Optimal, Real-Time Control – Colliders*

J.E. Spencer

Stanford Linear Accelerator Center, Stanford University
Stanford, California 94309

Abstract: With reasonable definitions, optimal control is possible for both classical and quantal systems with new approaches called PISC(Parallel) and NISC(Neural) from analogy with RISC(Reduced Instruction Set Computing). If control equals interaction, observation and comparison to some figure of merit with interaction via external fields, then optimization comes from varying these fields to give design or operating goals. Structural stability can then give us tolerance and design constraints. But simulations use simplified models, are not in real-time and assume fixed or stationary conditions, so optimal control goes far beyond convergence rates of algorithms. It is inseparable from design and this has many implications for colliders.

Introduction

Prediction is very difficult, especially about the future
— Niels Bohr

Predicting the future isn't difficult until one demands a high correlation between events and their predicted times of occurrence. For deterministic systems having models with fast predictive cycles compared to their time scales, stable control should be possible. The choice between closed and open loop control then revolves around which is better, the model or the feedback system. Whether either can be done fast enough or accurately enough in real time depends on resources relative to problem demands.

Because there are other stochastic effects beyond the known dynamic nonlinearities, we need both feedback and feedforward for optimal control. We use *causal feedforward* for the special combination that takes optimal advantage of the collapse of the probability distribution. External noise effects are shown in Fig. 1. An example is line noise in magnet supplies that maintain constant current via feedback but leave random field errors from eddy currents and hysteresis. Knowing if tolerances are violated from above or below improves the speed and quality of correction.

New stochastic effects are expected as the energy and luminosity of colliders increase associated with quantum effects and the growth of complexity (degrees of freedom) e.g. it becomes harder to define the system, its variables or their constraints even assuming the system is isolated. Further, even if the dynamics and control model are linear, their physical realization with measurement, roundoff or overflow errors can be nonlinear with chaotic regions. The meaning, possibility and implications of optimal real-time control under such circumstances are discussed.
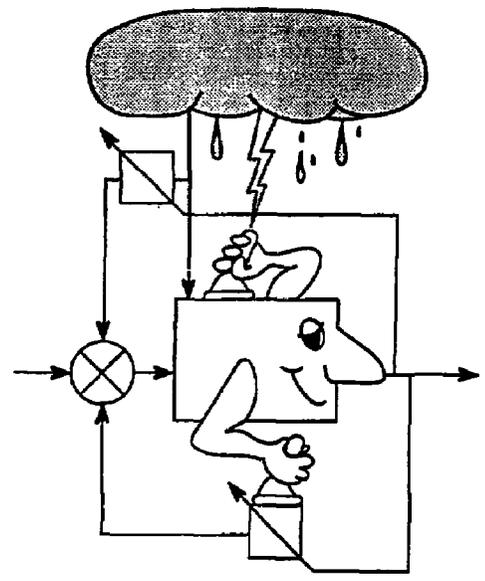
Figure 1: Real-Time Control in the Real World

Problems with Complexity

Despite books and conferences[1] relating it to entropy, order and information, there is still no universal measure of complexity. Real-time control of colliders is a good place to explain why. The largest machines ever are proposed to learn everything about the smallest distances in the least possible time. We consider whether this is consistent and, if so, at what cost and with what techniques?

Figure 2 compares the growing complexities of several systems where the vertical axis can be thought of as the bits of information needed to specify the system state. For ring colliders this is proportional to their radius (or diameter for two rings) which also relates to cost. For single chip DRAM and Intel micros, it is the number of bits or transistors. The time axis gives the first available date of the various products - physics or chips. Thus, the LHC and SSC are located by the lower bounds on their detector dates. The vertical scale is limited above by the human brain with a dozen billion neurons[2] and below by quanta such as photons in pure helicity states with one bit of information classically when we know the wave vector.
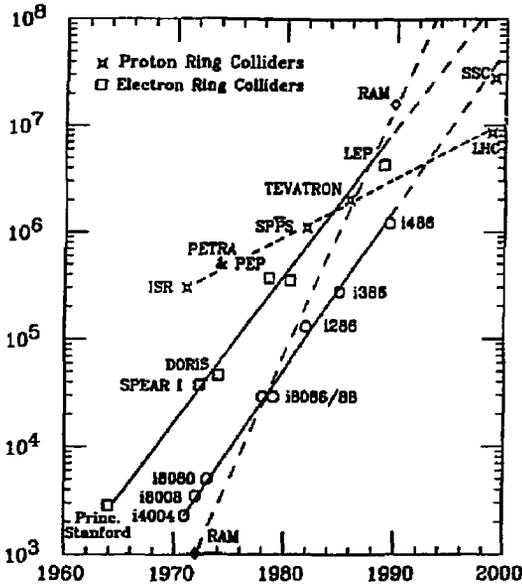
Figure 2: The Growth of Spatial Complexity

Excepting SSC, each class shows an exponential increase in spatial complexity i.e. the number of elements due to increases in size or density. Extrapolating, one can make predictions e.g. it's hard to see how SSC can produce any physics before 2005. Trends within a class help us predict capabilities e.g. what the i586-i786 chips will provide and when[3]. LEP uses some 2000 micros at the i286/386 level. Including memory, LEP's system approaches the human brain in complexity and the SSC may surpass it[2].

While LEP appears inefficient it provides flexibility and redundancy that's applicable for LHC. To compare to SLC, we could, in principle, count their control program statements and calculate execution times to determine the more complex system assuming both were optimal. But storage rings don't need a program most of the time so they must have lower temporal complexity. However, LEP is *much* larger or more spatially complex so the two emphasize complementary resources to solve the same problem. In either case the growth in complexity makes reliability, adaptability and flexibility increasingly important.

Real-time computation, as part of real-time control, faces similar but easier problems. In $n \times n$ bit multipliers, you can interchange space-time complexity going from pure memory methods with as little as one memory cycle but n^2 memory locations to serial multiplication with order n^2 machine cycles but no memory. The neural net is an ideal theoretical tool to study such problems e.g. we can use it as an associative memory or optimal parallel multiplier[3] or to simulate *any* complex process whose K-entropy is large or infinite. Neural nets are also practical and needed for complex systems such as colliders -- possibly in chips like the i686 or i786 by the year 2000 for the LHC[3,4].

Comparison of Colliders

Our ability to define our system, identify the variables and conditions they must satisfy in terms of a consistent measure of merit is itself a measure of problem complexity. Current linear collider (LC) designs for the same energy and luminosity differ wildly even before any consideration of how one finds or holds the optimal system state.

Some questions are: 1) What is the optimal collider for physics, 2) What is optimal for electrons, 3) How does one optimize time complexity for LC's and 4) Have we really explored all the possibilities? Some of these were explored by Richter[5], Rubbia[6], Panofsky[7] and by Palmer[8]. Restating 1): Why not build a 2 TeV electron machine rather than a 20 TeV SSC? Also, for 2): Why not build an LC since rings become dominated by radiative losses?

$$Cost \ = \ Fixed \ + \ Capital \ + \ Power$$
$$\approx \ C_o + C_1 R + C_2 \frac{\gamma^4}{R} \ \rightarrow \ C_o + 2 \ \overbrace{\sqrt{C_1 C_2} \gamma^2}$$

Both cost and radius scale with the square of the energy and the 'optimal' ring approaches an LC asymptotically. SLC and Tevatron scale linearly with length or energy. From Fig. 2 and the fact that the Tevatron has en times the top energy of LEP[8], it is preferred as long as its cost or complexity per unit length is $C_1^{T \cdot V} < 2 C_1^{LEP}$. For SLC, the equivalent path length is less than one-fourth LEP which also roughly measures their relative capital costs.

But no one can really say what C_1 or C_2 are un. l the physics is done since they depend on integrated lumir sity, detector capabilities and how many Z's are neede i.e. the underlying problem complexity. LEP, SLC an. he Tevatron with their various detectors are different a gorithms for this problem. Each has produced comparable results for different costs in time and other resources. One can estimate the enhanced reliability (and costs) needed for future machines from earlier ones[4,9]. When reliability and temporal complexity of LC's scale linearly, they are preferred and 0.5 TeV is competitive with LHC.

One approach to these problems is via correlations:

$$(\cdots f_{i_{\pm}}(t) f_{j_{\pm}}(t - \tau) \cdots) \leftrightarrow (\cdots f_{i_{\mp}}(t) f_{j_{\mp}}(t + \tau) \cdots)$$

with $f_{i_{\pm}}$ a state variable whose vector is $\mathbf{f}(t) = \{f_i(t) : i = 1 \cdots 2d\}$, with time series $\{\mathbf{f}(t_j) : j = 0 \cdots n\}$ where $t_j = t - j\tau$. We can predict $f_i(t_j > t)$ for chaotic systems when $n \geq 2d$[10] because we can measure and study the dimensionality d_i. We can also control and optimize f_i and thus $\mathbf{f}$. The correlations can be interpreted at fixed time as giving instantaneous envelop equations to any order. Similarly, from the time dependence at fixed location we can monitor the K-entropy which measures the average rate of information loss or phase space deformation i.e. the predictability with time. To keep all correlations bounded we must supply information at least as fast as it is lost through both design and control[11,12]. As $K \rightarrow \infty$ the problem becomes purely statistical but still allows optimal prediction when we know the statistics. However, adaptive prediction is still only worse by log n with n the number of steps in advance time for LMS algorithm.

2

Model-Driven, Adaptive Control System
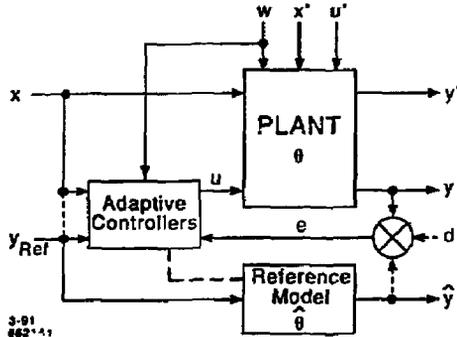

Figure 3: A General-Purpose, Real-Time Control System. It can cancel noise (w), control a plant, given the model, or find a model for a given plant and optimize it i.e. $\hat{\theta} \to \theta$. It is useful for complex systems where u', w, or x' are variables that may influence output but aren't controlled or monitored or when the system is changing and we can't set tolerances or respond fast enough. Ref.[3] gives a case where x and u get contaminated and have to be translated into hardware syntax while also learning.

Optimal, Real-Time Control

Although independent of space-time bounds, our most important postulate here is that of Turing and Church[4]. Also, we assume that complex systems need comparable or greater complexity in their controllers consistent with Shannon's measure of information. We can't drop degrees of freedom or the highest order of a plant or subsystem. Lastly, we can't have too much computing power because it improves the ability to adapt to and control unforeseen, unstable or chaotic conditions e.g. the K-entropy.

We begin by assuming a 'program' can be written that will simulate *any* physical process. It is also assumed that for any specific problem, one can translate this program into 'matched' hardware solutions such as ASIC's. This combination defines an algorithm for the problem. The classic example is the von Neumann computer but it can not be optimal because of its serial hardware and software.

Considering the program as an equivalent description of the problem (or process), information theory tells us the efficiency of any encoding or most efficient algorithm. But real-time control differs from computational algorithms in the importance of external, asynchronous effects that can be more important than the calculation. For optimal solutions, we therefore assume that we must use the process itself even though we may not be able to produce an equivalent software procedure. The hardware, in a very real sense, is the optimal 'program' or problem algorithm that we need to realize in our control system.

We still lack optimal control until we obtain the model and optimize it for the *existing* plant. But we know how to do that[12,3,4], at least in principle. Furthermore, ANN's

or artificial neural nets can replace computers in all ways[3] and don't suffer their bottlenecks[4]. They are also the paradigm for the ultimate, parallel, pipelined processor that handles concurrent, parallel data naturally. They apply to computers, colliders and detectors so they provide both the controller and model needed in Fig. 3. Some of these capabilities could be available in the i686 chip[3].

This is in direct contrast to conventional, model-based systems that lack any explicit reference to control. Since they are also used for design, we lack optimal designs as well as the tools to achieve them which can increase control costs C_{ic} drastically. In contrast to ring colliders, the beam characteristics in LC's have tighter tolerances and ultimately depend on source performance. Another paper addresses the source design from this perspective[11].

References

[1] *Measures of Complexity*, L.Peliti & A.Vulpiani (Eds.), Springer-Verlag, Berlin 1988. See also Ref.[10].

[2] *The Brain*, R.S.Thompson, W.H.Freeman and Co., New York 1985. The information to specify a state is different from that required to set it. Number and connectivity between neurons is then more appropriate because the complexity of one neuron can then be > 10^4. Likewise, there are large differences between bits and transistors in micros and DRAM[4].

[3] A.Corneliusen et. al., Computation and Control with Neural Nets, NIM A293(1990)507-516.

[4] J.E.Spencer, Real-Time Applications of NN, IEEE Trans. Nuc. Sci., NS36, No. 5(1989)1480. Information capacity of a 'forward' weight can be taken as 1 bit. To specify a simple net of two equal layers with n neurons apiece requires an $n \times n$ connection matrix but a linked list requires only $10^4 n$ while a planar, nonintersecting circuit takes at most $6(n-1)$ or $4(n-1)$ without interconnects within a layer i.e. order n.

[5] Burton Richter, Very High Energy Colliders, 1985 IEEE Part. Accel. Conf., Vol. NS-32, No.5(1985)3828.

[6] Carlo Rubbia, The 'Future' in High Energy Physics, 1988 European Part. Accel. Conf., Rome 1988, 290.

[7] Wolfgang K.H.Panofsky, A Perspective on Lepton-Photon Physics, XIVth Int'l. Symp. on Lepton-Photon Interactions, Stanford, CA(1989).

[8] R.B.Palmer, Prospects for High Energy Colliders, Ann. Rev. Nucl. Part. Sci., Vol. 40(1990)529.

[9] J.E.Spencer, Accel. Diag. and Cont. by NN, 1989 IEEE Part. Accel. Conf., 89CH2669(1989)1642.

[10] *Exploring Complexity An Introduction*, Gregoire Nicolis and Ilya Prigogine, W.H.Freeman and Co., New York 1989. Also: N.H.Packard et al., Phys. Rev. Lett., 45(1980)712 and J.D.Farmer and J.J.Sidorowich, ibid, 59(1987)845. We use ANN's to define models and to optimize. Parts of this paper were given in BAPS, Vol.35(1990)1047.

[11] J.E.Spencer, High Brightness Sources, This Conf.

[12] *Adaptive Signal Processing*, Bernard Widrow and S.D.Stearns, Prentice-Hall, NJ 1985.

DISCLAIMER