.C. 2002, Approaches to strategic Research and Technology (R&T) Analysis and road mapping, Acta Astronautica 51, No 1-9, p.3-21.

Merrow, Edward W., 1993. *Cost Growth in New Process Facilities*. Santa Monica, Ca.: Rand Corporation, report P6869. www.rand.org/pubs/papers/p6869

Millennium Project 2010, *State of the Future 2010.* Washington DC: the Millennium Project. www.stateofthefuture.org

Morton, I.A. 1970, A model of the innovation process in a science-based fragmented industry. In M. Cetron and J.D. Goldhar, *Science of Managing Organized Technology*. Routledge.

NSF 2007. *Emerging Frontiers in Science and Innovation 2008*. NSF document number NSF-07-579. Arlington, Virginia: National Science Foundation. http://www.nsf.gov/pubs/2007/nsf07579/nsf07579.htm

NSF 2002. *NASA-NSF-EPRI Joint Investigation of Enabling Technologies for SSP (JIETSSP).* NSF document number NSF-02-098. Arlington, Virginia: National Science Foundation. http://www.nsf.gov/pubs/2002/nsf02098/nsf02098.pdf

Poundstone, William 1993. *Prisoner's Dilemma*, Anchor Books

D. Prokhorov, D. 2008 Prius HEV neurocontrol and diagnostics, *Neural Networks,* 21, pp. 458-465

Raiffa, H. 1968, *Decision Analysis*, Addison-Wesley

Ramirez-Marquez, J.E. and Sauser, B.J. 2009. System development planning via system maturity optimization, IEEE Trans on Engineering Management, August 2009

Si, Jennie, Barto, A.G., Powell, W.G., and Wunsch, D. 2004 (Editors). *Handbook of Learning and Approximate Dynamic Programming* (IEEE Press Series on Computational Intelligence), Wiley-IEEE Press

Simon, H.A. 1996. *The Sciences of the Artificial (3rd ed.)*. Cambridge, MA: The MIT Press.

Von Neumann, J. and O.Morgenstern 1953. *The Theory of Games and Economic Behavior*, Princeton NJ: Princeton U. Press

Ward, Peter D. 2005, *Life as We Do Not Know It: The NASA Search for (and Synthesis of) Alien Life*, Penguin

P.Werbos 1994. *The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting*, Wiley.

Werbos, P. 2004. Backwards differentiation in AD and Neural Nets: Past Links and New Opportunities. In H. Martin Bucker, George Corliss, Paul Hovland, Uwe Naumann & Boyana Norris (eds), *Automatic Differentiation: Applications, Theory and Implementations*, Springer, New York, 2005. http://www.werbos.com/AD2004.pdf

Werbos, P. 2008. Bell's Theorem, Many Worlds and Backwards-Time Physics: Not Just a Matter of Interpretation, *Intl J Theoretical Physics*, e-pub date 2 April 2008. http://arxiv.org/abs/0801.1234

Werbos, P. 2009a. Intelligence in the Brain: A theory of how it works and how to build it, *Neural Networks*, Volume 22, Issue 3, April 2009, Pages 200-212.

Werbos, P. 2009b. Technological Solutions for Energy Security and Sustainability. in Gal Luft and Anne Korin, eds, *Energy Security Challenges for the 21st Century: A Reference Handbook* Praeger. Expanded version available from Nature Precedings http://hdl.handle.net/10101/npre.2008.2131.1> (2008).

Werbos, P. 2009c. Towards a Rational Strategy for the Human Settlement of Space, ***Futures***, Volume 41, Issue 8, October