Rodrick Wallace

Cognitive Dynamics on Clausewitz Landscapes

The Control and Directed Evolution of Organized Conflict

Springer

Cognitive Dynamics on Clausewitz Landscapes

Rodrick Wallace

Cognitive Dynamics on Clausewitz Landscapes

The Control and Directed Evolution of Organized Conflict

Springer

Rodrick Wallace
The New York State Psychiatric Institute
New York, NY, USA

ISBN 978-3-030-26423-9 ISBN 978-3-030-26424-6 (eBook)
https://doi.org/10.1007/978-3-030-26424-6

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Organized conflict—overt and covert—has long been of central concern to those entrusted with the maintenance of public health and public order. "Collateral damage" associated with, and following, large-scale conflict is an avalanche of death and social disintegration that entrains vast populations without consideration of combatant status, permanently distorting communities and their historical trajectories. Consequently, even the best-intentioned instances often rewind themselves, leading to repeated outbreaks in a kind of "punctuated equilibrium" that may be more familiar to evolutionary theorists than military strategists and their political overlords.

Klinger [1] ends her recent study of social science and national security with these words:

> In the final analysis, the strategy failure in Vietnam reflects the unfortunate confluence of problems with theory and practice as the strategy was implemented. Failure, in this case, illustrates the consequences of flawed strategic judgment. David Kaiser cites a letter written by William Bundy to American ambassadors in South Korea, Laos, and Japan dated June 1965. In it, Bundy listed three unpleasant choices for strategy in Vietnam:
>
> 1. Expand the bombing campaign and risk Chinese intervention.
> 2. Mine Haiphong Harbor.
> 3. Deploy more ground troops to raise the total to 300,000.
>
> The letter then noted that not one of these options would raise the chance of success much above 30 percent. That the United States pursued the war against such odds speaks not only to flawed theory with a misplaced sense of its scientific basis, but also to the folly of strategic judgment.

More than half a century later, the US remains deeply engaged in what has now become the Forever War, relentlessly directed against other non-Western populations and employing similar follies of strategic judgment.

Those of us who, as a kind of Moorsoldaten, must follow along behind the various armies with our shovels to bury the dead, would like to see some significant shift in this recurrent pattern. A deeper comprehension of the dynamics of organized conflict under imprecision, uncertainty and time constraint, might perhaps contribute to such a shift, and is the aim of this monograph.

You can't stop what you don't understand.

There are, of course, many formal models of warfare. From the differential equations of Lanchester and Richardson to the psychopathic fantasies of Game Theory, elaborate simulations are played out that, it can be claimed, express significant aspects of war. Western military thinkers, however, revere the works of the Prussian tactician and strategist Carl von Clausewitz, who insisted that uncertainties of intelligence and situational awareness will interact with "unanticipated difficulties" to dominate real-world battles and campaigns.

The work here will focus on such "Clausewitz Landscapes", most singularly characterized by "fog-of-war" and "friction", using formal theories of cognition and control acting under the influence of selection pressure representing the punctuated equilibrium of Lamarckian evolution: armies learn, and incorporate that learning into their doctrines.

The resulting mathematical structures, however, are of necessarily limited value. Bracken put the underlying problem thus [2]:

> …[T]here seem to be two kinds of people in the world: those who build mathematical models, and those who focus on the world… The modeler gains status by impressing other modelers and giving talks at professional societies. Those who focus on the world usually don't go to such meetings. They play to an audience of what's actually taking place on the battlefield, whether it is located in a foreign land or in a corporate board room.

The theoretical ecologist E. C. Pielou describes the conundrum in a slightly different manner [3], finding that the principal use of formal models in the study of complex ecologies is the important task of raising questions, not answering them. Models can suggest experimental and observational studies that are the only real sources of new knowledge as opposed to new speculation. The work here is such speculation. That being said, probability models can often be used to derive statistical tools—analogous to regression equations—that can, within constraints, serve to both analyze real-world data, uncovering hidden underlying mechanisms, and make limited predictions for policy planning.

For the models explored here, such a program remains to be addressed, and is both far from trivial and multifactorial. It is hard enough to transform probability models into reliable statistical tools for data analysis, but it is perhaps harder to both explore realms of applicability of the resulting tools, and to impose best-practice limits on their usage. The "natural" impulse, once a statistical model has been officially validated, is to manipulate both the model and the available data to achieve a politically predetermined outcome. A canonical example will be given in Chap. 11, where the infamous Rand Fire Project condensed well-known and highly punctuated traffic flow phase transitions down onto a simple travel time versus distance regression model for the full Trenton, NJ road network. A similar approach in New York City was literally devastating [e.g., 4].

Chapter 1 examines tactical time frames in the spirit of John Boyd's OODA loop, taking the perspective of the Data Rate Theorem that mandates a minimum rate of control information for successful regulation of an inherently unstable system. Strategic time frames are, by contrast, studied using a Lamarckian variant of the Eldredge/Gould vision of "punctuated equilibrium" in evolutionary theory. On this scale, moment-to-moment tactical "muddling through" just does not suffice.

Chapter 2 studies ways in which doctrine and groupthink can hobble institutional cognition in real-time conflict, using tools adapted from the theory of anytime algorithms that allow "good enough" solutions to problems under time constraints, and two surprisingly direct approaches using models based on stochastic differential equations. The work explicitly outlines how groupthink can greatly amplify fog-of-war burdens. Conversely, an unexpected implication is that sufficient fog-of-war may cripple AI systems that rely on anytime algorithms.

Chapters 1 and 2 are based on, but greatly expand and reinterpret, recently published peer-reviewed material [5, 6].

Chapter 3 focuses on asymmetric conflict in which one side is rich in material resources, and the other in information resources. The results, from the perspectives of the effectiveness and efficiency of institutional cognition, illuminate why the gods are not always on the side of the big battalions.

Chapter 4 examines the failure of institutional cognition under stress. Combat conditions often trigger a canonical and pathologically stable "ground state" condensation in which all possible targets become identified as enemies, i.e., the familiar dynamic of "Kill everyone and let God sort them out". At the other end of the spectrum is the collapse into denial-of-reality characteristic of MacArthur after Inchon or current right-wing attitudes on global climate change and factory-farm-induced mass-fatal pandemics.

Chapter 5 explores a heuristic "East Asian" model of extended conflict in which the aim is to "tune the channel" about the central "message" one is attempting to impose on an adversary. Their argument is that strategy is about defining the context—the "riverbanks"—along which conflict flows/evolves. This is contrasted with Western strategic perspectives, suggesting the possibility of "third stream" approaches to strategy less constrained by cultural "riverbanks", difficult as this would be to instantiate under real-world politics.

Chapter 6 reexamines the role of doctrine in driving the distortion between intent and effect on a Clausewitz landscape, in a sense taking the perspective of Chap. 5 to reconsider the dynamics of Chap. 2.

Chapter 7 uses the perspectives of the earlier developments to examine recent US security doctrine regarding the "resilience" of essential institutions and networks in the face of serious perturbation. While US doctrine seeks a rapid return to a more-or-less normal version of the pre-perturbation state—what can be characterized as "engineering resilience"—sufficient disturbance will set complex socio-technical networks on a new path-dependent evolutionary trajectory that will greatly resemble the pathological "ground state" of a eutrophied ecosystem. Again, disjunctions between doctrine and reality become manifest.

Chapter 8 studies the induction of culturally specific "emotional dysfunction" on one contending agent by another. Cultural blind spots, it appears, almost always provide an opportunity for exploitation of institutional cognition under severe time constraint. Such constraint, in fact, mandates the existence of rapid, culturally patterned responses by cognitive agents to specific incoming signals whose cultural blind spots can be exploited.

Chapter 9 applies a "Cambrian Explosion" formalism to Lamarckian evolutionary process on a Clausewitz Landscape, exploring mechanisms by which community fragmentation following a military "victory" by invading forces can lead to an avalanche of social disintegration triggering hyperviolent interactions. Hyperviolence is, unfortunately, one of the few ways available to punch through a "message" along a disintegrating social network. Hyperviolence will, of course, further degrade social structures, leading to a positive feedback producing the too familiar "cycles of violence".

Chapter 10 collapses the engineering curve implied by the earlier set of models to study a minimal structure describing cognitive dynamics on a Clausewitz landscape dominated by friction and fog-of-war, i.e., the Western cultural model of armed conflict. The results are not encouraging.

Chapter 11, summarizing and extending work conducted under an Investigator Award in Health Policy Research from the Robert Wood Johnson Foundation [4], presents a case history of an operations research model of a paramilitary system that caused a catastrophe by ignoring phase transitions in critical underlying dynamic structures. The failure of the Rand Corporation fire service models in New York City may have materially contributed to as many as 100,000 premature fatalities over a 30-year period. New York's status as the apex of the US urban hierarchy means that much of that harvest of death could not be confined to the city.

Chapter 12 is another sort of case history, an example of a policy recommendation constructed by combining bits and pieces from the earlier essays. It evaluates the civilian development of emotionally sentient artificial intelligence for "dual use" potential, which appears to be quite considerable.

Chapter 13, Final Remarks, argues that overt military combat has become obsolete as the routine tool-of-choice in strategic conflict, suggesting that, inverting the famous Clausewitz dictum, in the future, skilled political maneuver and negotiation may perhaps come to be seen as warfare by other means.

The Mathematical Appendix, Chap. 14, should be at least perused. Control over the evolution of conflict between cognitive agents under uncertainty involves matters that fully challenge current theoretical capabilities.

The nature of the volume, as a set of related but largely independent essays, mandates a certain amount of tutorial repetition among the chapters. In addition, the reference lists are by no means complete, and represent both the limitations and perspectives of the author, as do the particular formal methods chosen.

New York, USA Rodrick Wallace

References

1. Klinger, J. 2019. *Social Science and National Security Policy: Deterrence, Coercion, and Modernization Theories*. New York: Springer.
2. Bracken, P. Spring 2006. Net assessment: A practical guide. *Parameters* 90–100.
3. Pielou, E.C. 1977. *Mathematical Ecology*. New York: Wiley.
4. Wallace, D., and R. Wallace. 1998. *A Plague on Your Houses: How New York was Burned Down and National Public Health Crumbled*. New York: Verso.
5. Wallace R. 2019. Cognition, conflict, and doctrine: How groupthink fails on a Clausewitz landscape. *Journal of Defense Modeling and Simulation.* https://doi.org/10.1177/1548512919843642.
6. Wallace, R. 2019. Contrasting tactical and strategic dynamics on a Clausewitz landscape. *Journal of Defense Modeling and Simulation.* https://doi.org/10.1177/154851291984868.

Contents

1 Contrasting Tactical and Strategic Dynamics 1
1.1 Summary 1
1.2 Introduction 1
1.3 Tactics: The World of John Boyd 2
1.4 Strategy: The World of Eldredge and Gould 11
1.5 Discussion 18
1.6 Chapter Appendix: Uncertainty in τ 20
References 21

2 Doctrine and the Fog-of-War 23
2.1 Summary 23
2.2 Introduction 23
2.3 A Complex Model 24
2.4 Another Approach 32
2.5 Failure of Intelligence on a Cognitively Stressed Operational Network 32
2.6 Discussion 34
References 35

3 On Asymmetric Conflict 37
3.1 Summary 37
3.2 Introduction 37
3.3 Institutional Cognition in Conflict 39
3.4 An Operational Level Model 44
3.5 Iterating the Basic Idea 45
3.6 Discussion 47
3.7 Chapter Appendix: A Time-Constrained Operational Model 48
References 50

4 The Albigensian Ground State 51
4.1 Introduction 51
4.2 The Model 51
4.3 Discussion 55
References 55

5 Can There Be "Third Stream" Doctrines? 57
5.1 Introduction 57
5.2 The Analytic and Holistic Streams of Military Doctrine 60
5.3 Third Streams 64
5.4 Chapter Appendix: Reconsidering Arms Races 66
References 69

6 Reconsidering Doctrine and Its Discontents 71
6.1 Introduction 71
6.2 The Model 73
6.3 Discussion 74
References 75

7 Challenges to the US Security Doctrine of "Resilience" 77
7.1 Introduction 77
7.2 The Argument 77
7.3 Punctuated Equilibrium Redux 80
7.4 Discussion and Conclusions 85
7.5 Chapter Appendix: A Simplified Model of Phase Transition 88
References 89

8 Culture and the Induction of Emotional Dysfunction on a Clausewitz Landscape 91
8.1 Introduction 91
8.2 An Evolutionary Perspective on Emotion 92
8.3 Generalizing Anytime Algorithms 92
8.4 Discussion 97
References 98

9 Expected Unexpecteds: Cambrian Explosions in Lamarckian Systems 99
9.1 Introduction 99
9.2 Some Hard-Core Evolutionary Theory 100
9.3 The Lamarckian Variant 106
9.4 Chapter Appendix: "Simple" Lamarckian Phase Transitions 108
References 110

10 Reconsidering Clausewitz Landscape Dynamics 113
10.1 Introduction 113
10.2 The Models 113
10.3 Discussion 118
References 119

11 Failure of a Paramilitary System: A Case History of Catastrophe 121
11.1 Introduction 121
11.2 The Rand Fire Models 121
11.3 The Catastrophe 125
References 131

12 An Emerging Catastrophe: The Weaponization of Emotional Sentience in AI 133
12.1 A Brief Preface 133
12.2 Summary 134
12.3 Introduction 134
12.4 Revisiting an Evolutionary Perspective on Emotion 135
12.5 Anytime Algorithms Redux 136
12.6 Generalized Anytime Algorithms I 139
12.7 Generalized Anytime Algorithms II 142
12.8 Discussion 146
12.9 Mathematical Appendix 147
References 148

13 Final Remarks 151
References 154

14 Mathematical Appendix 157
14.1 Groupoids 157
14.2 Morse Theory 158
14.3 An RDT Proof of the DRT 160
14.4 An Information Black–Scholes Model 162
14.5 Symmetry and Information 163
References 165

About the Author

Rodrick Wallace is a research scientist in the Division of Epidemiology at the New York State Psychiatric Institute, affiliated with Columbia University's Department of Psychiatry. He has an undergraduate degree in Mathematics and a Ph.D. in Physics from Columbia, and completed a postdoctoral fellowship in the Epidemiology of Mental Disorders at Rutgers. He has worked as a public interest lobbyist, including two decades conducting empirical studies of fire service deployment, and received an Investigator Award in Health Policy Research from the Robert Wood Johnson Foundation. In addition to numerous works on public health and public policy, he has published a dozen peer-reviewed papers and chapters modeling evolutionary process, and many formal studies of human, institutional, and machine cognition.

Chapter 1
Contrasting Tactical and Strategic Dynamics

1.1 Summary

The asymptotic limit theorems of control and information theories allow formal examination of the fundamental differences between short-term tactical confrontations, dominated by relatively tractable engineering concerns, and extraordinarily subtle long-term "strategic" conflict dominated by punctuated Lamarckian evolutionary processes. The world of extended coevolutionary conflict is not the world of sequential "muddling through". The existential strategic challenge is to control long-term dynamics in which one maybe "losing" most short-term confrontations. Winning individual battles involves relatively direct, if not simple or easy, matters of sufficient local resources, training, and resolve. Winning extended conflicts is indirect, requiring management of coevolutionary phenomena subject to dismaying dynamics of punctuated equilibrium that are more familiar to evolutionary theory than to military doctrine. Directed evolution—"farming"—has given us the agricultural base needed for large-scale human organization. Directed coevolution of the inevitable conflicts between the various segments of such organization may be needed for our long-term persistence.

1.2 Introduction

Strategy and tactics act on different scales of time, space, and organization, but also face qualitatively different dynamic constraints and, perhaps loosely, "laws of motion". It is possible, as the Wehrmacht found on the Eastern Front in World War II, to retain significant tactical superiority during most confrontations, and still lose the war [1]. This observation is important, given current U.S. superiority in equipment, training, and resolve at the tactical level of military enterprise.

Krepinevich and Watts [2, 3] describe the current U.S. dilemma in these terms:

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_1

> The ability of the US national security establishment to craft, implement, and adapt effective long-term strategies against intelligent adversaries at acceptable costs have been declining...
>
> ...Strategy may be a game anyone can play, but the evidence is strong that very few can play it well.
>
> ...[T]he U.S. government not only has lost the ability to do strategy well, but... many senior officials do not understand what strategy is.

After January 2017, of course, the situation became dire.

Here, we develop two markedly different formal models of tactical and strategic conflict dynamics. These are based on the asymptotic limit theorems of control and information theories, and link tactics and strategy in a way that makes clear both their relationship and differences.

Statistical models based on probability models, for example, regression equations, can provide benchmarks against which real data can be compared. But, in the spirit of the mathematical ecologist Pielou [4], models should be used for indicating possible directions for observational or empirical study, since these are the only real bases for new knowledge as opposed to new speculation. What this essay presents are "models about", not "models of". In other terms, "the word is not the thing". Nonetheless, regression equations are also statistical models, and those fitted to data can, with some care, be used to make sense of empirical observations and provide limited tools for prediction.

1.3 Tactics: The World of John Boyd

John Boyd, the architect of the famous U.S. Left Hook in the First Gulf War, attempted to generalize his experiences of the thin, warm fog-of-war which embedded the tactics of mid-twentieth-century air combat. Boyd understood and explored something of the intimate relations between cognition and control basic to conflict. Such relations differ, of course, according to scale and level of organization. Boyd's understanding, explored and developed in a long series of lectures, evolved from a simple sequential dynamic of Observation, Orientation, Decision and Action (OODA) loop, to the elaborate model of cognition and control of his later presentations. Osinga [5], one of the foremost experts on Boyd's thinking, describes the later work as follows:

> The OODA loop model as [ultimately] presented by Boyd... represents his view on the process of individual and organizational adaptation in general, rather than only the military-specific command and control decision-making process that it is generally understood to depict. It refers to [a] conceptual spiral... to the process of learning, to doctrine development, to command and control processes and to the Popperian/Kuhnian ideas of scientific advance. The (neo-)Darwinists have their place, as do Piaget, Conant, Monod, Polanyi and Hall, while Prigogine and Goodwin are incorporated through Boyd's concluding statement in the final slide that follows [his presentation of] the OODA loop picture:
> "The key statements of this presentation, the OODA Loop Sketch and related insights represent an evolving, open-ended, far from equilibrium process of self-organization, emergence and natural selection."

> This relates the OODA loop clearly to Complex Adaptive Systems, the role of schemata and to the process of evolution and adaptation. Once again it shows that where the aim is "to survive and prosper" in a nonlinear world dominated by change, novelty, and uncertainty, adaptation is the important overarching theme in Boyd's strategic theory.

Figure 1.1, from Osinga's book [5], adapts Boyd's presentation, and displays a fully developed cognitive/control process model that seems particularly relevant to tactical and operational levels of conflict, where more tractable "engineering" problems dominate.

It is possible to examine Fig. 1.1 in a way that links control and information theories, focusing, at first, on tactical and operational levels. Extending these ideas to strategic scales, where coevolutionary dynamics may predominate, requires very different approaches, to understate the matter.

A—in a way, the—central point is that contending cognitive systems, forced to act in real time and under fog-of-war conditions, are inherently unstable, similar to a vehicle driven at night on a twisting, pot-holed roadway. That vehicle needs, in addition to an excellent driver, good headlights, and highly responsive steering.

The Data Rate Theorem (DRT) [6] is an extension of the Bode Integral Theorem that determines the minimum rate at which externally supplied control information must be provided for an inherently unstable control system to be stabilized and to remain stable.

The standard first approximation makes a linear expansion near a nonequilibrium steady state. An n-dimensional vector of system parameters at time t, which we represent as x_t, determines the state at time $t + 1$ according to the relation

$$x_{t+1} = \mathbf{A}x_t + \mathbf{B}u_t + W_t \tag{1.1}$$

$\mathbf{A}$ and $\mathbf{B}$ are taken as fixed $n \times n$ matrices. u_t is the vector of control information at time t, and W_t is an n-dimensional vector of Brownian "white noise".

Figure 1.2 is a control system schematic that can be compared with Boyd's model. That figure shows, projected down onto an irreducible minimum, the central features of any command and control structure in the presence of "white noise", i.e., the unanticipated difficulties and occurrences that are most often modeled (as we do here) as undifferentiated, but can become "colored" by internal structure in more complicated models. The dotted line in Fig. 1.2 represents a collapsed sum of the three rightmost "feedback" circuits of Boyd's OODA loop of Fig. 1.1. The other loops of Fig. 1.1 can be similarly collapsed into Fig. 1.2.

The Data Rate Theorem asserts that the minimum rate at which control information $\mathcal{H}$ must be provided for system stability is determined as

$$\mathcal{H} > \log[|\det[\mathbf{A}^m]|] \equiv a_0 \tag{1.2}$$

Taking $m \leq n$, then $\mathbf{A}^m$ is the subcomponent of $\mathbf{A}$ having eigenvalues ≥ 1. The right-hand side of Eq. (1.2) is then characterized as the rate at which the system generates topological information.

Stability fails if the inequality of Eq. (1.2) is violated.

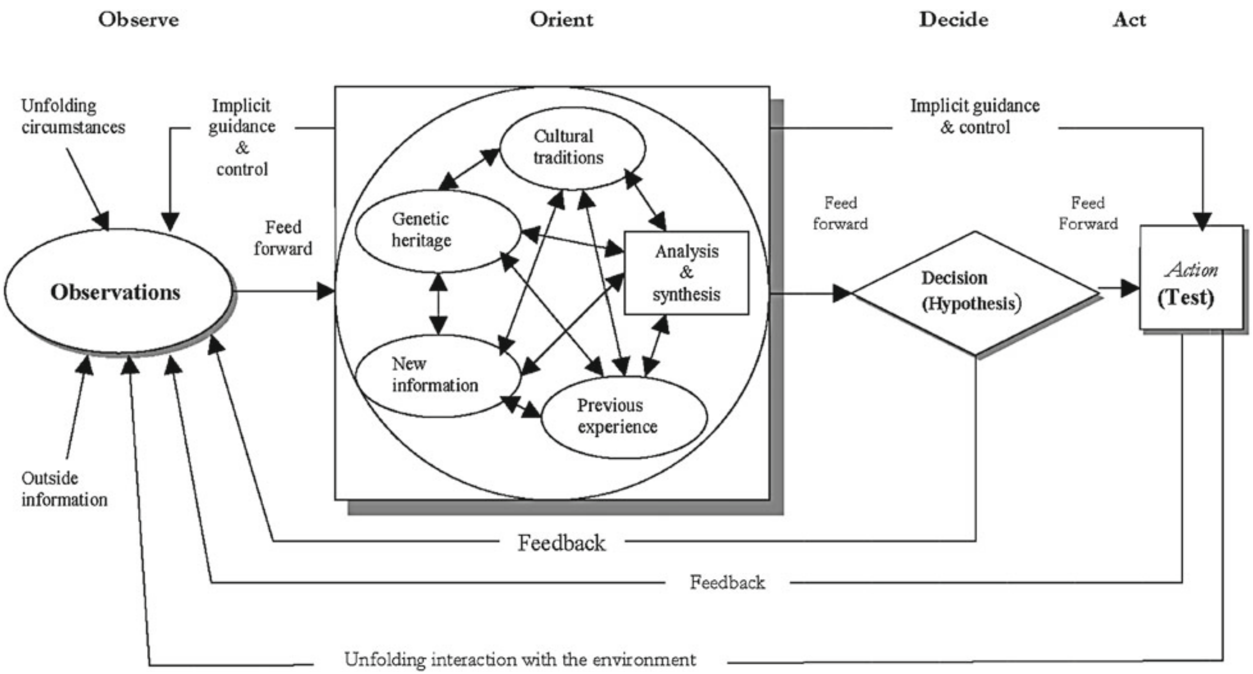


Fig. 1.1 From [5]. John Boyd's fully developed ideas convolute cognition with control, under the conditions of uncertainty inherent to Clausewitz landscapes strongly dominated by friction, the fog-of-war, and resource limitations

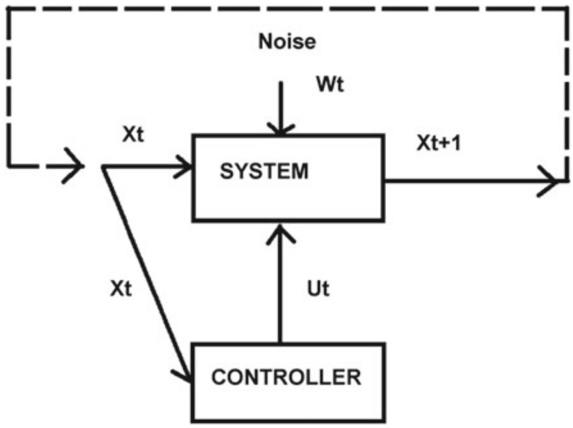


Fig. 1.2 Model of a control system under the constraints of the Data Rate Theorem. Compare with Boyd's model of Fig. 1.1

For the night driving example, if the headlights go out, a twisting roadway cannot be navigated. A different approach to the Data Rate Theorem is given in the Mathematical Appendix, using the Rate– Distortion Function that varies according to the nature of the control channel ([7], Sect. 12.5). The proof centers on the inherent convexity of all Rate–Distortion Functions.

Equations (1.1) and (1.2) both illustrate Boyd's model of a continuous cycle of interaction with the environment, assessing and responding to constant change. Boyd's fundamental claim is that victory in combat is assured by forcing circumstances to change more rapidly than an adversary can respond. According to the Data Rate Theorem analysis here, such challenge triggers fatal destabilization by making the rate at which topological information is generated greater than the rate at which the adversary can possibly provide control information.

At the tactical level, no cognitive system is immune to such an attack. However, on larger time, space, and resource scales—at the strategic level—an adversary can conventionally lose every tactical confrontation and still win the war.

Clausewitz's "fog-of-war" involves, among other things, the inevitability of limited intelligence regarding battlefield conditions, and his "friction" refers to the inevitable difficulty of imposing control on battle, due to weather, terrain, time lags, attrition, problems in resupply and logistics, and so on. For night driving on bad roads, this might involve a synergism between poor headlights and unresponsive steering, or, for that matter, actual fog and a muddy road that no vehicle can navigate.

Real-time cognitive systems engaged in conflict will usually have many such constraints acting simultaneously and synergistically. These circumstances include and incorporate resource limitations, perhaps expressed in time lags for resupply, and hence time lags in response to adversarial initiatives. We take these constraints as represented by a nonsymmetric $n \times n$ correlation-analog matrix ρ with elements $\rho_{i,j}$, characterizing those constraints and their pattern of interaction.

Such matrices will have n "invariants", r_i, $i = 1...n$, that remain fixed when transformations similar to rotation are applied to data, and we next construct an invariant scalar measure based on the standard polynomial relation

$$p(\gamma) = \det(\rho - \gamma I) = \\ \gamma^n + r_1\gamma^{n-1} + ...r_{n-1}\gamma + r_n \tag{1.3}$$

Here, det is the determinant, γ a parameter, and I is the $n \times n$ identity matrix. The first invariant is the matrix trace, and the last $\pm$ the determinant.

Based on the n invariants, we can define a composite scalar index $\Gamma = \Gamma(r_1, ..., r_n)$ as a monotonic increasing real function, a projection like the Rate–Distortion Manifold of [8], and the Generalized Retina of [9], or to the more familiar statistical technique of the principal component analysis.

The central extension of the condition of Eq. (1.2) is then

$$\mathscr{H}(\Gamma) > f(\Gamma)a_0 \tag{1.4}$$

An adaptation of the standard Black–Scholes approximation from mathematical finance ([7], Sect. 7.10) finds, in first order, that $\mathscr{H}(\Gamma)$ has a linear form, i.e., $\mathscr{H} \approx \kappa_1\Gamma + \kappa_2$. See Sect. 14.4 of the Mathematical Appendix for details.

Expanding $f(\Gamma)$ to first order, so that $f(\Gamma) = \kappa_3\Gamma + \kappa_4$, the limit condition is then

$$\mathscr{T} \equiv \frac{\kappa_1\Gamma + \kappa_2}{\kappa_3\Gamma + \kappa_4} > a_0 \tag{1.5}$$

where we will call $\mathscr{T}$ the "Clausewitz temperature" of the system.

Again, as constructed, this projects down fog-of-war, frictional, and resource constraints onto a single scalar measure.

For $\Gamma = 0$ the stability condition is $\kappa_2/\kappa_4 > a_0$. At large Γ, this condition becomes $\kappa_1/\kappa_3 > a_0$. If, however, $\kappa_2/\kappa_4 \gg \kappa_1/\kappa_3$, then the stability condition may be violated at high Γ, illustrated in Fig. 1.3.

$\mathscr{T}$, as we have constructed it from Eq. (1.3), can be seen as representing something of the "known unknowns" of a Clausewitz landscape.

It is possible to extend the argument to the expression of $\mathscr{T}$ in the mean distortion between sent orders and battlefield outcome, given fog-of-war uncertainties, but absent punctuated collapse. Such distortion decreases monotonically with increase in $\mathscr{T}$, according to a somewhat subtle argument that, surprisingly, recovers the Data Rate Theorem. The details are as follows.

How is a control signal u_t in Fig. 1.2 expressed in the system response x_{t+1}? In standard Rate–Distortion Theorem manner [10], we deterministically retranslate an observed sequence of system outputs $X^i = x_1^i, x_2^i, ...$ into a sequence of possible control signals $\hat{U}^i = \hat{u}_0^i, \hat{u}_1^i, ...$ and to compare that sequence with the original control sequence $U^i = u_0^i, u_1^i, ...$, with the difference between them having a particular value under some chosen distortion measure and hence having an average distortion

$$D \equiv \sum_i p(U^i)d(U^i, \hat{U}^i) \tag{1.6}$$

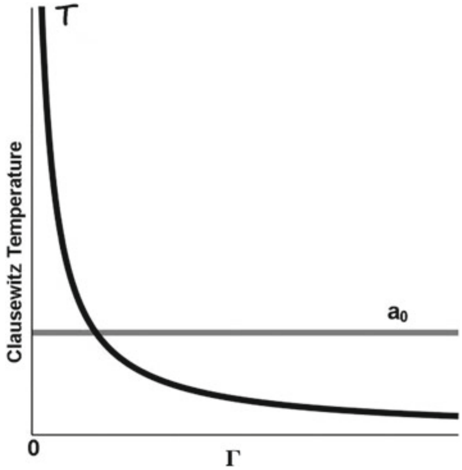


Fig. 1.3 The horizontal line is the Data Rate Theorem limit a_0. If $\kappa_2/\kappa_4 \gg \kappa_1/\kappa_2$, then, at an intermediate value of Γ, the Clausewitz temperature $\mathscr{T}$ falls below the critical value, and control fails. From Eq. (1.3), $\mathscr{T}$ has been constructed from the "known unknowns" of a Clausewitz landscape

$p(U^i)$ is the probability of the sequence U^i. $d(U^i, \hat{U}^i)$ is the distortion between U^i and the sequence of control signals that has been deterministically reconstructed from system output.

It is then possible to apply a Rate–Distortion argument. The Rate–Distortion Theorem asserts that there is a Rate–Distortion Function that determines the minimum channel capacity—$R(D)$—necessary to keep the average distortion below some fixed limit D. Again, see [10] for details.

The simplest approach is to construct a Boltzmann-like pseudoprobability in the Clausewitz temperature $\mathscr{T}$ as

$$dP(R, \mathscr{T}) = \frac{\exp[-R/g(\mathscr{T})]dR}{\int_0^\infty \exp[-R/g(\mathscr{T})]dR} \tag{1.7}$$

where the function $g(\mathscr{T})$ is to be determined. It must, however, be positive and monotonically increasing in $\mathscr{T}$.

The integral in the denominator has the simple value $g(\mathscr{T})$.

For a Gaussian channel, the Rate–Distortion Function is $R(D) = (1/2) \log(\sigma^2/D)$, so that $D = \sigma^2 \exp[-2R]$.

From Eq. (1.7) it is possible to calculate $< D >$ as

$$< D >= \frac{\int_0^\infty \sigma^2 \exp[-2R] \exp[-R/g(\mathscr{T})]dR}{\int_0^\infty \exp[-R/g(\mathscr{T})]dR} = \frac{\sigma^2}{2g(\mathscr{T}) + 1} \tag{1.8}$$

Above the critical threshold value, increasing $\mathscr{T}$ monotonically decreases the effect of σ^2 on the average distortion between what is ordered and what is implemented.

We can actually determine a critical value for $\mathscr{T}$ by explicitly calculating the function $g(\mathscr{T})$. This can be done by recognizing the denominator of Eq. (1.7) as a statistical mechanical partition function defining a free energy analog as

$$\exp[-F/g(\mathscr{T})] = \int_0^\infty \exp[-R/g(\mathscr{T})]dr = g(\mathscr{T})$$
$$F = -g(\mathscr{T})\log[g(\mathscr{T})] \tag{1.9}$$

Then an entropy-analog can be defined as

$$S(\mathscr{T}) = F(\mathscr{T}) - \mathscr{T} dF/d\mathscr{T} \tag{1.10}$$

We take the perspective of the Onsager treatment of nonequilibrium thermodynamics [11] to make a first-order approximation for the dynamic behavior of $\mathscr{T}$ as

$$d\mathscr{T}/dt \propto dS/d\mathscr{T} = \frac{\left(g\left(\mathscr{T}\right)\left(\ln\left(g\left(\mathscr{T}\right)\right)+1\right)\frac{\mathrm{d}^2}{\mathrm{d}\mathscr{T}^2}g\left(\mathscr{T}\right)+\left(\frac{\mathrm{d}}{\mathrm{d}\mathscr{T}}g\left(\mathscr{T}\right)\right)^2\right)\mathscr{T}}{g\left(\mathscr{T}\right)} \tag{1.11}$$

Imposing a nonequilibrium steady state, so that $d\mathscr{T}/dt \to 0$, Eq. (1.11) can be solved for $g(\mathscr{T})$:

$$g(\mathscr{T}) = \frac{C_1\mathscr{T} + C_2}{W(n, C_1\mathscr{T} + C_2)} \tag{1.12}$$

$W(n, x)$ is the Lambert W-function of branch n-integer that solves the equation $x = W(n, x)\exp[W(n, x)]$.

The appearance of a Lambert W-function is often a red flag for some deeply underlying network dynamics [12].

For $g(\mathscr{T})$ to be real, positive, and monotonically increasing—the boundary conditions—then $n = 0$ and $C_1\mathscr{T} + C_2 > \exp[-1]$. This requires C_2 to be negative, with $|C_2| > \exp[-1]$, and we recover the necessary condition on $\mathscr{T}$ of Eq. (1.5) as

$$C_2 < 0$$
$$\mathscr{T} > \frac{|C_2| - \exp[-1]}{C_1} \equiv a_0 > 0 \tag{1.13}$$

The forms of $g(\mathscr{T})$ and the average $< D >$ are shown in Fig. 1.4.

Fog-of-war constraints may entail added stochastic burdens imposed on the "ordinary" difficulties of conflict. These burdens might be taken as the "unknown

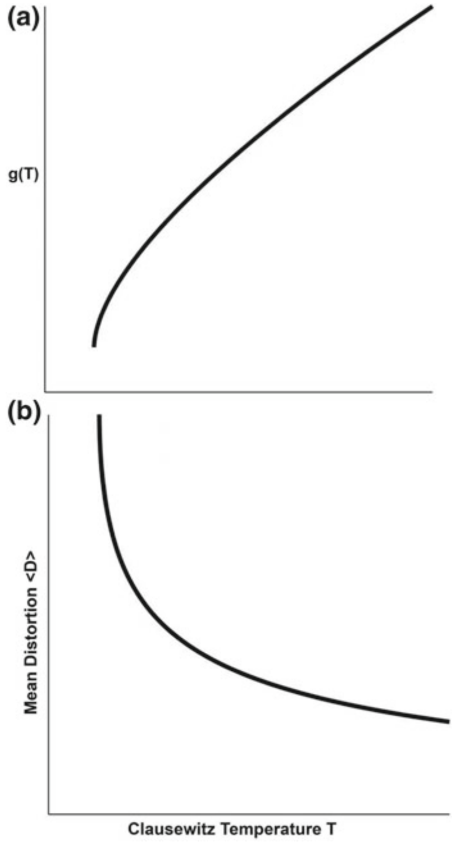


Fig. 1.4 **a** Form of the function $g(\mathcal{T})$ under the boundary conditions. **b** Mean average distortion $< D >$ as a function of $\mathcal{T}$ for a Gaussian channel model of the OODA loop of Fig. 1.2 at a fixed value of σ^2. Above the "catastrophe threshold" of Eq. (1.5), $< D >$ declines monotonically with increasing $\mathcal{T}$

unknowns" particular to Clausewitz landscapes. The assertion, then, is that, in terms of the fog-of-war, John Boyd's OODA loop is itself prone to uncertainties.

If the denominator in Eq. (1.7) can be treated with an integral approximation, then the analysis leading to Fig. 1.4 can be extended.

For a set of simultaneous complex situations that must be addressed at a single time by a command enterprise, or, in an "ergodic" sense, for a single situation that changes over time, the value of the Rate– Distortion Function, R, is itself subject to stochastic variation. We assign a probability distribution for values of R as a Gamma distribution with parameters k, θ having a density function

$$P(R, k, \theta) = \frac{R^{k-1} \exp[-R/\theta]}{\theta^k \Gamma(k)} \tag{1.14}$$

with $\theta, k > 0$ real.

The average of R across the distribution is $\hat{R} = k\theta$ and we redefine "free energy" F using the relation

$$\int_0^\infty \exp[-R/G(\mathcal{T})]P(R,k,\theta)dR = \exp[-F/G(\mathcal{T})] = \left(\frac{G(\mathcal{T})k}{G(\mathcal{T})k + \hat{R}}\right)^k \tag{1.15}$$

An entropy-analog is again defined as

$$S(\mathcal{T}) \equiv F(\mathcal{T}) - \mathcal{T} dF(\mathcal{T})/d\mathcal{T} \tag{1.16}$$

leading again to the Onsager nonequilibrium steady state condition

$$dS/d\mathcal{T} = 0 \tag{1.17}$$

The resulting differential equation in $G(\mathcal{T})$ can be solved to give the relation

$$Z \equiv -C_1\mathcal{T} + C_1$$
$$G(z) = \frac{Z}{W(-1, (kZ/\hat{R})\exp[kZ/\hat{R}]) - kZ/\hat{R}} \tag{1.18}$$

C_1 and C_2 are positive constants and $W(-1, x)$ is the Lambert W-function of order -1, real and finite only in the range

$$-\exp[-1] < (kZ/\hat{R})\exp[kZ/\hat{R}] < 0 \tag{1.19}$$

Thus, $G(\mathcal{T})$ can be real and finite only over a limited range of $\mathcal{T}$-values, as shown in Fig. 1.5, implying at least three different phases for the underlying system, in contrast to the "on/off" control theory analysis of the OODA loop using only the Data Rate Theorem.

The second example, that can be generalized using various "stochastic stabilization" theorems, is as follows.

First, suppose that $\mathcal{T}$ converges on some value $K > a_0$ according to the straightforward relation

$$d\mathcal{T}/dt = \mu\mathcal{T}(1 - \mathcal{T}/K) \tag{1.20}$$

When $d\mathcal{T}/dt = 0$, $\mathcal{T} = K > a_0$. Other functions converging on $K > a_0$ will give similar outcomes.

The most direct stochastic differential equation model that includes volatility is obviously

$$d\mathcal{T}_t = \mu\mathcal{T}_t(1 - \frac{\mathcal{T}_t}{K})dt + \sigma\mathcal{T}_t dW_t \tag{1.21}$$

where the term in σ represents volatility and dW_t is Brownian white noise.

Using the Ito Chain Rule [13] on $\log(\mathcal{T})$ shows that $\mathcal{T}$ converges in probability to zero if $\sigma^2/2 \geq \mu$, a standard outcome.

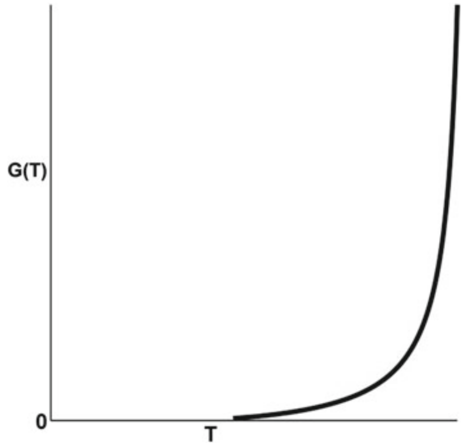


Fig. 1.5 Analysis for the "fog-of-war" OODA loop, assuming a Gamma probability distribution for values of the Rate–Distortion Function R, either as an average over a distribution of simultaneous command loops, or an "ergodic" time average over a single loop. $G(\mathcal{T})$ is real and finite only over a limited range, implying the existence of at least three different phases, for this model, in contrast with the on–off result for the "simple" control theory treatment of the OODA loop

Thus, and perhaps counterintuitively, a "strong" agent, facing what appear to be deterministically stable conflict dynamics, can be driven to extinction, given enough second-order fog-of-war "noise", σ.

A simulation, using the ItoProcess function of the computer algebra program Maple 2018, is shown in Fig. 1.6. There, $\mu = 1$, $K = 2$ and $\mathcal{T}$ at time zero is 0.3. σ is, respectively, 0.1 and 2.0. The first case satisfies the condition $\mu > \sigma^2/2$, and the system converges about $K = 2$, with some scatter. For $\sigma = 2$, example, the critical condition is violated, and the system collapses toward zero.

This series of arguments undercuts assertions that the OODA loop is an "end to history" in the study of armed conflict.

Boyd's model, expressed in control theory terms—extended by the arguments leading to Figs. 1.5 and 1.6—nonetheless seems to fit tactical, and, in some measure, operational, time frames. The strategic scale, as [2, 3], assert, appears dominated by quite a different set of mechanisms, by forms of coevolution subject to "punctuated equilibrium" dynamics only identified for natural populations by Eldredge and Gould in 1972 [14].

1.4 Strategy: The World of Eldredge and Gould

Punctuated equilibrium among contending agents is the idea that sets or systems of institutions that individually persist through some general forms of replication and reproduction, but are subject to Lamarckian or Darwinian selection pressures,

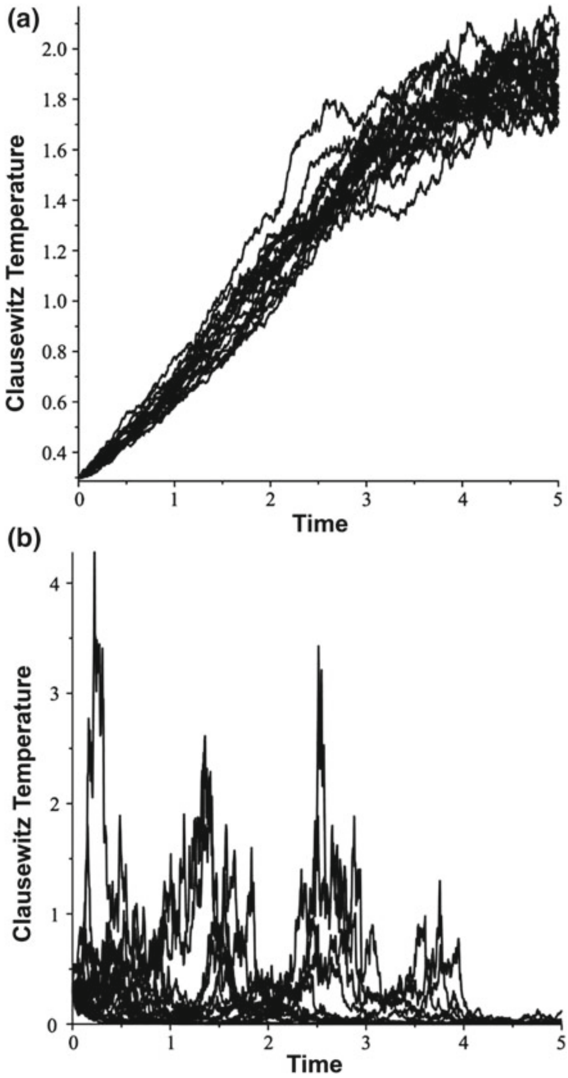


Fig. 1.6 Simulation of the relation $d\mathscr{T}_t = \mathscr{T}_t(1 - \mathscr{T}/2) + \sigma \mathscr{T}_t dW_t$, using the ItoProcess function of the computer algebra program Maple 2018. Twenty replications, 1000 timesteps, starting value $\mathscr{T}_0 = 0.3$. **a** $\sigma = 0.1$, consonant with the critical relation $\mu > \sigma^2/2$. **b** $\sigma = 2.0$, producing a barrage of "unknown unknowns" violating the critical condition

will often undergo long periods of apparent stasis while their phenotypes can adapt, before experiencing relatively rapid wholesale shifts in structure and behavior.

Among natural populations, the sudden shifts observed are usually extinction and/or speciation events in the fossil record. For cognitive institutions, which can learn and incorporate that learning into their means of reproduction, the dichotomy is usually extinction/structural change. Rogers [15], in 1993, applied the concept in a military setting, describing the processes of innovation associated with the use of cannon in the first Hundred Years' War. Others have followed his lead, applying the concepts to more contemporary situations (e.g., [16]).

A singularly perverse applications is by Leventoglu and Slantchev [17], who argue

> Our analysis suggests that one can usefully view war as a mutually coercive process that involves continuous fighting punctuated by occasional opportunities for peace.

We use the methodology to focus on the last level of the tactical–operational–strategic hierarchy, a more conventional perspective over a shorter time frame than studied by Rogers, but longer than that of a control theory analysis of tactical dynamics.

The Clausewitz Zweikampf between cognitive systems or entities, the "wrestling match", can be examined using coevolutionary theory, as Boyd implied and as has become a shibboleth in studies of asymmetric conflict (e.g., [18]). There are, however, certain twists on strategic, as opposed to tactical, scales, and levels of organization.

These ideas are neither conceptually nor mathematically trivial. While control theory is hard, evolutionary theory is even harder.

To begin, cognition, in conflict or not, most directly involves an agent's choice of an action, or a relatively small number of them, from a full repertoire of what is available [19]. Such choice permits the identification of a "dual" information source associated with cognition. More specifically, cognitive choice implies a reduction of uncertainty that, in turn, implies and requires the existence of some underlying information source characterizing that cognition.

We first study the set of cognitive modules within one contending entity. These modules might include, in addition to the hierarchy of direct combat units, similar hierarchies of supply, command, transport, and relief agents—the vast set of supporting enterprises needed to field combat units. These two compete within their embedding milieus for scarce resources. Thus, we see two highly networked entities that engage within themselves as well as between the larger structures. For each larger contending agent in the Zweikampf, we are forced to study matters across a vast network of internal contending and cooperating submodule information sources, a set $\mathscr{Y} \equiv \{Y_1, Y_2, ...\}$.

This implies the existence of a dense realm of interacting cognitive agents. Again, within a single side, under conditions of extended conflict, agents will both cooperate and compete for essential resources, including time.

Sequences of actions and behaviors of length $N \to \infty$ can, in general, be divided into a small high probability "typical" set that follows a characteristic "grammar" and "syntax", and a much larger set of paths having a vanishingly low probability that is not consistent with the underlying grammar and syntax [10, 20].

Coevolutionary dynamics, however, not only occur on a different timescale than tactical and operational actions, they may also require the invocation of "nonergodic" information sources, extending the arguments presented in Wallace ([21], Chap. 6). "Nonergodic" means that long-term averages do not converge on cross-sectional expectations, similar to "nonparametric" statistical models not based on the Central Limit Theorem. The source uncertainty of nonergodic systems, however, cannot be described in terms of the classic Shannon "entropy" [20], but regularities arising from Feynman's [22] identification of information as a form of free energy permit simplification of system dynamics via "phase transition" formalism familiar from statistical mechanics, based, however, on groupoid rather than group symmetry shifts.

The embedding natural or human ecosystem upon which conflict takes place, the underlying landscape in the ordinary sense of the term—land, sea, air, rural, urban, day, night, rain, snow, mud, etc.—has its own "grammar" and "syntax" defining what are high and low probability sequences of events, permitting representation as an information source, say X [21]. For example, seasonal transitions—from high summer to a rainy autumn choked in mud, to a hard-frozen surface that can support tracked vehicles—are not random, but have essential regularities in time and space.

Yet another fundamental information source arises from the dynamics of the "large deviations" to which contention is subject. Clausewitz, in fact, elaborates on the role of "probabilities" and the unexpected influence of random events. Champagnat et al. [23] describe how sudden transitions between nonequilibrium steady states also involve high probability paths whose description is in terms of the familiar Shannon uncertainty expression $\mathscr{I} = -\sum_k P_k \log(P_k)$, where the P_k form a probability distribution. Variations of this result appears repeatedly as Sanov's Theorem, Cramer's Theorem, the Gartner–Ellis Theorem, the Shannon–McMillan Theorem, and so on [24]. These results allow the characterization of large deviations in terms of yet another information source, L_D.

Study now focuses on a (not necessarily ergodic) joint information source uncertainty that we represent as $H(X, \mathscr{Y}, L_D)$.

This uncertainty is associated with one particular contending agent and is defined on a set of jointly typical [10] high probability paths z involving a sequence of states that conform to the "grammars" and "syntaxes" of the interacting information sources X, the set $\mathscr{Y}$, and L_D.

The argument then becomes concerned with the set of "jointly typical" high probability paths $z^n \equiv \{z_0, z_1, ..., z_n\} \rightarrow z$.

Khinchin [20] shows it is possible to define a path-dependent source uncertainty $H(z^n) \rightarrow H(z)$ that can vary across the manifold defined by the full set of paths z. $H(z)$, however, no longer has an "entropy like" mathematical expression, but its dynamics emerge via a familiar statistical mechanics argument.

We next invoke a Boltzmann pseudoprobability as

$$P[H(z_q)] \equiv \frac{\exp[-H(z_q)/\tau]}{\sum_z \exp[-H(z)/\tau]} \tag{1.22}$$

The sum (or some appropriate generalized integral) of the denominator is over all possible jointly typical paths z, and τ represents a generalization of the Clausewitz temperature $\mathscr{T}$ from Eq. (1.5) that we will later develop.

The pseudoprobability of Eq. (1.22) differs fundamentally from that of Eq. (1.7). R of Eq. (1.7) represents a Rate– Distortion Function associated with a single individual channel, while the $H(z_q)$ are uncertainty values associated with an enormously large set of individual paths. The dynamics associated with Eq. (1.22) will not be the same as those associated with Eq. (1.7), since these are different universes, in a sense.

The denominator of Eq. (1.22), however, can be used to define a "free energy" analog $\mathscr{F}$ as a Morse function [25, 26], via the relation

$$\exp[-\mathscr{F}/\tau] \equiv \sum_{z} \exp[-H(z)/\tau] \tag{1.23}$$

The Mathematical Appendix outlines something of Morse Theory.

The definition of a "strategic" τ may be subtle, even if $\mathscr{T}$ is kept within bounds during tactical confrontations. As Hannibal once put it, "Another such victory and we are undone..."

Or, as more recently said by bin Ladin [27],

> All that we have to do is send two mujahidin to the furthest point east to raise a piece of cloth on which is written al-Qaida, in order to make the generals race there to cause America to suffer human, economic, and political losses without their achieving for it anything of note... [Just] as we... bled Russia for 10 years, until it went bankrupt and was forced to withdraw in defeat... [s]o we are continuing this policy in bleeding America to the point of bankruptcy.

We could build τ as a scalar index from total or annual expenditures, casualties, opportunity costs, measures of degradation of the political capital needed at home and abroad for government function, and the time average of the tactical index $\mathscr{T}$, using an argument similar to that of Eq. (1.3). This leads to a new "temperature" measure analogous to Eq. (1.5), but focused on long-term, large-scale burdens and effects.

Given such a τ, in contrast to the relatively simple "on-off" dynamics of Eqs. (1.3)–(1.5) and those of the Data Rate Theorem, a supremely subtle phase transition model emerges on the strategic scale by identifying equivalence classes of a system's developmental pathways z, for example, "functional" and "doomed". A classification into equivalence classes allows the definition of a symmetry groupoid for the developmental process [7, 28, 29]. A groupoid is a generalization of an algebraic group in which a product is not necessarily defined between each element. The simplest example is, perhaps, a disjoint union of separate groups, but sets of equivalence classes also define a groupoid. See Weinstein [28] for details, and the Mathematical Appendix for an outline.

Given an information source associated with the system of interest, a full equivalence class algebra can then be constructed by choosing different system origin states and defining the equivalence of subsequent states at a later time by the existence of a high probability path connecting them to the same origin state. Disjoint partition by equivalence class, like orbit equivalence classes in dynamical systems, defines a symmetry groupoid associated with the cognitive process [28].

The equivalence classes across possible origin states then further define a set of information sources dual to different states available to the systems of interest. These create a large groupoid, with each orbit corresponding to an elementary "transitive" groupoid whose disjoint union is the full groupoid. Each subgroupoid is associated with its own dual information source, and larger groupoids must have richer dual information sources than smaller.

The "free energy" Morse Function of Eq. (1.23) is then liable to an analog of Landau's classical spontaneous symmetry breaking [25, 26]. Under symmetry breaking, higher "temperatures" are associated with more symmetric higher energy states in physical systems. Cosmology uses such arguments regarding the first moments after the "big bang", claiming that different physical phenomena broke out as the universe cooled through expansion.

In the circumstance of protracted conflict, decline in an agent's scalar index τ from Eq. (1.22) will trigger sharply punctuated collapse, from higher to lower symmetry states, causing ultimate system failure.

To reiterate, in this model, the Morse Function "free energy" of Eq. (1.23) can be driven into a sudden, highly punctuated phase transition reflecting fundamental symmetry shifts by changes in the strategic "temperature" parameter τ. The symmetry shift, however, is between groupoids associated with a synergism across contending institutional cognitions, environmental dynamics, and the structure of stochastic "large deviations". This is not the melting of ice, and does not generally occur on the timescale of tactical confrontation described by the "simple" John Boyd dynamics of the Data Rate Theorem.

Generalizations of τ and $\mathscr{F}$ are briefly examined in the Chapter Appendix.

The coevolutionary Clausewitz temperature parameter τ, applying to strategic rather than tactical time frames, may, however, be subject to stochastic dynamics as explored in Eqs. (1.20), (1.21) and Fig. 1.6—the "unknown unknowns". But system response may be more complicated as well as taking place on a different time scale.

That is, in striking contrast to the stable/unstable dichotomy associated with short-term conflict, phase transitions in spatially and temporally extended Zweikampf systems associated with $\mathscr{F}$, τ, their generalizations or their analogs, may be relatively subtle, akin to the Eldredge and Gould [14] pattern of "punctuated equilibrium" in evolutionary transition.

Following the arguments of [14], evolutionary and coevolutionary dynamics can undergo relatively long periods of apparent or near-stasis, where alterations are small or difficult to see, followed by relatively sudden massive changes leading to fundamentally different coevolutionary configurations. Something of contemporary debate on punctuated equilibrium, which has become a central matter in evolutionary biology, will be found in [30–34].

Dynamics of $\mathscr{F}$ at critical values of τ can be studied using the variants of the standard iterative renormalization techniques developed for cognitive processes in [35].

Most basically, however, phase transitions associated with changes in τ will be expressed in terms of changes in underlying groupoid symmetries, leading again to some version of the punctuated equilibrium dynamics identified by Eldredge and Gould [14]. The devil, as always, will be in the details.

A somewhat simplified model can be developed if the denominator in Eq. (1.22) is expressed as a distributed integral across the full set of information sources. Then R in Eqs. (1.14)–(1.19) can be replaced by an average $\hat{H}$ producing something like Fig. 1.5. However, the greater complexities of Eq. (1.23)—multiple phase transitions—seem likely.

A more complete description of system dynamics, beyond just modes of phase transition driven by the dynamics of τ itself, can be achieved using stochastic Onsager relations constructed in the usual manner from gradients in an "entropy" defined in terms of a vector of system parameters K:

$$S \equiv \mathscr{F}(K) - K \cdot \nabla_K \mathscr{F} \tag{1.24}$$

An Onsager approximation to system dynamics is made using a linear expansion for the rate of change of the parameter vector K in the gradients of S by the components of K [36].

Unlike a physical system, however, there cannot be any "Onsager reciprocal relations" in subsequent models, since the information sources are not locally reversible: palindromes have very low probability. See [11] for details on the relation between local time reversibility and the reciprocal relations.

The general stochastic model can be expressed as

$$dK_t = f(K, t)dt + g(K, t)dW_t \tag{1.25}$$

dW_t is a multidimensional Brownian noise. The f is assumed to be locally Lipschitz, and inherently stable or unstable, depending on battlefield conditions. Appleby et al. [37] show that a function g can always be found that stabilizes an inherently unstable function f, i.e., one that diverges to ∞. Conversely, in two or more dimensions, *a function g can always be found that destabilizes an inherently stable equilibrium for f*. Such contrasting behaviors can have important policy implications [38].

Extending the idea of "temperature" τ to more complicated algebraic entities seems possible, as does fitting such statistical models as these to real data. This work remains to be done.

Military case histories involving punctuated equilibrium transitions are not rare, well beyond the Punic Wars and the various colonial experiences in places like Afghanistan, Roman Germany, and Vietnam.

For the United States, a long-simmering conflict has ranged across the nation since well before the Civil War, culminating in that war, Reconstruction and the end of official slavery, through the 1877 agreement that ended Reconstruction and saw imposition of the Jim Crow system of reversed voting rights suppressive laws. This was followed, mid-twentieth century, with the Civil Rights Movement, often called Second Reconstruction, and by a systematic reversal beginning under the Reagan Administration, carried through further under the Clinton Administration, resulting in mass incarceration of African Americans. Mass incarceration is increasingly viewed as, essentially, an updated form of slavery. Currently, large-scale efforts are being made to again reverse African–American voting rights, often under the explicit or hidden rubric of a "white nationalism" that sometimes invokes Confederate and Nazi symbolism.

Western "nation building" in the Middle East post-WWI and post-WWII also comes to mind.

The progression on the Eastern Front of WWII from the battles of Moscow and Stalingrad to Kursk and Operation Bagration can be interpreted from these perspectives, i.e., as the punctuated equilibrium evolution of the Red Army under the selection pressure of German combined arms Bewegungskrieg. While the Wehrmacht maintained measurable tactical superiority throughout a good part of the war, at least in terms of man-to-man and unit-to-unit capabilities, and the Russians suffered vastly more casualties in the process of that evolutionary trajectory, the Wehrmacht still collapsed as Germany was literally bled dry, the final defense of Berlin given over to an army of 14 year olds and grandfathers.

The last—and not insignificant—Wehrmacht tactical victory, in March 1945, was "Operation Gemse", the defense of Lauban, Silesia, for which, notoriously, a 16-year-old high school student was awarded an Iron Cross.

More recent military examples can be found in the breakdown of civil order across Iraq in the aftermath of the Second Gulf War. Rayburn et al. [39], for example, in an official evaluation, write

> Although the invasion force was able to defeat the decrepit Iraqi military forces, it was not able to fill the void when the Iraqi state collapsed. Vast tracts of the country were left relatively unsecured and irregular forces, tribal connections, and complex social dynamics vexed coalition efforts to conclude the campaign. In the absence of central authority, Iraqis looted the country's infrastructure and communities began to fragment along ethno-sectarian lines. The ad hoc and anemic civilian and military headquarters that were established by the U.S. after the invasion proved similarly unable to restore order. Many of the initial decisions of these organizations... made the situation far worse. While the U.S. military had achieved operational success during the invasion, it was unable to consolidate its gains and achieve a strategic victory. The resultant governance and security vacuum in the summer of 2003 was quickly occupied by Sunni resistance organizations, Islamic terrorists, Shi'a militants, the Iranian regime, and Kurdish factions—circumstances that effectively ceded the initiative from coalition forces to Iraq's competing insurgent groups for years to come.

The results of this brilliant tactical finesse in the almost complete absence of competent strategic insight will be with us for the next century.

1.5 Discussion

Krepinevich and Watts [2, 3], using a different vocabulary, describe ways in which the world of strategy—of Eldredge/Gould coevolutionary dynamics—is far from the world of short-term tactics or of a sequential "muddling through". The fundamental challenge is to exert an explicit cognitive control of a coevolutionary phenomenon in which one may "lose" some, or even most, tactical confrontations. As Krepinevich and Watts indicate, winning individual battles is usually be a relatively direct, if neither simple nor easy, matter of having sufficient local resources, training, and resolve. Winning wars or other extended conflicts is not direct, and requires explicit long-term management of often subtle coevolutionary processes that may not be obvious, or even manifest, on individual tactical timescales.

Charles Darwin recognized that examples of cognitively directed evolution, of human-driven selection pressures, are not at all rare in other contexts. These examples gave us the broad and deep agricultural bases needed for large-scale human organization and technology. Indeed, modern crops are far different from their wild predecessors. Now, directed coevolution of the inevitable conflicts between the various segments of that organization may be needed for humanity's long-term persistence.

Analogous—but not identical—approaches, without the burden (or benefit) of a mathematical model of "punctuated equilibrium", are to be found in some traditional Chinese strategic thinking. Jullien's seminal study of Chinese doctrine [40] outlines something of this as follows:

> For something to be realized in an effective fashion, it must come about as an effect. It is always through a process (which transforms the situation), not through a goal that leads (directly) to action, that one achieves an effect, a result...
>
> Any strategy thus seems, in the end, to come down to simply knowing how to *implicate* an effect, knowing how to tackle a situation upstream in such a say that the effect flows "naturally" from it...
>
> All we need do is implant those ends [we wish to obtain] in the trajectory of things. In this way, left to its immanence, the desired effect is realized...
>
> So strategy is always a matter of knowing how to impinge upon the process upstream, in such a way that an effect will then tend to "come" of its own accord... It resembles a fruit that, changing imperceptibly, eventually ripens...
>
> ...[O]ne must reduce the opponent to a passivity by very gradually stripping him of his ability to react... In contrast to the event constituted by a battle, which gives rise to resistance, there is a continuously unfolding process in which the strength of the antagonist is progressively dissolved...

Such a paraevolutionary cultural perspective is well contrasted with US efforts in Korea after Inchon, in Vietnam, Iraq, and Afghanistan. Another example is the grossly arrogant stupidity and incompetence of the poorly executed Japanese attack on Pearl Harbor which gives rise to an enormous "resistance", in the Chinese sense. The Japanese failed to eliminate US aircraft carriers, but also left intact extensive repair facilities and the vast stores of fuel oil central to subsequent US prosecution of the war in the Pacific.

You really must know as much about the enemy as about yourself, and since, as the evolutionary anthropologist Robert Boyd once put it, "Culture is as much a part of human biology as the enamel on our teeth", this becomes quite tricky in cross-cultural conflict.

And again, and finally, the formal models studied here are far more about conflict rather than of it, and should be used in the manner Pielou suggests, i.e., as sources of new speculation, suggesting the empirical and observational studies that are the only possible sources of new knowledge. That being said, regression equations based on the Central Limit Theorem have proven useful, both as benchmarks against which to compare data, and, for limited purposes, as tools for prediction and control. The probability models introduced here, based on the asymptotic limit theorems of control and information theories, after sufficient further development, might, within their limitations, serve similar functions.

One important extension of the modeling exercise itself, however, would be to use stochastic differential equations with distributed delays, a different characterization of “friction” on a Clausewitz landscape. The work by Mao et al. [41] suggests that “well behaved” SDE models with fixed delays can be driven into instability by sufficient noise. Presumably, some similar dynamic would occur for distributed delays as well.

1.6 Chapter Appendix: Uncertainty in τ

The appearance of different scalar “temperatures” at tactical and strategic scales suggests, in a next iteration of the theory, the necessity of reformulating the idea of “temperature” itself. While, at the tactical, or even perhaps the operational level, the control theory arguments leading to the scalar index $\mathscr{T}$ of Eq. (1.5) may be reasonable, at the strategic level many things are happening on many different scales, including across spatial and social geographies. Under such circumstances, the strategic measure τ may not be a scalar. It may be a vector, tensor, or even an algebraic structure, like a groupoid or something even more complex.

A first step in this direction is to reformulate the argument leading to Eqs. (1.16) and (1.17) by allowing τ to be distributed according to some probability density function. Most simply, we can take that as the exponential, with density function $p(\tau) = \alpha \exp[-\alpha\tau]$ so that $< \tau >= 1/\alpha$. Then the expression for the pseudoprobability of Eq. (1.16) becomes

$$\mathscr{P}[H(z_q)] = \frac{\int_0^\infty \exp[-H(z_q)/\tau]p(\tau)d\tau}{\sum_z \int_0^\infty \exp[-H(z)/\tau]p(\tau)d\tau} \tag{1.26}$$

with

$$\int_0^\infty \exp[-H/\tau]p(\tau)d\tau = 2\sqrt{\frac{H}{<\tau>}}\texttt{BesselK}(1, 2\sqrt{\frac{\texttt{H}}{<\tau>}}) \tag{1.27}$$

This leads to the definition of a “free energy” Morse Function $\mathscr{F}$ as

$$2\sqrt{\frac{\mathscr{F}}{<\tau>}}\texttt{BesselK}(1, 2\sqrt{\frac{\mathscr{F}}{<\tau>}}) = \sum_z 2\sqrt{\frac{H(z)}{<\tau>}}\texttt{BesselK}(1, 2\sqrt{\frac{\texttt{H(z)}}{<\tau>}}) \tag{1.28}$$

For a Rayliegh distribution, having $p(\tau) = (\tau/\sigma^2)\exp[-\tau^2/(2\sigma^2)]$, and $< \tau >= \sqrt{\pi/2}\sigma$, the expression for $\mathscr{F}$ is in terms of the Meijer-G function:

$$\frac{\mathscr{F}}{< \tau >} G^{3,0}_{0,3}\left(\frac{1}{16}\frac{\mathscr{F}}{< \tau >}^2 \pi \Big|_{1/2,0,-1/2}\right) = \sum_z \frac{H_z}{< \tau >} G^{3,0}_{0,3}\left(\frac{1}{16}\frac{H_z}{< \tau >}^2 \pi \Big|_{1/2,0,-1/2}\right) \tag{1.29}$$

References

1. Dupuy, T. 1977. *A genius for war: the German army and general staff 1807–1945*. London: Hero Books.
2. Krepinevich, A., and B. Watts. 2009a. *Retaining strategic competence*. Washington, D.C.: Center for Strategic and Budgetary Assessment.
3. Krepinevich, A., and B. Watts. 2009b. Lost at the NSC. *The national interest* 99: 63–72.
4. Pielou, E. 1977. *Mathematical ecology*. New York: Wiley.
5. Osinga, F. 2007. *Science, strategy and war: the strategic theory of John Boyd*. London: Routledge.
6. Nair, G., F. Fagnani, S. Zampieri, and R. Evans. 2007. Feedback control under data rate constraints: an overview. *Proceedings of the IEEEE* 95: 108–137.
7. Wallace, R. 2017. *Computational psychiatry: a systems biology approach to the epigenetics of mental disorders*. New York: Springer.
8. Glazebrook, J.F., and R. Wallace. 2009. Rate distortion manifolds as models for cognitive information. *Informatica* 33: 309–345.
9. Wallace, R., and D. Wallace. 2016. *Gene expression and its discontents: the social production of chronic disease*, 2nd ed. Berlin: Springer.
10. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
11. de Groot, S., and P. Mazur. 1984. *Nonequilibrium thermodynamics*. New York: Dover.
12. Brah, F., A. Zaidi, J. Louveaux, L. Vandendorpe, 2011, On the Lambert-W function for constrained resource allocation in cooperative networks. *EURASIP Journal on Wireless Communications and Networking*. http://jwcn.eurasipjournals.com/content/2011/1/19
13. Protter, P. 2005. *Stochastic integration and differential equations*, 2nd ed. Berlin: Springer.
14. Eldredge, N., and S. Gould. 1972. Punctuated equilibrium: an alternative to phyletic gradualism. In *Models in paleobiology*, ed. T. Schopf, 82–115. San Francisco: Cooper and Co.
15. Rogers, C. 1993. The military revolutions of the hundred years' war. *The Journal of Military History* 57: 241–278.
16. Leonard, S. 2001. Inevitable evolutions: punctuated equilibrium and the Revolution in Military Affairs. A monograph of the School of Advanced Military Studies. United States Army Command and General Staff College, Ft. Leavenworth, KA (online).
17. Leventoglu, B., and B. Slantchev. 2007. The armed peace: a punctuated equilibrium theory of war. *American Journal of Political Science* 51: 755–771.
18. Johnson, D. 2009. Darwinian selection in asymmetric warfare: the natural advantage of insurgents and terrorists, 2009. *Journal of the Washington Academy of Sciences* 95: 89–112.
19. Atlan, H., and I. Cohen. 1998. Immune information, self-organization, and meaning. *International Immunology* 10: 711–717.
20. Khinchin, A. 1957. *Mathematical foundations of information theory*. New York: Dover Publications.

21. Wallace, R. 2018. *Carl von Clausewitz, the Fog-of-War, and the AI revolution: the real world is not a game of go*. New York: Springer.
22. Feynman, R. 2000. *Lectures on computation*. New York: Westview Press.
23. Champagnat, N., R. Ferriere, and S. Meleard. 2006. Unifying evolutionary dynamics: From individual stochastic process to macroscopic models. *Theoretical Population Biology* 69: 297–321.
24. Dembo, A., and O. Zeitouni. 1998. *Large deviations and applications*, 2nd ed. New York: Springer.
25. Pettini, M. 2007. *Geometry and topology in Hamiltonian dynamics and statistical mechanics*. New York: Springer.
26. Matsumoto, Y. 2002. *An introduction to Morse theory*. Providence, RI: American Mathematical Society.
27. bin Ladin, U. 2004. https://www.aljazeera.com/archive/2004/11200849163336457223.html
28. Weinstein, A. 1996. Groupoids: unifying internal and external symmetry. *Notices of the American Mathematical Association* 43: 744–752.
29. Golubitsky, M., and I. Stewart. 2006. Nonlinear dynamics and networks: the groupoid formalism. *Bulletin of the American Mathematical Society* 43: 305–364.
30. Hunt, G., M. Hopins, and S. Lidgard. 2015. Simple versus comples models of trait evolution and stasis as a response to environmental changes. *PNAS* 112: 4885–4890.
31. Landis, M., and J. Schraiber. 2017. Pulsed evolution shaped modern vertibrate body sizes. *PNAS* 114: 13224–13229.
32. Valverde, S., R. Sole. 2015. Punctuated equilibrium in the largescale evolution of programming languages. *Journal of the Royal Society Interface* 12:20150249 (online).
33. Lyne, J., H. Howe. 2017. "Punctuated equilibrium": rhetorical dynamics of a scientific controversy, Chap. 6. In *Landmark Essays on Rhetoric of Science: Case Studies*, ed. R. Harris, 2nd edn. Boca Raton: CRC Press.
34. Gould, S. 2002. *The structure of evolutionary theory*. Cambridge, MA: Harvard University Press.
35. Wallace, R. 2005. *Consciousness: a mathematical treatment of the global neuronal workspace model*. New York: Springer.
36. Wallace, R. 2018. New statistical models of nonergodic cognitive systems and their pathologies. *Journal of Theoretical Biology* 436: 72–78.
37. Appleby, J., X. Mao, and A. Rodkina. 2008. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control* 53: 683–691.
38. Wallace, R., et al. 2018. *Clear-cutting disease control: capital-led deforestation, public health austerity, and vector-borne infection*. New York: Springer.
39. Rayburn, J., F. Sobchak (eds.). Executive Summary: The U.S. Army in the Iraq War: Volume 1 Invasion-Insurgency-Civil War. Washington, D.C: Strategic Studies Institute and the U.S. Army War College
40. Jullien, F. 2004. *A treatise on efficacy: between western and Chinese thinking*. Honolulu: University of Hawaii Press.
41. Mao, X., C. Yuan, and J. Zou. 2005. Stochastic differential delay equations of population dynamics. *Journal of Mathematical Analysis and Applications* 304: 296–320.

Chapter 2
Doctrine and the Fog-of-War

2.1 Summary

Human conflict occurs on Clausewitz landscapes whose distinctive topology includes fog-of-war, friction, uncertainty, and misperception. Actively constraining an adversary using such features is a standard tactic of confrontation across courts of law, commerce, political campaigns, and, of course, the battlefield itself. Time-limited optimization models of cognitive effectiveness, extending standard "anytime algorithm" methods, show how doctrinal groupthink may amplify fog-of-war and friction to fatally limit the ability of a challenged institution to respond to shadow price demands imposed by an unfriendly agent or environment.

2.2 Introduction

As D.N. Wallace once put it, "Happiness is having a stupid enemy." Here, we will explore something of the dynamics of "self-stupidity" in organized conflict.

The Eastern Front of World War II determined the outcome of the war in Europe and provides a particularly striking example of conflict between cognitive institutions, the Nazi state's Wehrmacht and the Stalinist state's Red Army.

In the early stages of the conflict, up to the decisive battle near Moscow, the Wehrmacht's sophisticated, tactically flexible combined arms operations—a technological update of the classic Prussian Bewegungskrieg—did not face much resistance from the Red Army. Soviet forces, initially seen as competent and professional, suffered decimation of the officer corps in the Stalin purges of the late 1930s, and were further burdened by inexperienced replacements. The Red Army, however, entered the conflict with the Wehrmacht not only damaged by the loss of capable senior officers, but seriously constrained by Communist Doctrine that war was, in fact, a "scientific" enterprise, understandable through rigid and known "laws".

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_2

These doctrines were, early in the conflict, rigidly enforced by unit-level "political officers". Both circumstances, in the first months of the conflict, combined to ensure to the loss of 4 million men.

German forces, which stupidly foresaw only a lightning operation—the Nazi state was famously burdened by its own doctrines—became overextended into the infamous Russian winter. The Wehrmacht was then confronted with a counteroffensive driven by the USSR's Eastern reserve divisions, and was catastrophically rolled back from Moscow. Stalin had begun listening to experienced senior officers, and had brought many others back from the Gulag. Wehrmacht defeats at Stalingrad, Kursk, and finally in Operation Bagration, saw the collapse of the Wehrmacht's Army Group Center in Summer 1944, in spite of having man-to-man and unit-to-unit tactical superiority over the Russians throughout virtually all of the war.

The Red Army had, over time, evolved sufficiently good logistic, tactical, operational and strategic capabilities under the relentless German selection pressure. By contrast, Hitler, and his collaborators on the German General Staff, after the near-catastrophe at Moscow, and even after the losses at Stalingrad and Kursk, persisted in a war of attrition the Nazi state was fated to lose.

Roughly, the Wehrmacht showed greater flexibility in the early stages of the conflict, but the Red Army had greater flexibility—and resources—in the later.

There are other examples. One of the most striking was the blindsiding of US forces in Korea in 1950, after the spectacularly successful amphibious landing at Inchon. MacArthur then single-mindedly advanced toward Manchuria, while continually and steadfastly ignoring reliable intelligence information about Chinese military buildups and deployments. US Cold War doctrine held that "Communism" at the time was directed from Moscow, and Stalin would not allow the Chinese to interfere in Korea. American China specialists would certainly have known better, but had been purged from government and the universities for "losing China" as part of McCarthy Era doctrine. The Chosun Reservoir episode is worth study. The same dynamics, in various guises, have burdened US involvement in Vietnam, Iraq, Afghanistan, and so on. As has been said, those who do not remember history are condemned to repeat it.

There are many similar examples from law and commerce to "normal" political contention.

This essay formally compares institutional cognitive rigidity with flexibility in real-time conflict under resource constraints and time limitations Carl von Clausewitz characterized as "friction" and "fog-of-war", representing synergisms of uncertainty and unpredictability.

2.3 A Complex Model

Real-time cognition in living things, or cultural artifacts of machine, institutional, or composite structure, operates under constraints of time, situational awareness, precision of action, and material resources. Cognition requires a choice of some specific response to challenge from an embedding ecosystem, taken from a larger set of actions that are realistically available, according to what is known (or "emotionally intuited") from available information.

Even nonminded organisms learn or inherit emotional responses that increase survival rates by truncating response times to potentially deadly patterns of stimulation. In conflict, individuals, units, and large command structures must likewise act under challenge according to experience, intuition, or "doctrine", in some sense, to make the best possible choice under simultaneous multiple constraints.

This is both obvious and hard to formalize. To do so we abduct methods from computational "anytime algorithms" (AA) constructed to converge sufficiently rapidly on a "best" solution so that interruption of computation before final convergence produces a measurably "good enough" solution.

We take this route because AA have become central to computation and artificial intelligence under time constraints, and the mathematics has been worked out in some, if far from sufficient, detail.

We assert that military systems at the tactical, operational, and strategic levels of organization, as with many other real-world institutions, must likewise reach analogously "good enough" decisions using partial information under time and resource constraints. Anytime algorithm arguments appear, then, to be sufficiently powerful for a first approximation.

Zilberstein [1] explains the ubiquity and utility of anytime algorithms in the following manner,

> ...[Anytime algorithms] give intelligent systems the capability to trade deliberation time for quality of results. This capability is essential for successful operation in domains such as signal interpretation, real-time diagnosis and repair, and mobile robot control. What characterizes these domains is that it is not feasible (computationally) or desirable (economically) to compute the optimal answer...

As Fig. 2.1, from Zilberstein [1], shows, the "quality" of a "traveling salesman" problem calculation approaches an asymptotic limit with time, but—and importantly—with much variance.

Ashok and Patra [2], in Fig. 2.2, studied best-route problems over 100, 200, and 250 cities, fitting their empirical results to an Arrhenius function. Arrhenius functions seem a "natural" means of expressing the rate of cognition in terms of the rate at which essential resources are available. This is essentially similar to calculations of chemical reaction rate as a function of temperature.

However, and surprisingly, anytime algorithms are not necessarily stable. Indeed, sufficient conditions for stability must be established case-by-case, typically using stultifyingly obscure stochastic Lyapunov function methods similar to recent results on the stabilization and destabilization of stochastic differential equations (e.g., [3–5]).

Greco et al. [6] remark that

> With respect to most anytime algorithms... the fact that anytime controllers interact in feedback with dynamic systems introduces severe difficulties in their synthesis and in the analysis of the resulting closed-loop performance. Indeed, the stochastic switching system ensuing from executing different controllers at different times is even prone to instability.

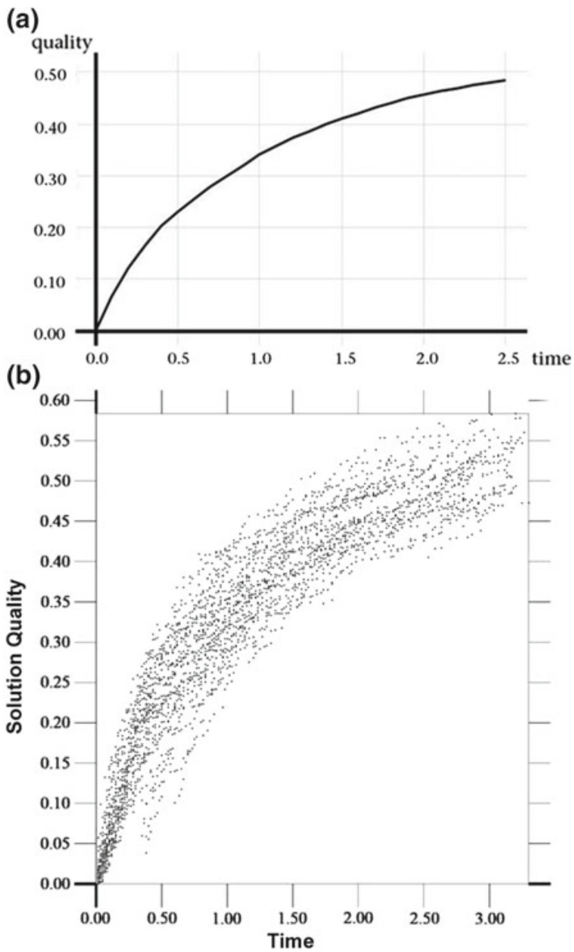


Fig. 2.1 From Zilberstein [1]. **a** Generic form of a quality-of-calculation result as a function of time under a generalized anytime algorithm. **b** Empirical realizations of quality measures at time *t* of a randomized tour-improvement algorithm for the traveling salesman problem over 50 cities. Under what circumstances does the variance explode?

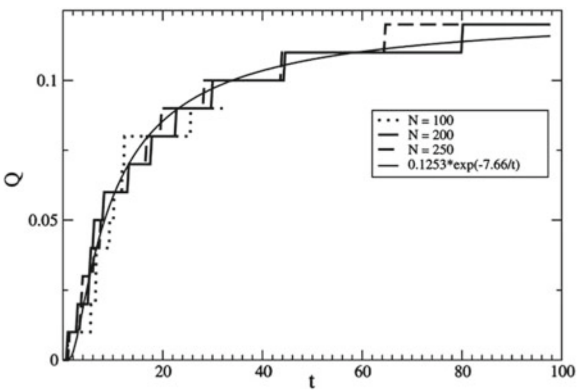


Fig. 2.2 From Ashok and Patra [2]. Quality measure for their anytime algorithm for the cities problem, N = 100, 200, 250. The fitted line is for an Arrhenius function

In contrast with most recent work on AA—the establishment of obscure sufficiency conditions for stability—we focus on necessary conditions for stability, and, conversely, on sufficient conditions for instability.

Our particular interest is how ideological rigidities and "fog-of-war" burdens become synergistic.

As described above, there are many historical examples.

The simplest approach is through a "computational bandwidth" model for institutional cognition under time constraint. Here, the "power" provided to the channel associated with an institutional cognitive module is measured by the time interval T_j over which essential resources are supplied. Assuming the usual Gaussian channels, their capacities follow the classic relation [7]

$$C_j = \beta \log(1 + \frac{T_j}{N_j}) \tag{2.1}$$

The modules are assumed to be structured in a serial and hierarchical manner—the usual tactical, operational, strategic arrangement—having the overall time constraint $T = \sum_j T_j$ as imposed by a cognitive opponent, "simple" environmental factors, or some synergism between them.

"Noise" N_j is assumed to be some composite—sum, product—of the different factors, including "fog-of-war" uncertainties and misperceptions, but also taking into account the influences of rigid doctrine that is at variance with reality, and the machinations of enforcers imposing that doctrine at and across tactical, operational, and strategic levels of organization.

Given a best-case independence of system subcomponents, a standard Lagrangian optimization is similar to the usual parallel channel calculation [7]

$$\begin{aligned} L = \sum_j \beta_j \log(1 + \frac{T_j}{N_j}) + \lambda(T - \sum_j T_j) \\ \partial L/\partial T_j = \frac{\beta_j}{N_j + T_j} = \lambda \\ T_j = \frac{\beta_j}{\lambda} - N_j \end{aligned} \tag{2.2}$$

λ is usually taken as an irrelevant "undetermined multiplier" needed to make the optimization converge, perhaps using the Kuhn–Tucker "reverse-water filling" mechanism [7]. λ, however, is not so simply interpreted in economics. There, it is taken as the "shadow price" imposed by the conditions of constraint. As Jin et al. [8] explicitly demonstrate, optimization can fail, if a shadow price can be imposed by an adversary as an "environmental demand" that the "firm" cannot meet.

If N_j is sufficiently large, $T_j \rightarrow \leq 0$. This is a condition that cannot be met by any real system. Similarly, if $\lambda \geq \beta_j / N_j$, the matter is similarly moot.

The ability to respond to a shadow price driven by an adversary is, in this model, inversely related to to the noise level N_j, a factor composed of what Clausewitz represented as "friction" and the "fog-of-war", compounded in various ways by the burden of "doctrinal" or "command blindness" noise.

"Ordinary" fog-of-war and frictional noise, N_q, may become synergistic as well as additive with doctrinal noise N_r, for example, as

$$N_j = N_q^j + N_r^j + \mathscr{A} N_q^j N_r^j \tag{2.3}$$

where $\mathscr{A} > 0$ may, as with the Red Army early in WWII, be debilitatingly large.

It is possible to make these arguments more realistic. For example, Eq. (2.1) assumes that the quality of a cognitive system or algorithm increases relentlessly with time. This is unlikely. In reality, cognitive processes or algorithms "top out" asymptotically. Typically, one uses an exponential model that for a quality measure Q_j as

$$Q_j = \frac{\beta_j}{\alpha_j}(1 - \exp[-\alpha_j T_j]) \rightarrow \frac{\beta_j}{\alpha_j} \tag{2.4}$$

The resulting Lagrangian optimization relations are

$$\begin{aligned} L = \sum_j \frac{\beta_j}{\alpha_j}(1 - \exp[-\alpha_j T_j]) + \lambda(T - \sum_j T_j) \\ \partial L/\partial T_j = \beta_j \exp[-\alpha_j T_j] = \lambda \end{aligned} \tag{2.5}$$

In the first order—without noise—the constraint on the shadow price becomes $0 \leq \lambda \leq \beta_j$.

For the multiplicative chain-of-command system, a more "natural" Lagrangian is expressed in terms of $\log(\Pi_j Q_j) = \sum_j \log(Q_j)$, giving

$$\partial L/\partial T_j = \frac{\alpha_j \, e^{-\alpha_j T_j}}{1 - e^{-\alpha_j T_j}} = \lambda \tag{2.6}$$

This is illustrated in Fig. 2.3. Sufficient shadow price λ still drives response times T_j to impossibly small values.

The effects of "noise" must be added in some way, here using an Ito stochastic differential equation (SDE) [9].

Given that

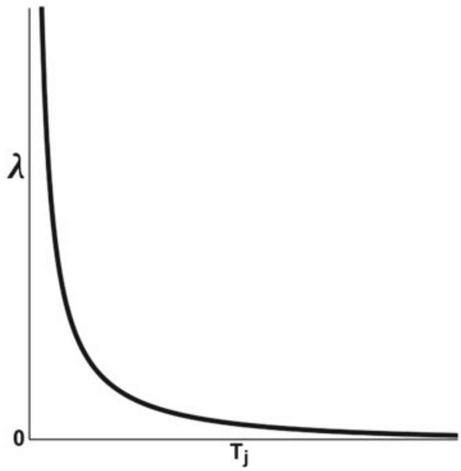
$$dQ/dt = \beta \exp[-\alpha t] = \beta - \alpha Q \tag{2.7}$$

we explore an SDE having the form

$$dQ_t = (\beta - \alpha Q_t)dt + \sigma Q_t dW_t \tag{2.8}$$

where the second term represents volatility under Brownian noise, here represented as dW_t.

Fig. 2.3 Shadow price relation for a synergistic chain-of-command system. Sufficiently high λ places impossible restrictions on response times



The expectation is $E(Q) \to \beta/\alpha$. However, a new dynamic appears by applying the Ito Chain Rule to Q^2 [9]. This permits an explicit calculation of the nonequilibrium steady-state variance as $E(Q^2) - E(Q)^2$. Then, after some algebra,

$$Var(Q) \to \left(\frac{\beta}{\alpha - \sigma^2/2}\right)^2 - \left(\frac{\beta}{\alpha}\right)^2 \tag{2.9}$$

The variance explodes as $\sigma^2/2 \to \alpha$, driving the system to extinction.

Thus, for this plausible model of an "anytime algorithm", sufficient noise is catastrophic.

Further, as done above, we can infer that

$$\sigma = \sigma_q + \sigma_r + \mathscr{A}\sigma_q\sigma_r \tag{2.10}$$

Thus "doctrinal" noise can be both additive to, and synergistic with, "operational" noise to affect the overall system stability.

Every "anytime algorithm" must show a similar pattern of instability under sufficient noise [3].

The Arrhenius version of an anytime algorithm [2], however, is most remarkable. Then

$$Q_j = \beta_j \exp[-\alpha_j/T_j] \to \beta_j \tag{2.11}$$

and the Lagrangian optimization is

$$\partial L/\partial T_j = \frac{\beta_j \alpha_j}{T_j^2} \exp[-\alpha_j/T_j] = \lambda \tag{2.12}$$

The limit on shadow price is

$$0 \leq \lambda \leq \frac{4}{\exp[2]} \frac{\beta}{\alpha} \approx 0.54134... \frac{\beta}{\alpha} \tag{2.13}$$

The synergistic chain-of-command form of this relation is $\partial L/\partial T_j = \alpha_j/T_j^2 = \lambda$. This leads again to a result like that of Fig. 2.3.

Some further development, under conditions of Brownian noise and volatility, gives an SDE expression as

$$dQ_t = \frac{Q_t}{\alpha} (\log[Q_t/\beta])^2 \, dt + \sigma Q_t dW_t \tag{2.14}$$

The Ito chain rule applied to $E(Q^2)$ gives

$$E(Q^2) \to \beta \exp[\sigma \sqrt{-\alpha/2}] \tag{2.15}$$

For $\sigma > 0$, this must be complex, since $\alpha > 0$.

Consequently, for an Arrhenius anytime algorithm, any level of noise may eventually become destabilizing.

A "tanh" anytime algorithm is described by the relations

$$Q = \frac{2}{1 + \exp[-\alpha t]} - 1 \to 1$$
$$dQ_t = \alpha(1 - Q_t^2)dt + \sigma Q_t dW_t \tag{2.16}$$

Some calculation shows the variance grows explosively, not as σ^2 as perhaps expected, but as σ^4.

Although these three models have markedly different stochastic properties, Fig. 2.4 shows their approaches to steady state all resemble Fig. 2.1a.

Examining anytime algorithms from a more comprehensive meta-computation perspective, they can be viewed as mappings of a particular "message" along what amounts to an information channel for which a quality measure improves with time. The difference between a scalar index of ultimate and instantaneous computational quality, say $D > 0$, then can be interpreted as a distortion measure in the usual Rate–Distortion argument. There is, then, some Rate–Distortion Function $R(D)$, convex in the scalar distortion index D, that determines the minimum metachannel capacity needed to keep the average distortion measure at or below some fixed value D. Taking Feynman's [10] characterization of information as another kind of free energy, it becomes possible to invoke a classic Onsager model from nonequilibrium thermodynamics [11] to study system dynamics.

Typically, one takes a Gaussian channel under the squared distortion measure. The Rate–Distortion Function [7] can then be expressed as $R(D) = \beta \log(\varepsilon^2/D)$ where

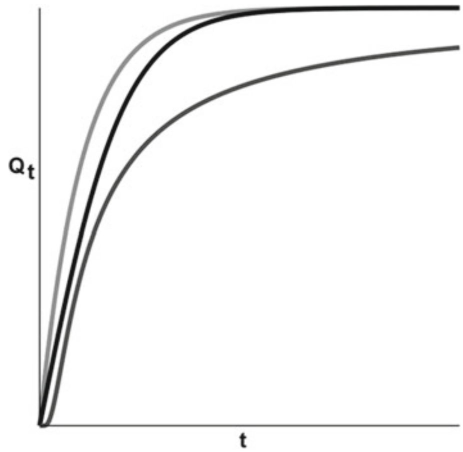


Fig. 2.4 Exponential, Arrhenius, and "tanh" models of an anytime algorithm appears to have similar direct time dynamics. However, the exponential model shows a phase transition at a critical value of "noise" σ, the Arrhenius is unstable at any nonzero value, and the variance of the "tanh" model grows as σ^4 rather than as σ^2 as expected. These are deep waters

ε is the noise inherent to the computation, as opposed to fog-of-war and frictional burdens that we express as σ. Since, following Feynman, $R(D)$ is a free energy measure, it becomes possible to define an entropy as $S = R - DdR/dD$, and to construct a first-order Onsager diffusion model in the entropy gradient as $dD/dt = -\mu dS/dD = \mu\beta/D(t)$. By inspection, this relation has the solution $D(t) \propto \sqrt{\mu t}$. This should be recognized as an important "correspondence reduction" to ordinary diffusion.

The next step in the argument involves constructing a stochastic differential equation in terms of the fog-of-war and frictional noise σ:

$$dD_t = [\frac{\mu\beta}{D_t} - F(t)]dt + \sigma D_t dW_t \tag{2.17}$$

$F(t)$ can be any function that asymptotically approaches the peak limiting value $F(\infty)$.

Given noise, applying the Ito Chain Rule to D^2 shows that a necessary, but not sufficient, condition for stability in variance for all such models is

$$F(\infty) \geq \sigma\sqrt{2\mu\beta} \tag{2.18}$$

Some similar relation holds for all possible channels as a direct consequence of the inherent convexity of the Rate–Distortion Function.

This argument strongly suggests that sufficient "system noise" will destabilize all possible anytime algorithms. As above, we see σ as both additive and synergistic in fog-or-war, frictional and doctrinal noise.

2.4 Another Approach

A more formal "mathematical" route to doctrine-as-noise is possible by examining a stochastic model for an *inverse casualty rate*, say Z_t, representing some total burden on men and material depending on interaction with an opponent. We take this as indexed by a variable Y_t, according to a relation of the form

$$dZ_t = Z_t f(Y_t)dt + \sigma Z_t g(Y_t)dW_t \tag{2.19}$$

Both f and g are seen as very general functions that may not be continuous and can include Levy jumps [9] with their own distribution properties. W_t is again Brownian diffusion. Equation (2.19) can be rewritten as

$$dZ_t = Z_t dX_t \tag{2.20}$$

From ([9], Theorem 37), the solution to Eq. (2.20) is

$$Z_t = \exp[X_t - \frac{1}{2}[X, X]_t^c]\Pi_{s \le t}(1 + \Delta X_t)\exp[-\Delta X_s] \tag{2.21}$$

Here, the infinite product converges, $[X, X]_t^c = \sigma^2 t$ is the quadratic variation of the continuous part of X, and the ΔX terms are Levy "jumps".

If X_t is bounded—$X_t \le at$ for some $a > 0$—then the leading exponential term imposes a condition on the term

$$\exp[(a - \frac{1}{2}\sigma^2)t]$$

Then, if $\frac{1}{2}\sigma^2 > a$, $Z \to 0$: sufficient noise, in this model, will "sterilize" a contending agent.

Doctrine and/or groupthink contributions to fog-of-war can tip the scales in this direction.

2.5 Failure of Intelligence on a Cognitively Stressed Operational Network

Fads, rumors, and epidemics propagate differently along networked structures according to scale and level of organization. Geographers, for example, generally recognize three such patterns:

(1) Hierarchical diffusion along relatively slow large-scale travel paths, from greater to lesser conurbations.

(2) Spatial contagion along the daily journey-to-work, from central places outward, a diffusion pattern much like the spread of a winestain on a tablecloth.
(3) Person-to-person spread along social networks of face-to-face contact.

The propagation of intelligence information across linked operational entities is often rapid and can sometimes resemble spatial contagion, although the geography of spread—the tablecloth—is structured according to some Probability-of-Contact matrix (POCM) within and between individual entities rather than simple spatial diffusion. Further, the importance of such information can be diluted according to some measure of operational burden: important information must be selected from a larger set of incoming data, an act of cognition that can be dimmed by stress, lack of resources, the burden of enemy engagement, the imposed blindness of doctrine, and other aspects of friction, the fog-of-war, and the inevitable failures of planning and command that Clausewitz describes.

Gould and Wallace [12] explore the analog of spatial contagion on such a structure, and we adapt that approach, extending it by a stochastic argument very similar to the previous section.

Again, the structure must first be characterized by the underlying POCM representing patterns of within and between contact for the individual entities.

The spread of a cognitively recognized intelligence signal on a particular network of interacting entities—between and within—can be described in first order at a nonequilibrium steady state, using the equilibrium distribution ε_i of a Markov process "per unit area" A_i, where the A_i scale with the different "size" each node must cover. This abstract configuration is distinguishable by the scale variable A_i as well as by its "position" i in the associated POCM. To obtain the underlying Markov process, the POCM is normalized to a stochastic matrix $\mathbf{Q}$ having unit row sums, and the vector ε calculated in the standard manner as the eigenrelation $\varepsilon = \varepsilon\mathbf{Q}$.

There is a vector set of dimensionless network flows of intelligence information $\mathscr{X}_t^i$, $i = 1, ..., n$ at time t, each determined by some very general functional relation

$$\mathscr{X}_t^i = g(t, \varepsilon_i/A_i) \tag{2.22}$$

Here, i is the index of the node of interest, $\mathscr{X}_t^i$ is the corresponding dimensionless scaled ith signal, t the time, and g an appropriate function. Again, ε_i is defined by the relation $\varepsilon = \varepsilon\mathbf{Q}$ for a stochastic matrix $\mathbf{Q}$, calculated as the network probability-of-contact matrix between operational nodes, normalized to unit row sums. Using $\mathbf{Q}$ breaks out the underlying network topology, a fixed between-and-within communications configuration weighted by operational burden that is assumed to change relatively slowly on the timescale of observation compared to the time needed to approach the nonequilibrium steady- state distribution.

Since the $\mathscr{X}$ are expressed in dimensionless form, g, t, and A must be constructed as dimensionless as well giving, for the monotonic increasing function G

$$\mathscr{X}_\tau^i = G[\tau, \frac{\varepsilon_i}{A_i} \times \mathscr{A}_\tau] \tag{2.23}$$

where $\mathscr{A}_\tau$ is the value of a *characteristic area* that represents the spread of intelligence information across the network at (dimensionless) characteristic time $\tau = t/T_0$.

G can be complicated, including other dimensionless structural variates for each individual node i.

The central idea of the analysis, following the previous section, is that the "characteristic area" $\mathscr{A}_\tau$ representing the degree of propagation of the intelligence information itself grows according to a stochastic process, even though G may be a deterministic mixmaster driven by systematic local probability-of-contact or other flow patterns.

The characteristic area of an intelligence report cannot grow indefinitely, so that there is a "carrying capacity" for the operational network, written as $K > 0$. An appropriate stochastic differential equation is then

$$d\mathscr{A}_\tau = [\alpha \mathscr{A}_\tau (1 - \mathscr{A}_\tau / K)] d\tau + \sigma \mathscr{A}_\tau dW_\tau \tag{2.24}$$

where α is an index of functional competence for the full operational enterprise, representing the "nature of the beast", as it were.

Using the Ito chain rule on $\log(\mathscr{A})$, as a consequence of the added Ito correction factor and the Jensen inequality for a concave function,

$$E(\mathscr{A}) \to 0, \ \ \alpha < \sigma^2/2 \tag{2.25}$$

If the "noise" index $\sigma^2/2$ exceeds the index of functional competence α, the characteristic area for the propagation of critical intelligence collapses to zero, as in the previous examples.

This should be recognized as the same argument leading to Eq. (1.7) and Fig. 1.4.

2.6 Discussion

Pielou ([13], p. 102), a theoretical ecologist, describes the utility and limitations of mathematical models of complex social and ecological systems in these terms:

> [Mathematical] models are easy to devise... the magic phrase "let us assume that" overrides objections temporarily... How is such a model to be tested? The [often fortuitous] correspondence between a model's predictions and observed events... cannot be taken to imply the model's simplifying assumptions are reasonable [in the real world]...
>
> [T]he usefulness of models... consists *not in answering questions but in raising them*. Models can be used to inspire new field investigations and these are the only source of new knowledge as opposed to new speculation.

From this perspective, then, what speculations emerge from this essay?

Some fairly elementary optimization arguments suggest that doctrinal noise may become intertwined or synergistic with fog-of-war uncertainties to limit the ability of an institution to respond to shadow price burdens imposed by an adversary or an environment, at various scales and levels of organization.

The second take on AA models, via SDE analyses, finds, in addition to the effect of adversarial or environmental shadow price constraints, that doctrinal and operational noise operating together may trigger literally explosive institutional failures.

Further development suggests that all "anytime algorithms" will be driven to failure by sufficient doctrinal and operational uncertainties. This result may have implications for the current rush-to-judgment on driverless vehicles, autonomous weapons, and the "AI revolution" in general, extending the work of [14].

The more formal stochastic sterilization model provides a quite general route to similar conclusions, as does the diffusion-of-intelligence analysis for an operational network under cognitive stress.

There is, perhaps, a general inference.

The Clausewitzian Zweikampf between cognitive institutional entities always involves general uncertainties and specific constraints on resources that may include time. The calculated imposition of such burdens on an opponent has always been one of the standard tactics of contention, from courts of law to commerce, political campaigns, and the battlefield. Groupthink enterprises—dominated by cultural artifacts of "doctrine" that filter out available information—appear inherently stupid when operating on Clausewitz landscapes of high frictional shadow price or fog-of-war randomness and misperception.

To restate Dr. D.N. Wallace's comment, "Happiness is having a stupid enemy".

The gods in charge of global climate change, Brexit, and the U.S. elections must be very happy indeed.

References

1. Zilberstein, S. 1996. Using anytime algorithms in intelligent systems. *AI Magazine* 17: 73–83.
2. Ashok, B., and T. Patra. 2010. Locating phase transitions in computationally hard problems. *PRAMANA Journal of Physics* 75: 549–563.
3. Quevedo, D., and V. Gupta. 2012. Sequence-based anytime control. *IEEE Transactions on Automatic Control* 58: 377–390.
4. Dang, T., K. Ling, D. Quevedo, 2018. Stability analysis of event-triggered anytime control with multiple control laws. arXiv:1804.07876v1 [math.OC] April 2018.
5. Appleby, J., X. Mao, and A. Rodkina. 2008. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control* 53: 126–132.
6. Greco, L., D. Fontanelli, A. Bicchi. 2007. Almost sure stability of anytime controllers via stochastic scheduling. In *46th IEEE Conference on Decision and Control*, 5640–5645.
7. Cover, T., and J. Thomas. 2006. *Elements of Information Theory*, 2nd ed. New York: Wiley.
8. Jin, H., Z. Hu, and X. Zhou. 2008. A convex stochastic optimization problem arising from portfolio selection. *Mathematical Finance* 18: 171–183.
9. Protter, P. 2005. *Stochastic Integration and Differential Equations*, 2nd ed. New York: Springer.
10. Feynman, R. 2000. *Lectures on Computation*. New York: Westview Press.

11. Wallace, R. 2017. *Computational Psychiatry: A Systems Biology Approach to the Epigenetics of Mental Disorders*. New York: Springer.
12. Gould, P., and R. Wallace. 1994. Spatial structures and scientific paradoxes in the AIDS pandemic. *GeograFiska Annaler B* 76: 105–116.
13. Pielou, E. 1977. *Mathematical Ecology*. New York: Wiley.
14. Wallace, R. 2018. *Carl von Clausewitz, the Fog-of-War, and the AI Revolution: The Real World is not a Game of Go*. New York: Springer.

Chapter 3
On Asymmetric Conflict

...American institutional culture, though necessarily disciplined by the prospective needs of warfare, must adopt and follow a way in warfare broadly compatible both with the overarching strategic culture, and, more particularly, with the country's public culture.

—C.S. Gray

Ride to the sound of the guns.

—Major General George Armstrong Custer

3.1 Summary

From the immune response of a higher animal against a pathogen to organized warfare between "imperialists" and "insurgents", agents of markedly different material resources have contended on Clausewitz landscapes characterized by uncertainties of "friction" and the "fog-of-war". Such asymmetric dynamics can be analyzed using the tools of information and control theories, significantly extending our understanding of why the gods are not always on the side of the big battalions.

3.2 Introduction

Cognitive entities engage in conflict and cooperation across many instances, scales, and levels of organization. For living things, there is multicellularity, mutualism, and for even simple organisms, shifting battles between pathogens, cancers, and a

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_3

relatively complex immune system. Social examples extend from cooperative farming through competition between firms and institutions, and, at larger scales, if negotiations or treaties fail, highly organized violent confrontations, as studied by Sun Zu, Carl von Clausewitz, John Boyd, and many others.

This chapter focuses on contention between cognitive agents—in a large sense—having markedly different material resources. Examples include the American War in Vietnam, and, curiously, the Eastern Front of WWII after the battles of Moscow, Stalingrad, and Kursk. Remarkably, in the face of massive Allied superiority in men and equipment, the German army was able to maintain significant cohesive resistance until early 1945.

Of course, current stand-offs in Afghanistan, Iraq, and other theaters of the Pentagon's "forever war", come to mind as well. Another—ongoing—case history is the grinding insurgency of the US Confederate States after 1864 that concluded in 1877 with the withdrawal of Federal troops and the establishment of a "Jim Crow" system that persisted well into the twentieth century, presently instantiated nationally as the mass incarceration of African Americans. The recent emergence of an explicit US "white nationalism" and parallel attempts to reimpose Jim Crow voting restrictions, are, then, just frostings on a long-baking cake.

We begin by reiterating a remark by the architect of the 9/11/2001 attacks on the USA [1]:

> All that we have to do is send two mujahidin to the furthest point east to raise a piece of cloth on which is written al-Qaida, in order to make the generals race there to cause America to suffer human, economic, and political losses without their achieving for it anything of note... [just] as we... bled Russia for 10 years, until it went bankrupt and was forced to withdraw in defeat... [s]o we are continuing this policy in bleeding America to the point of bankruptcy.

This is, of course, a very good summary.

Matters of asymmetric conflict have naturally received much attention in the US military science literature for some time. Schreiber's 1964 numerical model [2], based on analogs to the Lanchester equations, indicated that "...[A]n increase in the efficiency of an intelligence and command and control system can be equivalent to a substantial increase in numerical [force] strength".

Kress and Szechtman [3], using another elaboration of the Lanchester model, find that

> ...[R]egardless of the insurgents' initial force level, when the intelligence-gathering capabilities of the government grow slowly with the insurgents' strength, or the resentment generated by civilian casualties caused by government actions is not negligible, or the government reinforcement rate... is too low, the government is bound to lose, even with very favorable initial force ratios or attrition coefficients...
>
> ...[O]ur model represents a best case situation from the government perspective. Under the reasonable assumptions regarding the behavior of the intelligence function... we conclude that the government cannot completely eradicate the insurgency by force...

Here, in some contrast to these approaches, we will focus on the failure of institutional cognition for contending agents where one is rich in material resources and the other in information resources.

3.3 Institutional Cognition in Conflict

The approach is via an analog to the Arrhenius chemical reaction rate relation, used to model both the rate at which cognition proceeds, i.e., its "effectiveness", and the efficiency of that cognition, the effectiveness per unit resource consumed [4].

The central insight is to view the rate at which some index of essential resources, Z, is delivered to the cognitive system of interest as a temperature analog determining the rate at which institutional cognition-and-action can proceed. Lack of material, situational awareness, battlefield information, and effective communication between units will lower the rate at which a contending agent can either initiate or respond.

Following the arguments of Feynman [5], it is possible to interpret the Rate–Distortion Function R of the information channel connecting a cognitive system's intent with its actual effect [6] as a "free energy" measure [4], leading to a dynamic model similar to the Arrhenius treatment of the rate of a chemical reaction.

In more detail, the Rate–Distortion Function, $R(D)$, determines the minimum channel capacity needed to keep the average distortion between what is sent and what is received below a real number limit $0 < D$, by some appropriate measure. The RDF, for any information channel, is always a convex function of the measure D [6], and this repeatedly proves to be important.

Taking Feynman's perspective, it then becomes possible to define a Boltzmann-like probability measure defining a normalized, dimensionless, rate of cognition as

$$P[R \geq R_0] = \frac{\int_{R_0}^{\infty} \exp[-R/g(Z)]dR}{\int_{0}^{\infty} \exp[-R/g(Z)]dR} = \exp[-R_0/g(Z)] \tag{3.1}$$

where $g(Z)$ is a positive, monotonic increasing function of Z that must be determined, and R_0 is the threshold channel capacity necessary for initiation of a cognitive function, for recognition that a "signal", in a large sense, must elicit a response. For real systems, this value will never be zero, and in global broadcast settings [4], will involve sharp punctuation similar to the onset of consciousness.

The effectiveness—rate of response—and efficiency of a cognitive function indexed by j are then respectively measured as

$$\begin{aligned} F_j(Z_j) &= \exp[-K_j/g(Z_j)] \\ f_j(Z_j) &= \frac{F_j(Z_j)}{Z_j} \end{aligned} \tag{3.2}$$

for an associated local "detection level" parameter K_j.

The determination of $g(Z)$ is nontrivial, done, as in Chap. 1, by envisioning the denominator of Eq. (3.1) as a partition function.

That integral has the value $g(Z)$.

Next, exactly as in Chap. 1, we again define a "free energy" measure—basically a Morse Function—$\mathscr{F}$ as

$$\exp[-\mathscr{F}/g(Z)] = \int_0^\infty \exp[-R/g(Z)]dR = g(Z)$$
$$\mathscr{F}(Z) = -g(Z)\log[g(Z)] \tag{3.3}$$

This again leads to definition, as with Eq. (1.10), of an entropy-analog as

$$S(Z) = \mathscr{F}(Z) - Zd\mathscr{F}(Z)/dZ \tag{3.4}$$

Again imposing an Onsager nonequilibrium thermodynamics model, we recover Eq. (1.11):

$$dZ/dt \propto dS(Z)/dZ =$$
$$\frac{Z\left(g\left(Z\right)\left(\ln\left(g\left(Z\right)\right)+1\right)\frac{\mathrm{d}^2}{\mathrm{d}Z^2}g\left(Z\right)+\left(\frac{\mathrm{d}}{\mathrm{d}Z}g\left(Z\right)\right)^2\right)}{g\left(Z\right)} \tag{3.5}$$

As argued above, we assume a rapid approach to a nonequilibrium steady state, $dZ/dt \rightarrow 0$, so that the resulting differential equation can be explicitly solved for $g(Z)$ as before:

$$g(Z) = \frac{C_1 Z + C_2}{W(n, C_1 Z + C_2)} \tag{3.6}$$

Recall that $W(n, x)$ is the Lambert W-function of order n that solves the relation $x = W(n, x)\exp[W(n, x)]$ and the C_i are constants of integration. The effectiveness and efficiency measures of Eq. (3.2) are then

$$F_j(Z_j) = \exp[-K_j W(n, C_1 Z_j + C_2)/(C_1 Z_j + C_2)]$$
$$f_j(Z_j) = \frac{\exp[-K_j W(n, C_1 Z_j + C_2)/(C_1 Z_j + C_2)]}{Z_j} \tag{3.7}$$

From the properties of the W-function—and from the boundary conditions that $g(Z)$ is to be monotonically increasing and real— these relations make sense only if $n = 0$, $C_1 > 0$, $C_2 < 0$ and $C_1 Z + C_2 > -\exp[-1]$.

The boundary conditions generate a conundrum.

The forms of Eqs. (3.6) and (3.7) are shown in Fig. 3.1.

Taking into account the properties of the Lambert W-function, and imposing the boundary condition that $g(Z)$ is to be monotonically increasing and real, implies that $g(Z)$ will become complex for Z below the value $(|C_2| - \exp[-1])/C_1 > 0$.

This represents the onset of a phase transition in the rate of cognition if Z becomes too small. From general considerations regarding cognition [4], such a transition is likely to be in a disorganized, fragmented, and highly nonfunctional state. The effectiveness and efficiency measures using $g(Z)$, likewise, become complex/disorganized unless

$$Z > (|C_2| - \exp[-1])/C_1 > 0 \tag{3.8}$$

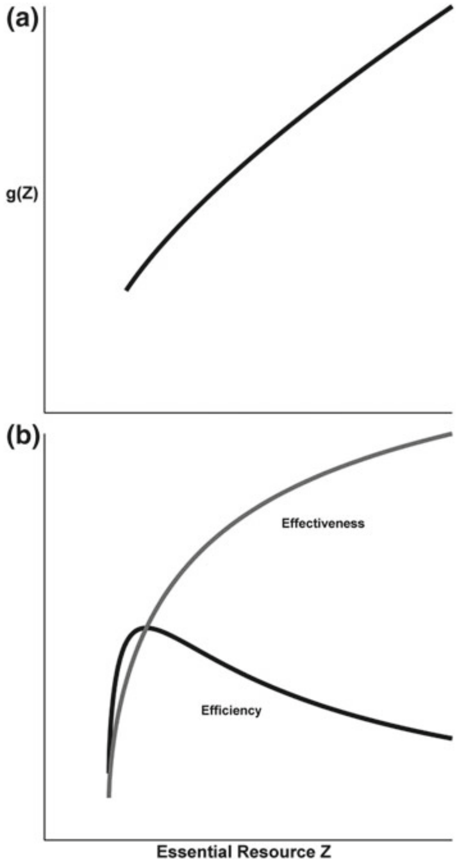


Fig. 3.1 **a** Form of $g(Z)$ from Eq. (3.6). The properties of the Lambert W-function, and imposition of the boundary condition that $g(Z)$ is to be monotonically increasing and real, implies that g becomes complex for Z below the value $(|C_2| - \exp[-1])/C_1 > 0$. This implies a phase transition if Z is too small. From studies of cognition in higher animals [4], such transition will likely be to a disorganized, highly nonfunctional state. **b** The effectiveness and efficiency measures using $g(Z)$. These are, again, complex/disorganized for $Z < (|C_2| - \exp[-1])/C_1$. Effectiveness rises monotonically with increase of the resource delivery rate Z, but efficiency peaks at some intermediate value, suggesting the onset of serious cost–benefit constraints. If $Z < (|C_2| - \exp[-1])/C_1 > 0$, both factors become disorganized

which should be compared with Eq. (1.19), essentially providing yet another RDT proof of the DRT.

What goes into the resource rate Z? Some $j = 1, ..., n$ cognitive modules must work together within an institution, a machine system, or a man/machine composite, a "cockpit" of some size. These crosstalking modules must be provided with material resources at some rate M_j. In addition, they must communicate with each other, reflected at each module by some "bandwidth" C_j representing the channel capacity available to it. Further, the embedding environment, represented here as an "external" information source H, must provide "intelligence" signals to each module at some rate H_j.

Thus, at level j, we assume a synergistic interaction, $Z_j = M_j C_j H_j$, rather than an additive one, because a system cannot persist under conflict if any individual component becomes zero.

Unlike a temperature in a chemical system, the total rate at which the essential resource Z can be delivered is necessarily limited, i.e., there is a constraint $Z = \sum_j Z_j$.

This constraint permits an optimization calculation.

Taking a simple scalarization as either $\sum_j F_j$ or $\sum_j f_j$, leads to standard Lagrangian optimization schemes as

$$
\begin{aligned}
L_F &= \sum_j F_j + \lambda_F (Z - \sum_j Z_j) \\
L_f &= \sum_j f_j + \lambda_f (Z - \sum_j Z_j)
\end{aligned}
\tag{3.9}
$$

and to the relations

$$
\begin{aligned}
\partial L_F / \partial Z_j &= \partial F_j / \partial Z_j = \lambda_F \\
\partial L_f / \partial Z_j &= \partial f(Z_j) / \partial Z_j = \lambda_f
\end{aligned}
\tag{3.10}
$$

Explicitly,

$$
\begin{aligned}
&\partial F(Z)/\partial Z = \\
&\frac{K\ (\mathrm{W}\,(C_1\,Z + C_2))^2\, C_1}{(C_1\,Z + C_2)^2\,(1 + \mathrm{W}\,(C_1\,Z + C_2))} \mathrm{e}^{-\frac{K\,\mathrm{W}(C_1\,Z + C_2)}{C_1\,Z + C_2}} \\
&= \lambda_F
\end{aligned}
\tag{3.11}
$$

See Fig. 3.2 for the graph of $\partial L_F / \partial Z_i$.

As argued in economics, since $\partial L_X / \partial Z = \lambda_X$, where $Z = \sum_i Z_i$, the parameters λ_X represents a "shadow price" index imposed on the system of interest both by the external "environmental" constraints and by internal "frictional" exigencies. If, for example, an externally driven λ_X is beyond the range of the functions $\partial L_X / \partial Z_j$, optimization is impossible, and the system will fail under protracted conflict with another system that can, in fact, be optimized (e.g., [7]).

More specifically for our effectiveness model, λ_F is an index of Clausewitzian friction. If it is sufficiently large, then Z, the rate at which essential resources can be delivered, collapses, and operations cannot be continued.

The rate of supply of critical resources, Z, which must be greater than the limit of Eq. (3.8), will be subject to the same burden of "unknown unknowns" as illustrated

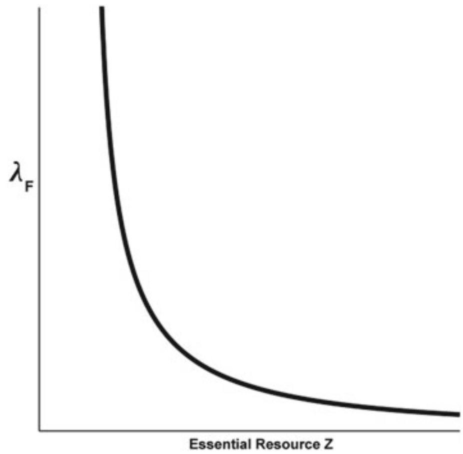


Fig. 3.2 The Lagrangian optimization of Eq. (3.10) for effectiveness. Recall that Z is the rate of delivery of an essential resource. λ_F is an index of the "shadow price" imposed by environmental demands, in a large sense that includes the deliberate actions of an adversary. Recall also that the effectiveness measure undergoes a punctuated transition to failure if $Z < (|C_2| - \exp[-1])/C_1 > 0$, producing the gap to the left. The essential point is that shadow price demand constrains the rate at which resources can be delivered, in this model. Sufficiently large λ_F chokes off delivery of essential resources, instantiating Clausewitzian friction

in Eq. (1.7) and Fig. 1.4. Even if Z is nominally within acceptable limits, sufficient fog-of-war and friction can drive an institution to cognitive, as well as control, failure.

More explicitly, it is possible to write a dynamic stochastic differential equation for Z as

$$dZ_t = \mu Z_t(1 - \frac{Z_t}{K})dt + \sigma Z_t dW_t \tag{3.12}$$

where the terms are similar to those in Eq. (1.7), leading to the stochastic stabilization of the unstable equilibrium condition $Z \to 0 < (|C_2| - \exp[-1])/C_1 > 0$ by a sufficiently large value of σ.

Indeed, one might perhaps link cognitive and control failure under the shadow price demands of conflict by writing something like

$$\mathcal{T}^F = \frac{\kappa_1^F \lambda_F + \kappa_2^F}{\kappa_3^F \lambda_F + \kappa_4^F} \tag{3.13}$$

Conflict on Clausewitz landscapes, in this model, involves complex and subtle interactions between institutional cognition and the ability of institutions to exert control in battlespace.

3.4 An Operational Level Model

We have, somewhat surprisingly, developed tools that will permit a first go at the operational level of conflict on a Clausewitz landscape, a matter not treated in the first two essays. The trick—or underlying gross oversimplification, if you will—is to recognize that operational scale dynamics are essentially a chain system in which cooperating institutions must each be successful for completion of the assigned task. The efficacy measure of Eq. (3.2), augmented by Eq. (3.6), must then be optimized as a product rather than as the sum of Eq. (3.9). Further, that product must be optimized across limited resources in the context of "frictional" effects that limit the delivery of those resources.

Putting this together is eased by using the product optimization scalarization

$$\log(\Pi_j F_j) = \sum_j \log(F_j) = \sum_j \frac{-K_j}{g_j(Z_j)} \tag{3.14}$$

Z_j is the rate at which essential resources are provided to subcomponent j. The Z_j are, in this model, taken as subject to an overall logistic constraint,

$$Z = \sum_j Z_j \tag{3.15}$$

The Lagrangian for optimization then becomes

$$L = \sum_j \frac{-K_j}{g_j(Z_j)} + \alpha(Z - \sum_j Z_j) \tag{3.16}$$

remembering Eq. (3.6).

The fundamental shadow price relation is then

$$\partial L/\partial Z_j = \frac{K_j \left(\mathrm{W}\left(C_1^j Z_j + C_2^j\right)\right)^2 C_1^j}{\left(C_1^j Z_j + C_2^j\right)^2 \left(1 + \mathrm{W}\left(C_1^j Z_j + C_2^j\right)\right)} = \alpha \tag{3.17}$$

where W is again the Lambert W-function of order zero, subject to the reality constraint

$$C_1^j Z_j + C_2^j > -\exp[-1] \tag{3.18}$$

This is illustrated in Fig. 3.3, where the Z_j are seen as forced by an externally imposed shadow price that drives α. That shadow price may be endogenous, associated with landscape and logistics, or exogenous, imposed by an adversary who may, in fact, have carefully arranged the embedding frictional impediments.

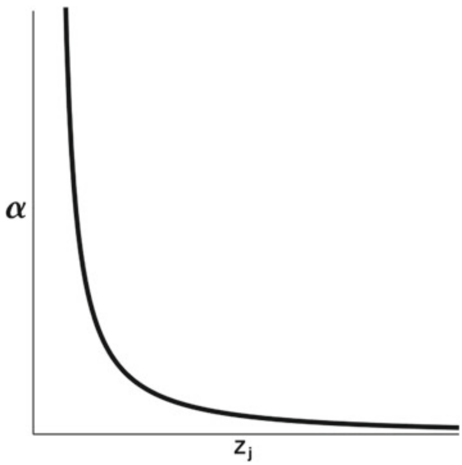


Fig. 3.3 Shadow price relation for the system of Eqs. (3.14)–(3.18). The shadow price is subject to manipulation by an adversary

One is, perhaps, reminded again of the Wehrmacht on the Eastern Front. In 1941, the initial German invasion involved a highly coordinated geographic spread of three powerful army groups across a vast reach of geography, ultimately stymied at Leningrad and Moscow. The next German attempt, in 1942, involved two operational-level armies, but only in the South, stymied at Stalingrad. The Wehrmacht's Kursk offensive of 1943 involved only a single large operational group, split into North and South pincers, and, again, failed. The Russian response, Operation Bagration, in Summer of 1944, was a well-coordinated, multipronged onslaught across a huge front, resulting in the collapse of the Wehrmacht's Army Group Center and allowing the Russian March to the Baltic. Only the Russian desire to ensure postwar control of Eastern Europe prevented an immediate full strike against Berlin, finally made in Spring 1945.

Throughout the war in the East, the Wehrmacht largely retained superiority at the level of individual combatants, units, and tactical competence. As did the US in Vietnam, Iraq, and Afghanistan.

Operational and strategic dynamics are different from tactical dynamics.

A second operational level model, focused on time constraint, is explored in the Chapter Appendix.

3.5 Iterating the Basic Idea

Another view emerges if we examine the efficacy measure of Eq. (3.2) from the perspective of the time dependence of the variable $G(t) \equiv g(Z(t))$, rather than in terms of Z itself. The basic idea is to allow G to have its own dynamics, e.g., an exponential approach to the nonequilibrium steady state implied by Eq. (3.5) and the relation $dZ/dt \to 0$. Then,

$$G(t) = \frac{\beta}{\alpha}(1 - \exp[-\alpha t]) \rightarrow \frac{\beta}{\alpha} \equiv G_\infty$$
$$dG/dt = \Omega(G(t), t) = (\beta - \alpha G(t)) \quad (3.19)$$

We can, after some considerable algebra, write a stochastic differential equation for the efficacy function F as

$$dF_t = (dF/dt \times dG/dt)dt + \sigma_F F_t dW_t =$$
$$F_t(\frac{K}{G_t^2}\Omega(G_t, t)dt + \sigma_F F_t dW_t =$$
$$\frac{F_t}{K}(\log(F_t)^2(\beta + \frac{\alpha K}{\log(F_t)})dt + \sigma_F F_t dW_t \quad (3.20)$$

Applying the Ito chain rule to F_t^2 in the last expression produces the necessary condition for the second-order stability as

$$\frac{\alpha^2 K}{2\beta} = \frac{\alpha K}{2G_\infty} > \sigma_F^2 \quad (3.21)$$

where, it is important to recognize, G is monotonic increasing in Z, and σ_F is specific to the configuration under study.

Another critical stability condition emerges by recognizing that the G_j may themselves be subject to stochastic variation. From Eq. (3.19), the appropriate SDE is then

$$dG_t = (\beta - \alpha G_t)dt + \sigma_G G_t dW_t \quad (3.22)$$

where the various σ_G can be expected to have markedly different properties than the σ_F above.

Applying the Ito chain rule to G_t^2, the condition for stability in variance of the G-factors, becomes

$$\alpha > \frac{\sigma_G^2}{2} \quad (3.23)$$

Solving Eqs. (3.21) and (3.23) for α, we obtain the necessary condition for total system second-order stability as

$$\frac{K}{G_\infty} > \left(2\frac{\sigma_F}{\sigma_G}\right)^2 \quad (3.24)$$

This implies that high levels of needed resources—as measured by G_∞—must be accompanied by a high rate of reaction—K—for stability in variance, at some given value of σ_F.

An adversary in command of time and place of confrontation, and/or in the degree of "noise"—σ_F—will likely impose violation of this condition when one's need for resources is sufficiently great.

Taking an operational level perspective—a qualitatively different configuration—in which there is a synergistic chain-of-efficacy as in Eq. (3.14), a certain initial simplicity emerges. Then the individual terms in F are replaced by the scalarized product $\hat{F} = \log(\Pi_j F_j)$ so that

$$\hat{F} = \sum_j \frac{-K_j}{g_j(Z_j)} \equiv \sum_j \frac{-K_j}{G_j} \tag{3.25}$$

By observation, and as expected, stability in variance of the entire operational configuration—the full chain—is determined by the stochastic properties of the least stable individual component.

This is not particularly reassuring.

Welcome to the dynamics of defeat and humiliation in the Forever War.

3.6 Discussion

Again, the mathematical ecologist Pielou ([8], p. 102) cautions that models such as we have employed here are impossible to adequately test, and concludes that "In my opinion the usefulness of models is great [however] it consists *not in answering questions but in raising them*. Models can be used to inspire new field investigations and these are the only source of new knowledge as opposed to new speculation."

The most obvious inference raised by this analysis is that, in terms of the resource rate measure $Z = M \times C \times H$, a contending agent "weak" in the rate of material supply M may be relatively stronger in terms of internal bandwidth C, situational awareness/intelligence bandwidth H. In addition, if able to define the pace of conflict, such an agent can build up resources according to its own timetable, taking whatever time is needed to ensure serial local superiority.

A further degree-of-freedom emerges from an optimization study. Given a (possibly precarious) balance in the composite measure Z between a contending agent rich in material resources and another rich in "information" resources, the ability to impose a shadow price index λ appears as an independent factor.

Recent history indicates that agents both rich in "information resources", and able to determine the place and pace of conflict, will most likely also be able to impose a brutal shadow price on an opponent.

Consonance of the cognition approach with the Clausewitz temperature analysis of the control theory section in Chap. 1 is obtained if we replace Γ in Eq. (1.5) with λ_F as in Eq. (3.13).

Then, a sufficient shadow price index imposed by an adversary can drive a low enough value of the "temperature" measure $\mathscr{T}$ to cause control failure, in addition to the condition of Eq. (3.8) showing how failure to provide essential resources at a sufficient rate leads to cognitive failure.

Over longer, strategic, timescales, such failure—or at least transition to a markedly different coevolutionary regime—may be by the punctuated evolutionary mechanisms of Chap. 1, involving the strategic measure τ rather than the tactical control theory version $\mathscr{T}$.

Equations (3.14) through (3.18) extend something of the argument to the operational level, assuming a chain-link model of units acting under constraints of both resources and time, and Eqs. (3.19)–(3.25) extend these ideas to the Forever War.

Although the analysis has been couched in military terms, there are more general implications. The living state is particularly marked by cognition (in the sense of [9]) at every scale and level of organization. Pathogen and parasite populations must contend with active immune response. The outcome, at the individual level, depends in part on the ability of pathogens or parasites to evade immune attack long enough for reproduction and spread. The most successful such intruders do not necessarily kill their hosts, but engage in a population ecology balance permitting transmission between hosts. Rapid "defeat" of a host—sudden death—may preclude population survival of the pathogen or parasite. Indeed, adaptation of hosts to parasites and pathogens converges on a kind of stable mutualism. In fact, the human genome (and, consequently, human evolution), is seeded with viral DNA in a kind of ultimate mutualism.

In the domain of public health, contending institutional agents and associated populations contest power relations, determining the quality of the living and working conditions that drive patterns of chronic and infections disease within human populations ([10–13], and references therein).

The gods, however, are not always on the side of the big battalions, and this can cut many ways.

3.7 Chapter Appendix: A Time-Constrained Operational Model

For completeness, we consider a two-stage hierarchical model constrained by time, not by the availability of resources. The time development of the resource Z_j is taken as

$$Z_j = \frac{b_j}{a_j}(1 - \exp(-a_j T_j)) \tag{3.26}$$

but under the time constraint

$$T = \sum_j T_j \tag{3.27}$$

The Lagrangian is then

$$L = \sum_j -\frac{K_j}{g_j(Z_j(T_j))} + \beta(T - \sum_j T_j) \tag{3.28}$$

where, again, $g(Z)$ is given by Eq. (3.6). Then, omitting the j subscripts,

$$\partial L/\partial T = \frac{K C_1 b e^{-aT} a^2}{\left(-C_1 b e^{-aT} + C_1 b + C_2 a\right)^2} \times \left(\mathrm{W}\left(\frac{-C_1 b e^{-aT} + C_1 b + C_2 a}{a}\right)\right)^2 \times$$
$$\left(1 + \mathrm{W}\left(\frac{-C_1 b e^{-aT} + C_1 b + C_2 a}{a}\right)\right)^{-1} = \beta \qquad (3.29)$$

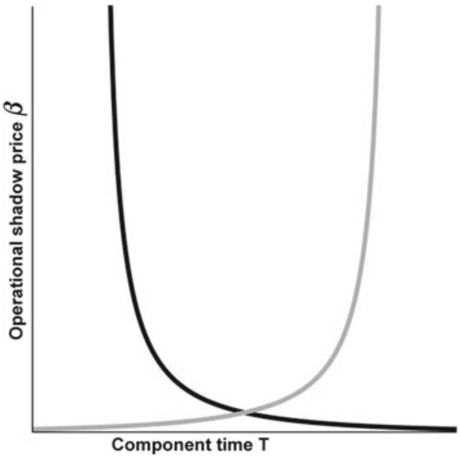


Fig. 3.4 Optimization of a two-level hierarchy having the same driving parameter values. Here, optimization occurs at $T_1 = T_2 = T/2$, assuming β is a floating "undetermined multiplier"

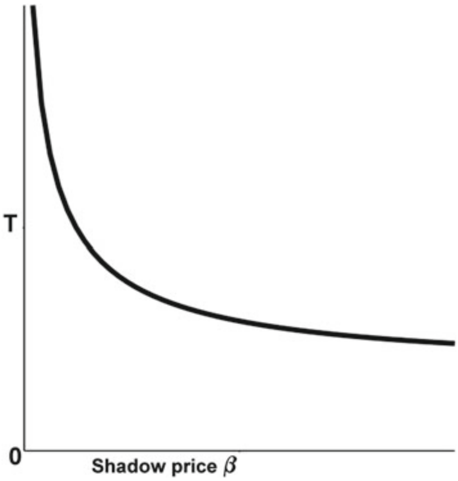


Fig. 3.5 It is possible to solve the relations of Fig. 3.4 for the total time T in terms of β, then seen, not as an 'undetermined multiplier, but as a shadow price imposed by environmental or adversarial factors or actions. Rising β flattens the total allowed response time of Eq. (3.27) to the limit condition of Eq. (3.26), raising the risk of system failure. However, sufficiently low T in this model, as driven by β, triggers a phase transition

subject to the condition

$$\frac{-C_1 b e^{-aT} + C_1 b + C_2 a}{a} > -e^{-1} \tag{3.30}$$

For a two-component command/response system, having the time partition T_1, $T_2 = T - T_1$, the shadow price relation is shown in Fig. 3.4. For both subsystems, $a = b = K = C_1 = 1$ and $C_2 = -0.75$. The horizontal axis represents the component times T_j, and the vertical the value of the “undetermined multiplier” β. Only for a single value of β is there an “optimal” solution, here $T_j = T/2$. At higher values of β—imposed by an unfriendly environment and/or a skilled adversary—T is forced sharply downward and the system risks failure. More explicitly, T can be solved for in terms of the shadow price β, as shown in Fig. 3.5. Rising β flattens the total available system time $T = T_1 + T_2$. Sufficiently large β may, in fact, trigger a phase transition to complex dynamics, in this model.

References

1. Bin Ladin, U. 2004. https://www.aljazeera.com/archive/2004/11200849163336457223.html
2. Schreiber, T. 1964. Note on the combat value of intelligence and command control systems. *Operations Research* 12: 507–510.
3. Kress, M., and R. Szechtman. 2009. Why defeating insurgencies is hard: The effect of intelligence in countreinsurgency operations—a best-case scenario. *Operations Research* 57: 578–585.
4. Wallace, R. 2017. *Computational Psychiatry: A systems biology approach to the epigenetics of mental disorders*. New York: Springer.
5. Feynman, R. 2000. *Lectures on Computation*. Boulder CO: Westview.
6. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
7. Jin, H., Z. Xu, and X. Zhou. 2008. A convex stochastic optimization problem arising from portfolio selection. *Mathematical Finance* 18: 171–183.
8. Pielou, E.C. 1977. *Mathematical Ecology*. New York: Wiley.
9. Atlan, H., and I. Cohen. 1998. Immune information, self-organization, and meaning. *International Immunology* 10: 711–717.
10. Wallace, D., and R. Wallace. 2018. *Right-to-work laws and the crumbling of American public health*. New York: Springer.
11. Wallace, R.G. 2016. *Big farms make big flu: Dispatches on infectious disease, agribusiness, and the nature of science*. New York: Monthly Review Press.
12. Wallace, R.G., and R. Wallace (eds.). 2016. *Neoliberal ebola: Modeling disease emergence from finance to forest and farm*. New York: Springer.
13. Wallace, R., L. Fernando Chaves, L. Bergmann, C. Ayres, L. Hogerwerf, R. Kock, and R.G. Wallace. 2018. *Clear-cutting disease control: Capital-led deforestation, public health austerity, and vector-borne infection*. New York: Springer.

Chapter 4
The Albigensian Ground State

Caedite eos. Novit enim Dominus qui sunt eius.

—Arnaud Amalric

4.1 Introduction

Failure of institutional cognition on Clausewitz landscapes can take many forms. The previous chapter examined ant-on-a-hotplate asymmetric conflict. A related, and particularly chilling, example is the collapse into a "ground state" for which all possible targets are enemies. From the Albigensian Crusade to My Lai and Srebrenica, examples abound.

Conversely, another possible ground state for individual or institutional cognition is blanket denial, for example, MacArthur's systematic failure to recognize the reality of intelligence reports after the Inchon landings and the beginnings of his infamous March to the Yalu. Again, examples abound.

We explore something of the underlying mechanisms using recently developed models of multilevel cognition.

4.2 The Model

As Atlan and Cohen [1] argue, the essence of cognition is the choice by an entity of an action in response to some set of environmental signals, taken from a larger set of actions available to it. Choice reduces uncertainty, and the reduction of uncertainty implies the existence of an information source "dual" to the cognitive process.

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_4

Information sources have "grammar" and "syntax" that differentiate high from low probability sequential paths of action [2, 3].

In consequence, multilevel, multi-scale institutional cognition under conflict involves a network of lower level information sources dual to individual cognitive modules whose probability of interaction can be measured by the strength of the crosstalk—information exchange—between them.

For a random network of cognitive modules, it is well known that a "giant component" involving most of the network nodes emerges in a punctuated manner once the strength of interaction exceeds a well-defined threshold. In contrast, a star-of-stars of stars network has a zero threshold: everything is connected. Network topology counts and changing topology can be used to tune cognitive system network sensitivities.

A random network analysis assumes a variable average number of fixed-strength linkages between the information sources composing the network nodes. Mutual information, by contrast, can continuously vary in magnitude, suggesting the need for a parameterized renormalization. The modular network structure linked by crosstalk has a topology depending on the degree of interaction. The central argument is as follows.

Define an interaction parameter, ω, as a real number ≥ 0, and examine the geometric structures defined by linkages set to zero if crosstalk mutual information is less than that value, and set to unity if greater than ω. A given ω will define a regime of giant components of network elements—cognitive submodules—linked by mutual information greater than or equal to it. This should be recognized as a standard renormalization transformation.

The argument can be inverted. Some given topology for the giant component will define a critical value ω_C such that network elements interacting by crosstalk mutual information less than this value will be unable to participate, that is, will be locked out and not "consciously" perceived on the time scale of the shifting "spotlight" global broadcast of interest. As a result, ω is a tunable, syntactically dependent, detection limit that depends on the instantaneous topology of the corresponding shifting, tunable, giant component of linked cognitive modules defining the dynamic global broadcasts required by no-free-lunch constraints. Wallace [2] shows how this argument can be extended to a multiple set of simultaneous global broadcasts, characterized by a vector $\Omega = (\omega_1, ..., \omega_m)$.

All global broadcast systems, because of their necessary dynamic flexibility, are inherently unstable in the control theory sense and require fairly draconian regulation, as constrained by the Data Rate Theorem. Failure of regulation under the uncertainties of friction and the fog-of-war is the essence of the pathological dynamic studied here, and we further parse the mechanism explored in Fig. 1.3.

An almost-classic phase transition model for multilevel, multi-scale cognitive structures emerges if it is possible to identify equivalence classes of a system's developmental pathways, for example "functional" and "doomed". This allows the definition of a symmetry groupoid for the developmental process. Again, a groupoid is a generalization of an algebraic group in which a product is not necessarily defined between each element. The simplest example is a disjoint union of separate groups, but sets of equivalence classes also define a groupoid. See the Mathematical Appendix for details, and for a more complete exploration of symmetry in information theory.

Another "free energy" can be defined that is liable to an analog of Landau's classical spontaneous symmetry breaking, in the Morse Theory sense. Again, see the Mathematical Appendix for an outline of Morse Theory.

Under symmetry breaking, higher "temperatures" are associated with more symmetric higher energy states in physical systems. Here, for the control of cognitive processes, decline in $|\Omega|$ can result in sharply punctuated collapse from higher to lower symmetry states, resulting in serious cognitive failure.

This is worked out as follows.

Given a "dual" information source associated with the inherently unstable cognitive system of interest, an equivalence class algebra can be constructed by choosing different system origin states and defining the equivalence of subsequent states at a later time by the existence of a high probability path connecting them to the same origin state. Disjoint partition by equivalence class, analogous to orbit equivalence classes in dynamical systems, defines a symmetry groupoid associated with the cognitive process.

The equivalence classes across possible origin states define a set of information sources dual to different cognitive states available to the inherently unstable cognitive system or coalition of systems of interest. These create a large groupoid, with each orbit corresponding to an elementary "transitive" groupoid whose disjoint union is the full groupoid. Each subgroupoid is associated with its own dual information source, and larger groupoids must have richer dual information sources than smaller.

Take X_{G_i} as the system's dual information source associated with groupoid element G_i. A Morse Function can now be built using the temperature analog $|\Omega|$.

Take $H(X_{G_i}) \equiv H_{G_i}$ as the Shannon uncertainty of the information source associated with the groupoid element G_i. It is possible to define a Boltzmann-like pseudoprobability as

$$P[H_{G_i}] \equiv \frac{\exp[-H_{G_i}/|\Omega|]}{\sum_j \exp[-H_{G_j}/|\Omega|]} \tag{4.1}$$

The sum is over the different possible cognitive modes of the full system.

The "free energy" Morse Function $\hat{F}$ is then defined as

$$\begin{aligned} \exp[-\hat{F}/|\Omega|] &\equiv \sum_j \exp[-H_{G_j}/|\Omega|] \\ \hat{F} &= -|\Omega| \log[\sum_j \exp[-H_{G_j}/|\Omega|]] \end{aligned} \tag{4.2}$$

Given the underlying groupoid generalized symmetries associated with high-order cognition, as opposed to simple control theory, it is possible to apply a version of Landau's symmetry-breaking approach to phase transition [4]. The shift between such symmetries should remain highly punctuated in the "temperature" $|\Omega|$, but in the context of what is likely to be far more complicated groupoid rather than group symmetries.

Arguing by abduction from physical theory, there should be only a few possible phases, with sharp and sudden transitions between them as $|\Omega|$ decreases.

An essential inference surrounds the "ground state" of such a sequence of phase transitions. In a military setting, under the pressures of protracted stress and uncertainty, as well as of time and resource constraints—fog-of-war and friction—that ground state becomes something much like the Albigensian Crusade injunction to "kill everyone and let God sort them out". WWII case histories range from Babi Yar to Zywocice. US examples include the No Gun Ri bridge incident, and the massacre at My Lai. The Patriot missile fratricides of the First Gulf War provide a man/machine "cockpit" example.

It is possible to study the stability of such a state.

The first approach is via the treatment of Eqs. (1.6), (1.7) and Fig. 1.4. That is, the cognitive function index $|\Omega|$ is taken as subject to blunting by noise, leading to a first-order stochastic differential equation having the form

$$d|\Omega|_t = \mu|\Omega|_t(1 - \frac{|\Omega|_t}{K})dt + \sigma|\Omega|_t dW_t \tag{4.3}$$

where, as before, $K > 0$ is the limiting value of $|\Omega|_t$, σ the "noise" amplitude, and dW_t represents Brownian fluctuation. Again, following the argument of Eq. (1.7) and Fig. 1.4, if $\sigma^2/2 > \mu$, then $|\Omega|_t$ converges in probability to zero, driving the system into the Albigensian ground state.

An inverse and very general approach leading to the same result is also possible. We invoke a multidimensional parameter vector X characterizing the *deviation* of the cognitive system of interest from the pathological ground state. The free energy analog $\hat{F}$ from Eq. (4.2) allows definition of an "entropy" as

$$S = \hat{F}(X) - X \cdot \nabla_X \hat{F} \tag{4.4}$$

It is now possible to write a first-order nonequilibrium thermodynamic "Onsager" equation describing system dynamics in terms of the gradients of S by X that will have the generic form

$$dX_t = g(X_t, t)dt + h(X_t, t)dW_t \tag{4.5}$$

where dW_t is taken as Brownian white noise and corrective doctrine implies that $dX/dt = g(X_t, t)$ has the solution $|X(t)| \to \infty$. This is the context imposed by the Laws of Land Warfare. However, it is always possible to invoke the arguments of Appleby et al. [5] that, for a one-dimensional system, under wide conditions, a function $h(X_t, t)$—"noise"—can always be found that "stabilizes an unstable equilibrium", i.e., locks in the Albigensian ground state for which all possible targets are classified as enemies.

4.3 Discussion

Massive "collateral damage", it seems, is something of a norm on Clausewitz landscapes, often having long-term policy consequences that may obviate short-term "victories". In particular, conflict systems seem to rewind, almost in proportion to the sum total of direct and collateral damages. The progression from WW I to WW II seems a canonical example of the kind of "resilience" transformation described by Holling [6], like the permanent eutrophication of a lake from a massive spilling of blood.

The reader can, no doubt, provide other case histories.

The question then arises regarding the time scale on which there is success in conflict, given the path-dependent trajectory of historical development.

Are they now all Pyrrhic victories?

Conversely, and sure to contribute to the recurrence of armed confrontation, is the cognitive collapse to blanket denial of reality. Beyond such matters as MacArthur's March to the Yalu, or the German General Staff's response to the catastrophe at Kursk (there was a Colonel's revolt in 1944 instead of a General's revolt in 1943 that would have saved a million lives) are the many examples from public health. These range from administrative attempts to cover up the early stages of the Great Plague in London, to similarly slow response to the influenza pandemic of 1918, and failure to confront the AIDS pandemic in the USA. Currently, public health denial in the USA focuses on ignoring the role of agribusiness factory farming in the likelihood of another 1918 pandemic [7], as well, of course, as persistent right-wing madness over the reality of human-caused global warming.

As described, ground state cognitive collapse, at all scales and levels of organization, can be pathologically stable, leading to cycles of conflict or catastrophe, as long as fundamental causes remain unaddressed.

Acknowledgements The author thanks Dr. D.N. Wallace for suggesting persistent denial as another ground state example.

References

1. Atlan, H., and I. Cohen. 1998. Immune information, self-organization, and meaning. *International Immunology* 10: 711–717.
2. Wallace, R. 2012. Consciousness, crosstalk, and the mereological fallacy: An evolutionary perspective. *Physics of Life Reviews* 9: 426–453.
3. Wallace, R. 2017. *Computational Psychiatry: A systems biology approach to the epigenetics of mental disorders*. New York: Springer.
4. Pettini, M. 2007. *Geometry and topology in Hamiltonian dynamics and statistical mechanics*. New York: Springer.
5. Appleby, J., X. Mao, and A. Rodkina. 2008. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control* 53: 683–691.
6. Holling, C. 1973. Resilience and stability of ecological systems. *Annual Reviews of Ecological Systematics* 4: 1–23.

7. Wallace, R., and R.G. Wallace. 2015. Blowback: new formal perspectives on agriculturall driven pathogen evolution and spread. *Epidemiology and Infection* 143: 2068–2080.
8. Hanley, C., M. Mendoza, and S. Choe. 2001. *The bridge at No Gun Ri: A hidden nightmare from the Korean War*. New York: Henry Holt and Company.

Chapter 5
Can There Be "Third Stream" Doctrines?

The serious trouble begins when Americans have to interact with alien societies in unfamiliar terrain...
A certainty of material superiority can breed overconfidence and minimize incentives to outthink the enemy...
A practical consequence of the messianic, crusading dimension to American strategic behavior, is a self-righteousness that is not friendly, or even receptive, to unwelcome cultural and political facts...
The country's culture simply did not register the unwanted Vietnam experience. Cultures, including strategic cultures, are capable of ignoring what they wish to ignore.

—C.S. Gray

冲突像水一样发展和流动

5.1 Introduction

Mao Tse-Tung ([1], pp. 227–228) writes

> From the particular characteristics of war there arise a particular set of organizations, a particular series of methods and a process of a particular kind. The organizations are the armed forces and everything that goes with them. The methods are the strategy and tactics for directing war. The process is the particular form of social activity in which the opposing armed forces attack each other or defend themselves against one another, employing strategy and tactics favourable to themselves and unfavourable to the enemy. Hence, war experience is a particular kind of experience. All who take part in war must rid themselves of their customary ways and accustom themselves to war before they can win victory.

The Western evolutionary anthropologist Robert Boyd might be seen as providing a devastating riposte: "Culture is as much a part of human biology as the enamel

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_5

on our teeth". Our "customary ways" are deeply burned into our individual and collective patterns of behavior.

Indeed, Mao describes something of this ([1], p. 195):

> Epistemologically speaking, the source of all erroneous views on war lies in idealist and mechanistic tendencies on the question of war. People with such tendencies have a subjective and one-sided approach to problems. They either indulge in groundless and purely subjective talk, or, basing themselves upon a single aspect or a temporary manifestation, magnify it with similar subjectivity into the whole of the problem. But there are two categories of erroneous views, one comprising fundamental and therefore consistent errors which are hard to correct, and the other comprising accidental and therefore temporary errors which are easy to correct. Since both are wrong, both need to be corrected. Therefore, only by opposing idealist and mechanistic tendencies and taking an objective and all-sided view in making a study of war can we draw correct conclusions on the question of war.

These remarks can be viewed, among other things, as an implicit criticism of Western doctrine, of an "analytic" perspective that often bases itself on single aspects or temporary manifestations, magnifying them into the whole of the problem. This is the famous Mereological Fallacy, imputing the whole from only a part, and characterizes a fundamental difference between "Western" and "East Asian" modes of perception.

At the tactical level, Mao ([1], pp. 79–80) comments

> Why is it necessary for the commander of a campaign or a tactical operation to understand the laws of strategy to some degree? Because an understanding of the whole facilitates the handling of the part, and because the part is subordinate to the whole. The view that strategic victory is determined by tactical successes alone is wrong because it overlooks the fact that victory or defeat in a war is first and foremost a question of whether the situation as a whole and its various stages are properly taken into account. If there are serious defects or mistakes in taking the situation as a whole and its various stages into account, the war is sure to be lost.

More recently, the Chinese military strategist Jie [2] described the incorporation of "information" perspectives into contemporary PLA doctrine as follows:

> Today, the technological state of human society is undergoing a transition from an industrialized to an informationized age, [entraining the practice of war] from the mechanized to the informationized... [Experts] hold [that]... with the transformation from an industrial society to an information society, informationized war is beginning to replace mechanized war... a long-term change that will lead to qualitative changes in... war... [centering on the structure of the campaign]...
>
> ...[T]he so-called informationized war is, as human society entered the Information Age, a new type of war carried out on an informationized battlefield, with information as the leading factor, with informationized armed forces as the main strengths, with informationized weapons and equipment as the main operational tools, and with informationized operations as the main operational form... [I]nformation not only has become the leading factor in social development, but also has become the leading factor in gaining victory in war...
>
> The battlefields for informationized operations not only include the tangible battlefields composed of the land, sea, air, space and network EM physical fields, but also include the intangible battlefields composed of the information field and cognitive field. Informationized operations not only have covered the military field, but also involve the political, economic,

> diplomatic, legal, and cultural fields. This requires conducting [operations[.. in the land, sea, air, space, network EM, and cognitive multidimensional space; conducting integrated-whole operations in all dimensions of the battlefield; and thus bringing into play the maximum operational functions of entire systems...

Jullien [3] describes Chinese views on strategy in these terms:

> For something to be realized in an effective fashion [in the Chinese view], it must come about as an effect. It is always through a process (which transforms the situation), not through a goal that leads (directly) to action, that one achieves an effect, a result...
>
> Any strategy thus seems, in the end, to come down to simply knowing how to implicate an effect, knowing how to tackle a situation upstream in such a way that the effect flows "naturally" from it...

The fundamental difference between East Asian and Western perception and reasoning has been the subject of considerable study. For example, Nisbett et al. [4], following in a long line of research [5, 6], review an extensive literature on empirical studies of basic cognitive differences between individuals raised in East Asian and Western cultural heritages, which they characterize, respectively, as "holistic" and "analytic".

They find

1. Social organization directs attention to some aspects of the perceptual field at the expense of others.
2. What is attended to influences metaphysics.
3. Metaphysics guides tacit epistemology, that is, beliefs about the nature of the world and causality.
4. Epistemology dictates the development and application of some cognitive processes at the expense of others.
5. Social organization can directly affect the plausibility of metaphysical assumptions, such as whether causality should be regarded as residing in the field vs. in the object.
6. Social organization and social practice can directly influence the development and use of cognitive processes such as dialectical versus logical ones.

They conclude that tools of thought embody a culture's intellectual history, that tools have theories built into them, and that users accept these theories, albeit unknowingly, when they use these tools.

In addition, Masuda and Nisbett [7] find that research on perception and cognition suggests that whereas East Asians view the world holistically, attending to the entire field and relations among objects, Westerners view the world analytically, focusing on the attributes of salient objects. Compared to Americans, East Asians were more sensitive to contextual changes than to focal object changes, These results suggest that there can be cultural variation in what may seem to be basic perceptual processes.

Similarly, Nisbett and Miyamoto [8] argue that fundamental perceptual processes are influenced by culture. These findings establish a dynamic relationship between the cultural context and perceptual processes. They suggest that perception can no longer be regarded as consisting of processes that are universal across all people at all times.

Wallace [9] explores analogous dynamics involving inattentional blindness and culture.

The obvious inference is that such basic biological differences will also be expressed at the institutional level under conditions of conflict, although perhaps via different modalities at and across different scales and levels of organization.

Somewhat surprisingly, two different poles of perception can be derived from the asymptotic limit theorems of information theory, a matter we explore in the next section. But this is not really the point: we already know this. The point is that there may well be a third—intermediate—stream possible in the Lamarckian evolution of strategic conflict.

Such a third stream would blindside both "analytic" and "holistic" modalities of military doctrine.

5.2 The Analytic and Holistic Streams of Military Doctrine

Extended conflict can be viewed as the transmission of a particular kind of "message" by one party to another using a spectrum of means that may include, but not be limited to, armed combat. The message is often "encoded" physically in ordinary time and space, but also as appropriate to "cognitive space" and, more recently, "cyberspace" as well. Conflict is, then, "communicated" to an opponent through a literally compelling series of multimedia events. In getting the central "message" of conflict through to an opponent, it is important that the structure of that message remains fixed and be "received" in the form transmitted. This is, in a sense, an inverse problem of the typical information theory problem, where channels are viewed as fixed, and messages are tuned around the channel properties to achieve the maximum possible transmission rate. We focus on making sure the message is transmitted intact.

This can, somewhat surprisingly, at least in theory, actually be done. An East Asian metaphor might be "Cultivate the channel of the stream so water flows how and where you want".

Some development is required, taken from [10].

Messages from an information source, seen as symbols x_j from some alphabet, each having probabilities P_j associated with a random variable X, are "encoded" into the language of a "transmission channel", a random variable Y with symbols y_k, having probabilities P_k, possibly with error. Someone receiving the symbol y_k then retranslates it (without error) into some x_k, which may or may not be the same as the x_j that was sent.

More formally, the message sent along the channel is characterized by a random variable X having the distribution

$$P(X = x_j) = P_j,\ j = 1, ..., M.$$

The channel through which the message is sent is characterized by a second random variable Y having the distribution

$$P(Y = y_k) = P_k, k = 1, ..., L.$$

Let the joint probability distribution of X and Y be defined as

$$P(X = x_j, Y = y_k) = P(x_j, y_k) = P_{j,k}$$

and the conditional probability of Y given X as

$$P(Y = y_k|X = x_j) = P(y_k|x_j).$$

Then the Shannon uncertainty of X and Y independently, and the joint uncertainty of X and Y together, are defined respectively as

$$H(X) = -\sum_{j=1}^{M} P_j \log(P_j)$$

$$H(Y) = -\sum_{k=1}^{L} P_k \log(P_k)$$

$$H(X, Y) = -\sum_{j=1}^{M}\sum_{k=1}^{L} P_{j,k} \log(P_{j,k}) \tag{5.1}$$

The *conditional uncertainty* of Y given X is defined as

$$H(Y|X) = -\sum_{j=1}^{M}\sum_{k=1}^{L} P_{j,k} \log[P(y_k|x_j)] \tag{5.2}$$

For any two stochastic variates X and Y, $H(Y) \geq H(Y|X)$, as knowledge of X generally gives some knowledge of Y. Equality occurs only in the case of stochastic independence.

Since $P(x_j, y_k) = P(x_j)P(y_k|x_j)$, then $H(X|Y) = H(X, Y) - H(Y)$.

The information transmitted by translating the variable X into the channel transmission variable Y—possibly with error—and then retranslating without error the transmitted Y back into X is defined as

$$I(X|Y) \equiv H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y) \tag{5.3}$$

See [11] for details. The essential point is that if there is no uncertainty in X given the channel Y, then there is no loss of information through transmission. In general this will not be true, and herein lies the essence of the theory.

Given a fixed "vocabulary", in a large sense, for the transmitted variable X, and a similarly fixed vocabulary and probability distribution for the channel Y, we may vary the probability distribution of X in such a way as to maximize the information sent. The capacity of the channel is defined as

$$C \equiv \max_{P(X)} I(X|Y) \tag{5.4}$$

subject to the subsidiary condition that $\sum P(X) = 1$.

The critical trick of the Shannon Coding Theorem for sending a message with arbitrarily small error along the channel Y at any rate $R < C$ is to encode it in longer and longer "typical" sequences of the variable X; that is, those sequences whose distribution of symbols approximates the probability distribution $P(X)$ above which maximizes C.

If $S(n)$ is the number of such "typical" sequences of length n, then

$$\log[S(n)] \approx nH(X) \tag{5.5}$$

where $H(X)$ is the uncertainty of the stochastic variable defined above. Some consideration shows that $S(n)$ is much less than the total number of possible messages of length n. Thus, as $n \to \infty$, only a vanishingly small fraction of all possible messages are meaningful in this sense. This observation, after some considerable development, is what allows the Coding Theorem to work so well. In sum, the prescription is to encode messages in typical sequences, which are sent at very nearly the capacity of the channel. As the encoded messages become longer and longer, their maximum possible rate of transmission without error approaches channel capacity as a limit. Again, standard references on information theory provide details.

This approach can be, in a sense, inverted to give a "tuning theorem" variant of the coding theorem.

Telephone lines, optical waveguides, and the tenuous plasma through which a planetary probe transmits data to earth may all be viewed in traditional information-theoretic terms as a *noisy channel* around which we must structure a message so as to attain an optimal error-free transmission rate.

Telephone lines, waveguides, and interplanetary plasmas are, relatively speaking, fixed on the timescale of most messages, as are most sociogeographic networks. Indeed, the capacity of a channel, is defined by varying the probability distribution of the "message" process X so as to maximize $I(X|Y)$.

Suppose there is some message X so critical that its probability distribution must remain fixed. The trick is to fix the distribution $P(x)$ but *modify the channel*—i.e., tune it—so as to maximize $I(X|Y)$. The *dual* channel capacity C^* can be defined as

$$C^* \equiv \max_{P(Y), P(Y|X)} I(X|Y) \tag{5.6}$$

But

$$C^* = \max_{P(Y), P(Y|X)} I(Y|X) \tag{5.7}$$

since $I(X|Y) = H(X) + H(Y) - H(X, Y) = I(Y|X)$.

Thus, in a purely formal mathematical sense, *the message transmits the channel*, and there will indeed be, according to the Coding Theorem, a channel distribution $P(Y)$ which maximizes the dual- channel capacity C^*.

One may do better than this, however, *by modifying the channel matrix* $P(Y|X)$. Since

$$P(y_j) = \sum_{i=1}^{M} P(x_i)P(y_j|x_i) \tag{5.8}$$

$P(Y)$ is entirely defined by the channel matrix $P(Y|X)$ for fixed $P(X)$ and

$$C^* = \max_{P(Y), P(Y|X)} I(Y|X) = \max_{P(Y|X)} I(Y|X) \tag{5.9}$$

Calculating C^* requires maximizing the complicated expression $I(X|Y) = H(X) + H(Y) - H(X, Y)$, which contains products of terms and their logs, subject to constraints that the sums of probabilities are 1 and each probability is itself between 0 and 1. Maximization is done by varying the channel matrix terms $P(y_j|x_i)$ within the constraints. This is, in general, a very difficult problem in nonlinear optimization. However, for the special case $M = L$, C^* a solution may be found by inspection:

If $M = L$, then choose $P(y_j|x_i) = \delta_{j,i}$, where $\delta_{i,j}$ is 1 if $i = j$ and 0 otherwise. For this special case

$$C^* \equiv H(X) \tag{5.10}$$

with $P(y_k) = P(x_k)$ for all k. *Information is thus transmitted without error when the channel is constructed to become "typical" with respect to the fixed message distribution* $P(X)$.

If $M < L$ matters reduce to this case, but for $L < M$ information must be lost, leading to Rate–Distortion limitations.

Thus modifying the channel becomes the canonical means of ensuring the transmission of an important message in the manner desired, rather than encoding that message in a "natural" language which maximizes the rate of transmission of information on a fixed channel.

We have examined the two limits in which either the distributions of $P(Y)$ or of $P(X)$ are kept fixed. The first provides the usual Shannon Coding Theorem, and the second a tuning theorem variant, i.e., a tunable, retina-like, Rate–Distortion Manifold, in the sense of [12].

This result is, somewhat surprisingly, essentially similar to Shannon's [13] observation that evaluating the rate–distortion function corresponds to finding a channel that is just right for the source and allowed distortion level.

But again, this is not the point. We already qualitatively understand a good deal of this.

What, then, is the point?

5.3 Third Streams

Recall Jullien's remarks above. One can, apparently, give this metaphor something of a heuristic foundation in information theory: tune the channel. The Western metaphor emerges as "adjust the tactics to the landscape", and Westerners—certainly from Rommel to Boyd—have sometimes been master tacticians.

Strategy, however, is another matter, particularly, at this writing, for the US national security team.

In both cognitive and cyber spaces, it appears that an increasingly important tool for redirecting the stream/channel of conflict may be information itself [2]. A report by the Center for Strategic and Budgetary Assessments ([14], redacted, p. 41) put the problem thus:

> ...[T]here is the issue of Clausewitzian friction. It almost inevitably intrudes at the tactical level of war. But its intrusions tend to be even more consequential at the operational and strategic levels because operational and strategic problems are usually "wicked" ones, meaning that they are ill structured, open-ended, and not amenable to closed, engineering solutions... Witness some of the frictions the Allies and the Germans encountered during the Normandy invasion in June 1944... The problem with friction, of course, is that no conceivable advances in weaponry or technology are capable of eliminating it despite recurring hopes to the contrary. This is because general friction arises from [1] human physical and cognitive limitations, [2] the inherent uncertainties in the information on which actions in war are based, and [3] the structural nonlinearity of combat processes and interactions. The deepest uncertainty affecting U.S.-PRC military competition in the information aspects of war, then, is whether the PLA will be able to do a better job of coping with the frictions of high-tech local wars under informationized conditions than U.S. forces. It may be that the PLA's prescriptive, top-down planning approach and quest for "trumpcard" stratagems will prove to be impediments to the PLA's capacity to deal with friction. But... there is little evidence that the American military is inclined to embrace as holistic and comprehensive an approach to the growing role of information in modern warfare as the Chinese. Insofar as information's future role in war is concerned, it is difficult to avoid the conclusions that the Chinese and American militaries are operating on very different frequencies.

More recently, Allen [15] has explored the Chinese transition from "informationized" to "intelligenceized" military enterprise as follows:

> ...[M]ost of China's leadership sees increased military usage of AI as inevitable and is aggressively pursuing it... Beyond using AI for autonomous military robotics, China is also interested in AI capabilities for military command decisionmaking... Zenh Yi [senior executive in a large Chinese defense company has said that today] "mechanized equipment is just like the hand of the human body. In future, AI systems will be just like the brain of the human body".

Allen goes on to make the point that

> Several months after AlphaGo's momentous March 2016 victory over Lee Sedol, a publication by China's Central Military Commission Joint Operations Command Center argued that Alpha Go's victory "demonstrated the enormous potential of artificial intelligence in combat command, program deduction, and decisionmaking".

The Chinese, we argue above, have some considerable foundation for their basic approach: the "tuning theorem" formulation. On the other hand, Western military practice has much foundation in the "coding theorem". However, and of considerable importance, in marked contrast to the assertions of [2] and to the successes of AlphaGo, a metaholistic perspective regarding conflict on Clausewitz landscapes does not find "information" or machine intelligence to be the universal solvents of current Chinese military thinking. "Information" is not a local stream of the "Flowing Water" of historical Chinese strategy [3]. Machine intelligence is not immune from the perturbations inherent to Clausewitz landscapes of uncertainty, imprecision, and overreach [16]. The perspective is, in a sense, a kind of inverse to the typically Western Mereological Fallacy, i.e., assuming that some few singular features constitute the whole. The relations between information and control under conditions of uncertainty and the frictional distortion of intent are not at all straightforward. Chapter 1 strongly suggests that mastery of information and rapidity of cognition do not imply mastery of control in conflict, just as tactical competence does not imply competence on operational or strategic scales and levels of organization.

US and PRC military enterprises have, it appears, devoted considerable effort into creating "Blue" versus "Red" team exercises, attempting to understand and counter each other's strong points, which are built-in to high level command dynamics. Indeed, who could tell MacArthur not to march helter-skelter along the valleys toward the Yalu after Inchon? Who could ever tell Westmoreland anything? The sequence of German blunders on the Eastern Front comes to mind. PRC forces will have their own cultural constraints, e.g., their 1975 Vietnam debacle.

One inference of our formal analysis is that there should be an "intermediate theorem" involving simultaneous tuning of both the "message" and the "channel" for maximal strategic impact under specific time constraints. The relative effectiveness of this intermediate approach would depend on setting those time constraints—John Boyd fashion—but might often be effective beyond what would be possible from either extreme. This would require the singularly difficult task of learning to act outside of both our own cultural constraints, as well as of those defined by an opponent.

The benefits, however, might be similarly singular.

And that is our central point.

What we suggest, then, goes somewhat beyond current "deception" doctrine that attempts to use an opponent's inherent thought processes as a strategic tool. For example, [17] describes "ambiguity-decreasing deception" in these terms:

> Ambiguity-decreasing deceptions manipulate and exploit an enemy decision maker's preexisting beliefs and bias through the intentional display of observables that reinforce and convince that decision maker that such pre-held beliefs are true. Ambiguity-decreasing deceptions cause the enemy decision maker to be especially certain and very wrong. Ambiguity-decreasing deceptions aim to direct the enemy to be at the wrong place, at the wrong time, with the wrong equipment, and with fewer capabilities. Ambiguity-decreasing deceptions are more challenging to plan because they require comprehensive information on the enemy's processes and intelligence systems. Planners often have success using these deceptions with strongminded decision makers who are willing to accept a higher level of risk...
>
> ...[I]t is generally easier to induce the deception target to maintain a preexisting belief than to deceive the deception target for the purpose of changing that belief... exploit[ing] target biases and the human tendency to confirm exiting beliefs...

> Any bias is potentially exploitable. Most targets are unaware of how deeply their biases influence their perceptions and decisions.

Such operations are not without risk [17]:

> Deceptions may produce unintended, often unwanted consequences. Believing that a threat is real, an enemy can act unpredictably. Proper planning and coordination and knowing the enemy can reduce the chance that deceptions will result in unfavorable action. Successful planners consider second- and third-order effects of the deception plan to mitigate unintended consequences...

Barton Whaley ([18], p. 190) remarks on cultural matters as follows:

> Whenever the target of your deception belongs to a different culture than yours, problems of cross-cultural communication of the intended deception will arise. These communication problems can range from the trivial to the decisive, that is, from minor glitches to complete failure. But wise deceivers will appreciate this problem, try to discover how it is apt to work in specific situations, and plan accordingly. Similarly, the opposing deception analysts will benefit from doing the same.

It may be possible, via "third stream thinking", to act well beyond the culturally and historically driven expectations of an adversary that will be based, largely, on their inherently constrained understanding of one's own culture and history. Gray ([19], p. 6), however, raises a central caution:

> Strategic cultural understanding is difficult to achieve and even more difficult to operationalize. The fact that it is an important concept, robust in its essentials against challenge, is irrelevant. The practical implications of the promotion of culture to intellectual and doctrinal leading edge status may well, indeed probably will, prove to be unduly demanding.

These are deep waters, and even if the analytic tools developed in these essays evolve into statistical tools for data analysis and policy purposes, proper use will remain a highly skilled enterprise indeed.

Something of an example is worked out in the Chapter Appendix, where an attempt is made to explore a technologically driven arms race in the context of dynamics determined, ultimately, by cultural fragmentation.

5.4 Chapter Appendix: Reconsidering Arms Races

The classic model of Richardson [20] describes an arms race between two polities in terms of a coupled set of differential equations. Let X_i be the level of weapons adopted by polity i. σ and ρ are taken as "defence coefficients", α, γ as "fatigue coefficients", and g, h as "grievance coefficients". Richardson's model then has the form

$$\begin{aligned} dX_1/dt &= \sigma X_2 - \alpha X_1 + g \\ dX_2/dt &= \rho X_1 - \gamma X_2 + h \end{aligned} \tag{5.11}$$

Some development shows that there may not be a steady-state solution to these relations, that is, one possible outcome is an explosion, a runaway arms race. Variations of the model abound, incorporating, e.g., time delays, resource constraints, etc.

It is worth noting that for systems in two or more dimensions, like Eq. (5.11), it is always possible to introduce a pattern of "noise" that destabilizes a stable steady state [21].

Here, we take a different tack.

We suppose that, in addition to "Western" and "East Asian" patterns of perception and decision, there are many other "third stream" doctrines possible, characterized by some index j. Each such, including the Western and East Asian modes, will have a distinct cognitive process involving choice, in the context of some global intensity index K of available technologies and related resources. Choice reduces uncertainty, and the reduction of uncertainty implies the existence of an information source H_j. Most simply, we can construct a Boltzmann pseudoprobability as

$$P[H_j] = \frac{\exp[-H_j/K]}{\sum_i \exp[-H_i/K]} \tag{5.12}$$

in the context of the usual groupoid symmetries that characterize cognition (e.g., [9, 10, 12, 22]). There may, in fact, be quite a number of such groupoids, so that j has a large range. It then becomes possible to define a "groupoid free energy" F in terms of the "partition function" denominator of Eq. (5.12) as

$$\exp[-F/K] \equiv \sum_i \exp[-H_i/K] \tag{5.13}$$

See [21] for more details.

We can now apply the groupoid version of Landau's heuristic phase transition argument [23], in the technology rate index K (e.g., some technology index per capita, per dollar, or an appropriate synergism), leading to complicated patterns of punctuated equilibrium phase transition in F as K varies in space, time, and social structure across a landscape of contention.

However, as is well known from the experiences of WWI, WWII and the Cold War, technologies may, themselves, become entrained into—and feed—conflict dynamics.

Suppose there is an explosive proliferation of contending cognitive agents during a protracted conflict, so that the right hand side of Eq. (5.13) can be approximated as an integral, i.e.,

$$\begin{aligned} \sum_i \exp[-H_i/K] &\approx \int_0^\infty \exp[-H/K]dH = K \\ F &\approx -K\log(K) \end{aligned} \tag{5.14}$$

It is then possible to define an entropy-analog, S, as

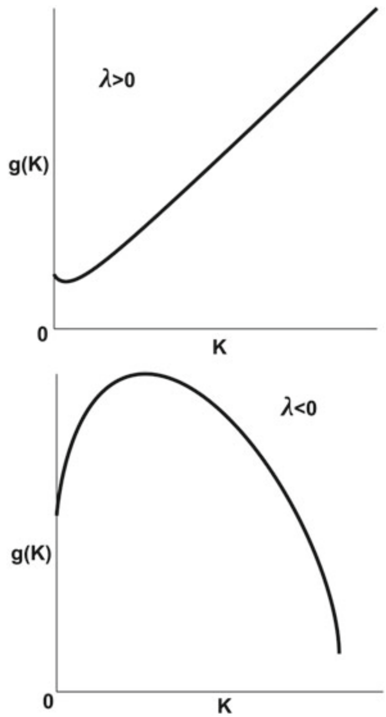


Fig. 5.1 "Cultural" arms race outcomes based on cognitive dynamics. Like Eq. (5.16), a positive value of μ leads to a runaway arms race involving the convergence of technologies. Negative μ leads to complicated dynamics ending in some punctuated "speciation" phase transition

$$S \equiv F(K) - KdF/dK \approx K \tag{5.15}$$

A crude first approximation to the dynamics of K during extended conflict emerges from the standard Onsager nonequilibrium thermodynamics formalism [24] as

$$\begin{aligned} dK/dt &\approx \mu dS/dK = \mu \\ K(t) &\approx \mu t + K_0 \end{aligned} \tag{5.16}$$

where μ is a "cultural diffusion coefficient".

That is, in this model, the technology intensity level K can become self-referential, leading, in the case of positive μ, to an "inverse Cambrian explosion", a technology-driven collapse of possible strategic approaches that may converge on such mirror-image pathologies as Mutual Assured Destruction, the proliferation of multiple-warhead missiles, or time-driven conflict analogs to stock market "flash crashes" under the control of each side's tactical artificial intelligence.

Indeed, the cultural diffusion of the Wehrmacht's superb combined arms methodologies to the Red Army helped ensure subsequent German catastrophes on the Eastern Front, given the Soviet Union's superior manpower, production capabilities, and quite significant U.S. contributions of trucks and aircraft to the Red Army.

By contrast, a locally negative value of μ might well lead to social fragmentation-driven "Cambrian explosions" of proliferating doctrines and conflict methodologies (i.e., asymmetries) that will be examined using another formalism [25] in a subsequent chapter.

A somewhat more sophisticated approach follows the pattern of Sect. 3.3, replacing K in Eq. (5.14) with a function $g(K)$ and imposing a relation $dK/dt = f(K) = dS/dK$. Then the same development leading to Eq. (3.6) gives

$$g(K) = \frac{X}{W(n, X)}, X = -\left(C_1 K - K \int \frac{f(K)}{K} dK - C_2 + \int f(K) dK\right) \tag{5.17}$$

W is the Lambert W-function of order n and the C_i are appropriate constants.

Taking, as above, $f(K) = \pm\mu$ gives the results of Fig. 5.1. In this model $g(K)$ plays the role of the temperature index. $\mu > 0$ drives a dynamic roughly similar to Eq. (5.16), a runaway arms race, while a negative value shows complicated dependence on K, leading to a punctuation event.

References

1. Tse-Tung, Mao. 1963. *Selected military writings of Mao Tse-Tung*. Peking: Foreign Languages Press.
2. Jie, Quao. 2012. *Lecture 10 from lectures on the science of campaigns*. Beijing: Military Science Press.
3. Jullien, F. 2004. *A treatise on efficacy: Between western and Chinese thinking*. Honolulu: University of Hawaii Press.
4. Nisbett, R., K. Peng, C. Incheol, and A. Norenzayan. 2001. Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review* 108: 291–310.
5. Markus, H., and S. Kitayama. 1991. Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review* 98: 224–253.
6. Heine, S. 2001. Self as cultural product: An examination of East Asian and North American selves. *Journal of Personality* 69: 881–906.
7. Matsuda, T., and R. Nisbett. 2006. Culture and change blindness. *Cognitive Science: A Multidisciplinary Journal* 30: 381–399.
8. Nisbett, R., and Y. Miyamoto. 2005. The influence of culture: Holistic vs. analytic perception. *Trends in Cognitive Science* 9: 467–473.
9. Wallace, R. 2007. Culture and inattentional blindenss: A global workspace perspective. *Journal of Theoretical Biology* 245: 378–390.
10. Wallace, R. 2017. *Computational psychiatry: A systems biology approach to the epigenetics of mental disorders*. New York: Springer.
11. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
12. Glazebrook, J., and R. Wallace. 2009. Rate distortion manifolds as model spaces for cognitive information. *Informatica* 33: 309–346.
13. Shannon, C. 1959. Coding theorems for a discrete source with a fidelity criterion, Institute of Radio Engineers International Convention Record, 142–163.
14. Anonymous. 2013. Countering enemy 'informationized operations' in war and peace, Center for Strategic and Budgetary Assessments, HQ0034-09-D-3007-0014 (Redacted litigation release), https://www.esd.whs.mil/.../Litigation%20Release%20-%20Countering%20Enemy%20.

15. Allen, G.C. 2019. *Understanding China's AI strategy: Clues to Chinese strategic thinking on artificial intelligence and national security*. Washington, D.C.: Center for New American Security.
16. Wallace, R. 2018. *Carl von Clausewitz, the fog-of-war and the AI revolution: The real world is not a game of Go*. New York: Springer.
17. U.S. Army. February 2019. Army support to military deception. FM 3-13.4. https://atiam.train.army.mil/catalog/dashboard.
18. Waley, B. 2016. *Practise to deceive: Learning curves of military deception planners*. Annapolis: Naval Institute Press.
19. Gray, C.S. 2006. Out of the wilderness: Prime time for strategic culture. Prepared for the Defense Threat Reduction Agency Advanced Systems and Concepts Office, Contract Nos. DTRA01-03-D-0017, SP0600-04-C-5982.
20. Richardson, L.F. 1960. Arms and insecurity, Boxwood Press.
21. Appleby, J., X. Mao, and A. Rodkina. 2008. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control* 53: 683–691.
22. Wallace, R. 2011. On the evolution of homochriality. *Comptes Rendus Biologies* 334: 263–268.
23. Pettini, M. 2007. *Geometry and topology in Hamiltonian dynamics and statistical mechanics*. New York: Springer.
24. de Groot, S., and P. Mazur. 1984. *Non-Equilibrium thermodynamics*. New York: Dover.
25. Wallace, R. 2013. A new formal perspective on 'Cambrian explosions'. *Comptes Rendus Biologies* 337: 1–5.

Chapter 6
Reconsidering Doctrine and Its Discontents

6.1 Introduction

Military doctrine—learned or inherited, Lamarckian or Darwinian—provides a backbone from which, over space, time, and social structure, operational, and eventually, tactical scale activities, are expressed. In a sense, this construct is the "body" that engages in actions driven by a strategy that may itself be another expression of doctrine. Doctrine tells how to fight a war at the different scales and levels of organization of the confrontation. Here, we examine the coupling of doctrine to tactical, operational, and strategic activities as analogous to gene expression in organisms. In itself, this is not a particularly deep or surprising perspective. What we have developed, however, is a formal framework for its exposition.

The similarity of the use of doctrine to gene expression is less artificial than it seems. The development of an organism takes place according to a complex template that includes, not only a rigidly patterned time sequence of events—growth—but is also affected by "epigenetic" environmental cues of various forms that are either inherited from recent experience within a reproducing population, or carry information from the local, embedding, environment.

Some detail on gene expression.

Following Wallace [1], a cognitive paradigm for gene expression emerges in which contextual factors determine the behavior of a "reactive system", not some deterministic, or even a simple stochastic, mechanical process [2–6].

O'Nuallain [7], for example, puts gene expression directly in the realm of complex linguistic behavior, for which context imposes meaning. He finds the analogy between gene expression and language production useful both as a research paradigm and also, given the relative lack of success of natural language processing by computer, as a cautionary tale for molecular biology. As described in Sect. 4.2, a relatively simple model of cognitive process as an information source permits use of Dretske's [8] insight that any cognitive phenomenon must be constrained by the limit

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_6

theorems of information theory, in the same sense that sums of stochastic variables are constrained by the Central Limit Theorem.

The approach is consistent with recent theory and observation in epigenetics and epigenetic epidemiology. Jablonka and Lamb [9, 10] argue that information can be transmitted from one generation to the next in ways other than through the base sequence of DNA. It can be transmitted through cultural and behavioral means in higher animals, and by epigenetic means in cell lineages. All of these transmission systems allow the inheritance of environmentally induced variation. Such Epigenetic Inheritance Systems are the memory systems that enable somatic cells of different phenotypes but identical genotypes to transmit their phenotypes to their descendants, even when the stimuli that originally induced these phenotypes are no longer present.

The epigenetic perspective has received much empirical confirmation [11–14].

Foley et al. [15] argue that epimutation is estimated to be 100 times more frequent than genetic mutation and may occur randomly or in response to the environment. Periods of rapid cell division and epigenetic remodeling are likely to be most sensitive to stochastic or environmentally mediated epimutation. Disruption of epigenetic profile is a feature of most cancers and is speculated to play a role in the etiology of other complex diseases including asthma, allergy, obesity, type 2 diabetes, coronary heart disease, autism spectrum and bipolar disorders, and schizophrenia.

More to our particular perspective, Scherrer and Jost [16, 17] explicitly invoke information theory in their extension of the definition of the gene to include the local epigenetic machinery, a construct they term the “genon”: coding information is not simply contained in the coded sequence, but is, in their terms, *provided by* the genon that accompanies it on the expression pathway and controls in which peptide it will end up. In their view the information that counts is not about the identity of a nucleotide or an amino acid derived from it, but about the relative frequency of the transcription and generation of a particular type of coding sequence that then contributes to the determination of the types and numbers of functional products derived from the DNA coding region under consideration.

The genon, as Scherrer and Jost describe it, is precisely a localized form of global broadcast linking cognitive regulatory modules to direct gene expression in producing the great variety of tissues, organs, and their linkages that comprise a living entity.

“Epigenetic” signals, in our context, affect the “regulatory machinery” by which doctrine is expressed in conflict.

One formal tool for understanding phenomena that “provide” information—that are information sources—in the context of noise is the Rate–Distortion Theorem.

Chapter 3 introduced a composite “index of essential resources” delivered at some rate Z, under burdens of attrition and the relevance of information. Typically, this would be a scalar function of personnel, material supplies, internal communication bandwidth, and intelligence bandwidth that characterizes the risks faced and opportunities available. It might be constructed using the nonsymmetric matrix method used in Eqs. (1.3)–(1.5). In the biological context, Wallace [18] has identified Z as the rate at which metabolic free energy is provided during development. Details will vary according to the underlying nature of the Clausewitz landscape under study. Certainly, the unique characters of air, ground, sea/riverine and mixed-modal operations come to mind.

6.2 The Model

Suppose that essential resources are provided at some rate Z. Let $R(D)$ be the Rate–Distortion Function (RDF) [19] describing the relation between strategic intent and operational effect in the context of doctrinal regulation defining the "riverbanks" between which the "stream of action", as opposed to the "stream of consciousness", is allowed to take place. The distortion measure, taken as a scalar D, represents a disjunction between the intent of the regulatory system and the actual impact on operations and tactics, seen as the transmission of a "message" along the "noisy channel" of combat or other forms of confrontation. Let R_t be the RDF at time t. Under conditions of noise and volatility, we can write the stochastic differential equation of dynamic behavior as

$$dR_t = f(t, R_t)dt + bR_t dW_t \tag{6.1}$$

where dW_t is again taken as Brownian noise.

Let $Z(R_t, t)$ be the incoming rate of essential resources needed to achieve R_t at time t. *We can then apply the Black–Scholes argument of Sect.* 14.4—expanding Z in R using the Ito Chain Rule [20]—to conclude that, for a first order approximation at nonequilibrium steady state,

$$Z_{nss} \approx \kappa_1 R + \kappa_2 \tag{6.2}$$

This is, perhaps, not a particularly unexpected result.

For a Gaussian channel under the squared distortion measure, $R = 1/2 \log[\sigma^2/D]$.

Recall that, as argued in Sect. 9.3, recognition of R as another form of free energy allows definition of an "entropy" variate, S, as

$$S = R(D) - DdR/dD = \frac{1}{2}\log[\sigma^2/D] \tag{6.3}$$

This definition, in turn, leads to a nonequilibrím Onsager equation in the gradient of S [21]

$$dD/dt = -\mu dS/dD = \frac{\mu}{2D} \tag{6.4}$$

having the classic solution

$$D(t) \propto \sqrt{\mu t} \tag{6.5}$$

This represents an important "correspondence reduction" to ordinary diffusion that, in turn, suggests an iterative approximation to a full stochastic differential equation for the distortion D

$$dD_t = [\frac{\mu}{2D_t} - M(Z)]dt + \frac{\beta^2}{2} D_t dW_t \tag{6.6}$$

$M(Z) \geq 0$ is monotonically increasing in Z.

The nonequilibrium steady-state expectation is

$$D_{nss} = \mu / 2M(Z) \tag{6.7}$$

Application of the Ito Chain Rule to D_t^2 for Eq. (6.6) finds a necessary condition for stability as

$$M(Z) \geq \frac{\beta^2}{2}\sqrt{\mu} \tag{6.8}$$

similar to Eqs. (2.16) and (9.8).

From Eqs. (6.2) and (6.7), however,

$$M(Z) \geq \frac{\mu}{2\sigma^2} \exp[2(Z - \kappa_2)/\kappa_1] \geq \frac{\beta^2}{2}\sqrt{\mu} \tag{6.9}$$

Solving for Z gives an explicit necessary condition on Z as

$$Z \geq \frac{\kappa_1}{2} \log[\frac{\beta^2\sigma^2}{\sqrt{\mu}}] + \kappa_2 \tag{6.10}$$

for there to be a real second moment in D under the condition $Z \geq 0$.

6.3 Discussion

We have used a somewhat counterintuitive Black–Scholes argument leading to Eq. (6.2) that significantly extends the result of Eq. (2.16). We find, in effect, a "Data Rate Theorem" result for the rate at which essential resources must be supplied for "genetic" (or "epigenetic") doctrine to actually work under conflict on a Clausewitz landscape of uncertainty, attrition, and delay.

σ is defined by the embedding landscape itself, determining the "noise" against which intent must contend to attain some desired effect, measured by the distortion D. The values of κ_1, κ_2 and β, however, represent something of the mismatch between doctrine and reality, extending the results of Chap. 2: the larger these parameters, the higher the rate of resources needed to manage any particular Clausewitz landscape under a given doctrine.

Lamarckian selection insures that, for surviving institutions, doctrine will come to match the realities of the embedding Clausewitz landscape. One is, of course, reminded of the Red Army on the Eastern Front.

Some genes are more fitted to a particular environment than others. Fish do not fly well, and fishing by birds is highly specialized. While humans in small, highly trained groups are among the most deadly predators of all—current agribusiness-driven pandemics excepted [22]—as the evolutionary anthropologist Robert Boyd has put it, "culture is as much a part of human biology as the enamel on our teeth" (e.g., [23]). The observation holds at and across the tactical, operational, and strategic levels of armed conflict. Doctrine is always an expression of the larger, embedding, culture, as outlined, somewhat, in the previous chapter, and doctrinal learning outside of cultural constraints is difficult indeed. US experiences in Korea, Vietnam, Iraq, and Afghanistan provide case histories.

As one commentator put it, "The US has learned nothing from Vietnam".

References

1. Wallace, R. 2017. *Computational psychiatry: A systems biology approach to the epigenetics of mental disorder*. New York: Springer.
2. Cohen, I. 2006. Immune system computation and the immulogical hommunculus. In *MoDels 2006, LCNS*, vol. 4199, ed. O. Nierstrasz, J. Whittle, D. Harel, and G. Reggio, 499–512. Heidelberg: Springer.
3. Cohen, I., and D. Harel. 2007. Explaining a complex living system: Dynamics, multiscaling, and emergence. *Journal of the Royal Society: Interface* 4: 175–182.
4. Wallace, R., and D. Wallace. 2008. Punctuated equilibrium in statistical models of generalized coevolutionary resilience: how sudden ecosystem transitions can entrain both phenotype expression and Darwinian selection. *Transactions on Computational Systems Biology IX, LNBI* 5121: 23–85.
5. Wallace, R., and D. Wallace. 2009. Code, context, and epigenetic catalysis in gene expression. *Transactions on Computational Systems Biology XI LNBI* 5750: 283–334.
6. Wallace, R., and D. Wallace. 2016. *Gene expression and its discontents: The social production of chronic disease*, 2nd ed. New York: Springer.
7. O'Nullain, S. 2008. Code and context in gene expression, cognition, and consciousness. In *The codes of life: The rules of macroevolution*, ed. M. Barbiere, 347–356. New York: Springer. Chap. 15.
8. Dretske, F. 1994. The explanatory role of information. *Philosophical Transactions of the Royal Society A* 349: 59–70.
9. Jablonka, E., and M. Lamb. 1995. *Epigenetic inheritance and evolution: The Lamarckian dimension*. Oxford, UK: Oxford University Press.
10. Jablonka, E., and M. Lamb. 1998. Epigenetic inheritance in evolution. *Journal of Evolutionary Biology* 11: 159–183.
11. Backdahl, L., A. Bushell, and S. Beck. 2009. Inflammatory signalling as mediator of epigenetic modulation in tissue-specific chronic inflammation. *The International Journal of Biochemistry and Cell Biology* 41: 176–184.
12. Turner, B. 2000. Histone acetylation and an epigenetic code. *Bioessays* 22: 836–845.
13. Janisch, R., and A. Bird. 2003. Epigenetic regulation of gene expression: How the genome integrates intrinsic and environmental signals. *Nature Genetics Supplement* 33: 245–254.
14. Jablonka, E. 2004. Epigenetic epidemiology. *International Journal of Epidemiology* 33: 929–935.

15. Foley, D., J. Craid, R. Morley, C. Olsson, T. Dwyer, K. Smith, and R. Saffery. 2009. Prospects for epigenetic epidemiology. *American Journal of Epidemiology* 169: 389–400.
16. Scherrer, K., and J. Jost. 2007a. The gene and the genon concept: A functional and information-theoretic analysis. *Molecular Systems Biology* 3: 87–95.
17. Scherrer, K., and J. Jost. 2007b. Gene and genon concept: Coding versus regulation. *Theory in Bioscience* 126: 65–113.
18. Walace, R. 2015. *An information approach to mitochondrial dysfunction: Extending Swerdlow's hypothesis*. Singapore: World Scientific.
19. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
20. Protter, P. 2005. *Stochastic integration and differential equations*, 2nd ed. New York: Springer.
21. de Groot, S., and P. Mazur. 1984. *Non-equilibrium thermodynamics*. New York: Dover.
22. Wallace, R., L. Fernando Chaves, L. Bergmann, C. Ayres, L. Hogerwerf, R. Kock, and R.G. Wallace. 2018. *Clear-cutting disease control: Capital-led deforestation, public health austerity, and vector-borne infection*. New York: Springer.
23. Richerson, P., and R. Boyd. 2006. *Not by genes alone: How culture transformed human evolution*. Chicago, IL: Chicago University Press.

Chapter 7
Challenges to the US Security Doctrine of "Resilience"

7.1 Introduction

Institutions, commercial enterprises, communities, political entities, and the social milieus in which they are enmeshed, must engage in cognitive processes of control and command to address day-to-day contingencies and incorporate the learning needed to successfully adapt to larger evolutionary selection pressures. Applying an evolutionary perspective based on the asymptotic limit theorems of information theory—extending, somewhat, the arguments of Sect. 1.4—we will show a failure to adapt can be driven by environmental challenges or internal dynamics. These represent de-facto selection pressures that can force such systems into persistent pathological "ground states".

The subsequent and consequent developmental path then remains pathological in the absence of sufficiently sustained external correction pressures. These powerful dynamics preclude the engineering resilience approach of current US security doctrine: while it may be possible to ensure the return to normal function for relatively simple power and communications networks under moderate perturbation, extension of the idea to socioeconomic entities is a ludicrous fantasy. Ecosystem and evolutionary perspectives that recognize the possibility of path-dependence and long-term eutrophication, in various forms, are more relevant, and may lead to realistic and sustainable policy objectives.

7.2 The Argument

What, more precisely, do US national security planners mean byresilience? The US Dept. of Homeland Security [1] characterizes resilience in these terms:

> The term "resilience" [as used in current US security doctrine] refers to the ability to adapt to changing conditions and withstand and rapidly recover from disruption due to

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_7

> emergencies…The United States officially recognized resilience in national doctrine in the 2010 National Security Strategy, which states that we must enhance our resilience - the ability to adapt to changing conditions and prepare for, withstand, and rapidly recover from disruption…

One is powerfully reminded of the French military doctrine of "Elan Vital" on the eve of WW I, which Tuchman [2] describes as follows:

> Translated into military terms Bergson's "elan vital" [the all conquering will] became the doctrine of the offensive. In proportion as a defensive strategy gave way to an offensive strategy, the attention paid to the Belgian frontier gradually gave way in favor of a progressive shift of gravity eastward toward the point where a French offensive could be launched to break through to the Rhine.

The results of this shift included catastrophic levels of casualties.

Western cultural atomism, it has been argued ([3] Chap. 1, and references therein), limits, and indeed badly deforms, theory in economics, evolution, and human psychology. An example of ideological deformation in"resilience" research is the work of Gao et al. [4] published by the prestigious international journal Nature—and supported by the Army Research Laboratories and the Defense Threat Reduction Agency—that seeks "universal resilience patterns in complex networks". Gao et al. claim

> [Our] analytical results unveil the network characteristics that can enhance or diminish resilience, offering ways to prevent the collapse of ecological, biological or economic systems, and guiding the design of technological systems resilient to both internal failures and environmental changes.

This is a predetermined result clearly driven by the National Security Strategy doctrine quoted above. Military-funded research is centrally tasked with implementing doctrine, here constrained to deliver a simplistic engineering resilience involving the ability to bounce back to near normal function after significant perturbation. Current US security doctrine thus inhibits exploration more complex—and far more likely—scenarios that do not have politically palatable outcomes.

Failure to adapt national security doctrine to the underlying and evolving realities of tactical, operational, and strategic circumstances is a central, ongoing, and disastrous, failing of the current US security enterprise [5, 6]. In consequence, the bland parroting of an inherently flawed statement-of-doctrine in a major international journal is potentially an ethical lapse of some consequence, even though the associated research funding is highly desirable.

This is no small matter, and one is truly reminded of France's "Elan vital", a doctrine that involved, in addition to command blindness regarding German incursion through Belgium, such tactics as mass charges into concentrations of machine guns entrenched behind barbed wire. Today, US security doctrine calls for "Resilience". Apparently, a 100 years has not been enough time for Western military practitioners to appreciate the deadly burdens of doctrinal fantasy. Of course, the American war in Vietnam and the recent occupation of Iraq also come to mind. It can be argued that, in spite of debilitatingly massive and continuing expenditures, the US defense establishment has had few significant strategic successes since the War in the Pacific (essentially, the USSR won the war in Europe). The Inchon landing of 1950, and the initial stages of the two Gulf Wars, were indeed tactical masterstrokes, but these were wasted by subsequent gross—continuing, and indeed accelerating—incompetence at the highest levels (e.g., [6]).

Apparently, this is all of a piece.

A growing economics literature raises significant questions regarding the utility of “resilience” as an organizing theme. Bene et al. [7], for example, write

> Although it is appealing, one should not rely on the term too heavily. It is not a panacea and certainly not the new catch all for development. Instead, it needs to be considered more carefully, especially with the recognition of “good” and “bad” resilience.
>
> On the basis of this, practitioners need to step back, consider the objectives of their interventions and then consider how resilience may support or actually hinder these objectives.
>
> In particular, a resilience-based systems approach might end up leading us toward abandoning interest in the poor(est) for the sake of system-level resilience.
>
> The politics of resilience (who are the winners who are the losers of resilience interventions) need to be recognized and integrated more clearly into the current discussion.

These are deep waters, and analyses of even the simplest real systems via “engineering resilience”—using standard Lyapunov methods—must be very carefully done (e.g., [8]).

Andergassen et al. [9] describe network dynamics as an evolutionary process via firms’ interaction, leading to a lock-in phenomenon that “need not be positive for all types of routine. Those [in the network] that are mostly or entirely local perform poorly and increase the system’s vulnerability…”

“Resilience”, unlike the *Ex Cathedra* pronouncement of US security doctrine, is indeed far from a policy panacea.

Rocha et al. [10] examine a broad spectrum of cascading ecosystem regime shifts within and across scales. They found that some 45% of regime shift pairwise combinations studied present at least one plausible structural interdependence, concluding that “Management of regime shifts should account for potential connections”. We shall provide an economic example below.

We know that enterprises and institutions—cognitive entities that can learn from experience and incorporate that learning into corporate culture—are subject to evolutionary selection pressures strongly enforcing path dependence ([3, 11, 12], and references therein). “Buggy whip” industries become extinct in the face of significant market shifts if they do not adapt rapidly enough.

A fairly direct, if not entirely elementary, study of the dynamics of institutional networks under selection pressure and uncertainty profoundly challenges the engineering resilience requirements of US defense and homeland security doctrine for enterprises, institutions, and socioeconomic systems. Intractable difficulties emerge from the powerful cognitive processes that must incorporate learning into corporate culture for successful adaptation, while crafting responses under shifting day-to-day market demands. Most particularly, facing environmental pressures, in a large sense, cognitive entities can be driven into a pathologically stable ground state of paralysis and fixation that may ensure their ultimate demise. Here, we will build on the model of evolutionary dynamics used to study strategy that was presented earlier.

7.3 Punctuated Equilibrium Redux

The essential ideas, then, replicate, in some measure, the evolutionary treatment of strategy in Sect. 1.4. Networks of interacting institutions and other cognitive enterprises individually persist through some form of replication and reproduction, but are subject to selection pressures. Such networks will, it has been observed, often undergo long periods of apparent stasis while their phenotypes (in a large sense) can adapt, before experiencing relatively rapid wholesale shifts in structure and behavior. Once begun, such shifts are path dependent and permanent: as it is said, one never crosses the same river twice. A pristine lake, once eutrophied, remains contaminated [13]. A species, once extinct, is no more.

For natural populations, such shifts are usually extinction and/or speciation events. For networked systems of institutions and other cognitive phenomena, the dichotomy is usually extinction/fundamental structural change. Rogers [14] was perhaps the first to apply the concept in a military setting, using it to describe the processes of innovation associated with the use of cannon in the first Hundred Years War. Others have followed his lead, applying the ideas to contemporary circumstances (e.g., [15]). The model is that of Fig. 7.1, taken from the notebooks of Charles Darwin.

Extending, somewhat, the arguments of Sect. 1.4, and for pedagogic reasons, we reconstruct the formal scaffolding.

Cognition—individual or institutional—involves an agent’s choice of an action, or a relatively small number of them, from the complete repertoire of those available [16]. This permits the identification of a number of “dual” information sources representing a set of interacting agents, say $Y_1, Y_2, \ldots, Y_m$, where m may be large. That is, cognitive choice implies a reduction of uncertainty which, in turn, implies the existence of an underlying information source characterizing that cognition.

Sequences of actions and behaviors of some length $N \to \infty$ can, in general, be divided into a small high probability “typical” set consonant with a “grammar” and “syntax”, and a very large set of paths having vanishingly low probability that is not consistent with the underlying grammar and syntax [17].

However, coevolutionary dynamics may require invocation of “nonergodic” information sources, following the arguments presented in [18]. “Nonergodic” means that long-term averages do not converge on cross-sectional expectations, roughly analogous to the idea of “nonparametric” statistical models that are not based on the Central Limit Theorem. The source uncertainty of such systems then cannot be described in terms of the classic Shannon “entropy” [17], but regularities arising from Feynman’s [19] identification of information as a form of free energy still permit considerable simplification of system dynamics via familiar “phase transition” formalism, based, however, on groupoid rather than group symmetry shifts. More of this below.

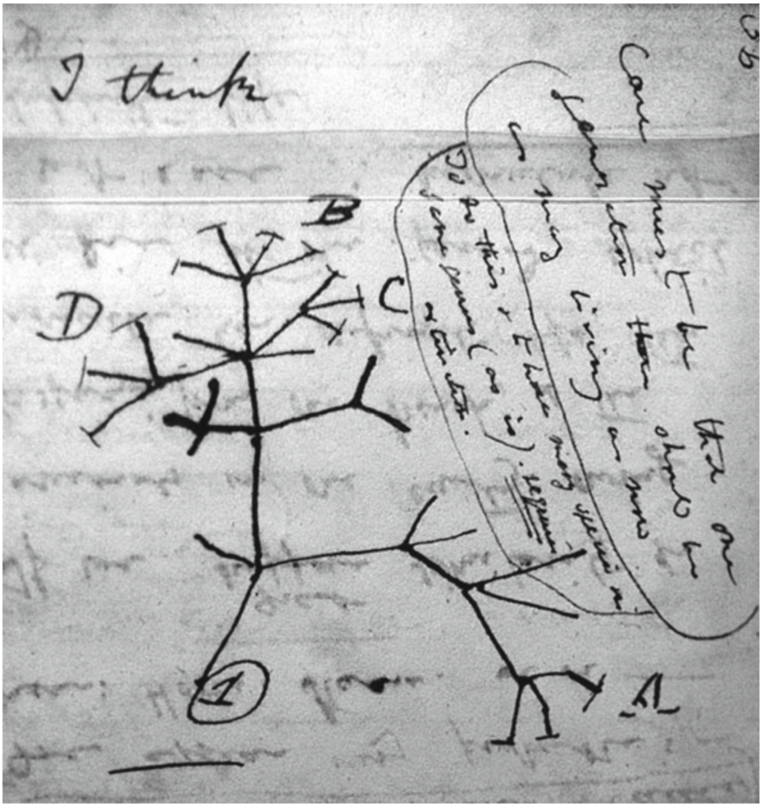


Fig. 7.1 From the notebooks of Charles Darwin, the first diagram of evolutionary process. Complicated social or biological systems, that incorporate the impacts of selection pressures through adaptation or learning, change form and move on, seldom returning to their earlier states. *Photo* Rodrick Wallace

Following [20], the embedding natural or human ecosystem upon which the business of human enterprise takes place, the underlying “landscape” in some sense, can be characterized by its own “grammar” and “syntax” defining what are high and low probability sequences of events, permitting representation as an information source, say X. That is, such matters as seasonal transitions—from high summer to a rainy autumn choked in mud, to a hard-frozen surface that can support tracked vehicles, or from low to high demand seasons in the demand for goods, services, or electricity—are not random, but have persistent regularities.

Another essential information source arises from the dynamics of the “large deviations” to which all agencies or agents in dynamical systems are subject. As Champagnat et al. [21] describe, sudden transitions between nonequilibrium steady states also involve high probability paths whose description is in terms of the familiar Shannon uncertainty expression

$$\mathscr{I} = -\sum_k P_k \log(P_k)$$

where the P_k form a probability distribution. Variations of this result appears repeatedly as Sanov's Theorem, Cramer's Theorem, the Gartner–Ellis Theorem, the Shannon–McMillan Theorem, and so on [22]. This result allows the characterization of a system's large deviations in terms of yet another information source, L_D.

Study now focuses on a very complicated—and not necessarily ergodic—joint information source significantly different from that characterizing Zweikampf as studied in Sect. 1.4:

$$H(X, Y_1, Y_2, \ldots Y_m, L_D) \tag{7.1}$$

This joint uncertainty is defined on a set of jointly typical [23] high probability paths z defined by a sequence of high-dimensional states that conform to the "grammars" and "syntaxes" of the interacting information sources $X, Y_1, Y_2, \ldots Y_m$ and L_D. Again, there can be very many information sources Y_k linked in extremely topologically rich networks of crosstalk (and other forms of exchange that can be reinterpreted as crosstalk).

The argument then becomes concerned with the set of "jointly typical" high probability paths $z^n \equiv \{z_0, z_1, \ldots, z_n\} \rightarrow z$.

Following [17], it is possible to define a path-dependent source (joint) uncertainty $H(z^n) \rightarrow H(z)$ that can vary across the manifold defined by the full set of paths z. $H(z)$, however, no longer has an "entropy like" mathematical expression, but its dynamics can be inferred using a familiar statistical argument, invoking a Boltzmann pseudoprobability as

$$P[H(z_q)] \equiv \frac{\exp[-H(z_q)/\tau]}{\sum_z \exp[-H(z)/\tau]} \tag{7.2}$$

where the sum, or a generalized integral, of the denominator is over all possible jointly typical paths z, and τ is an appropriate "temperature" analog that we will develop further below.

From that pseudoprobability, a "free energy" $\mathscr{F}$ can be defined as a Morse function [24], using the denominator of Eq. (7.2) and the relation

$$\exp[-\mathscr{F}/\tau] \equiv \sum_z \exp[-H(z)/\tau] \tag{7.3}$$

The definition of τ here differs from that in Sect. 1.4, since the rates of supply of essential resources for large, complicated socioeconomic networks may greatly differ from circumstances of conflict. We might construct τ economically as a scalar index from a broad spectrum of expenditures per unit time, (inverse) loss rates, rates of other opportunity costs, measures of degradation rates of the political or economic capital needed for institutional function, and so on. Typically, one might conduct a principal component analysis across some elaborate correlation matrix, and choose the magnitude of the largest vector, representing a volume measure, in a large sense.

Given some such scalar rate measure τ, a subtle phase transition model again emerges by identifying equivalence classes of a system's developmental pathways z, again, say, "functional" and "catastrophic". Identification of equivalence classes allows definition of a symmetry groupoid for the developmental process [20, 25, 26]. A groupoid, again as outlined in the Mathematical Appendix, is a generalization of an algebraic symmetry group in which a product is not necessarily defined between each element. The simplest example is, perhaps, a disjoint union of separate groups, but sets of equivalence classes also define a groupoid. See [25] for details.

Expanding the argument, given an information source associated with the networked system of interest, a full equivalence class algebra can be constructed by choosing different system origin states and defining the equivalence of subsequent states at a later time by the existence of a high probability path connecting them to the same origin state. Disjoint partition by equivalence class, analogous to orbit equivalence classes in dynamical systems, defines a symmetry groupoid associated with the cognitive process [25].

The equivalence classes across possible origin states define a set of information sources dual to different states available to the systems of interest. These create a large groupoid, with each orbit corresponding to an elementary "transitive" groupoid whose disjoint union is the full groupoid. Each subgroupoid is associated with its own dual information source, and larger groupoids must have richer dual information sources than smaller.

The "free energy" Morse Function of Eq. (7.3) is then, as argued above, liable to an analog of Landau's classical spontaneous symmetry breaking [24]. Recall that, under symmetry breaking, higher "temperatures" are associated with more symmetric higher energy states in physical systems. Cosmology makes much of such arguments regarding the first moments after the "big bang". The assertion is that different physical phenomena broke out as the universe cooled through expansion.

Under the impact of external (or internal) "selection pressure", decline in the scalar rate index τ from Eq. (7.2) will cause sharply punctuated collapse from higher to lower symmetry states, triggering ultimate system failure, or at least transitions to significantly less functional, or even deteriorated, states.

Again, the "free energy" Morse Function of Eq. (7.3) can be driven into sudden, highly punctuated "phase transition" analogs reflecting fundamental symmetry shifts by changes in the "temperature" parameter τ. The symmetry shift, however, is between groupoids associated with a synergism across networked institutional cognitions, environmental dynamics, and the structure of stochastic "large deviations". This is not simply the melting of ice or the boiling of water.

The essential point—as in Sect. 1.4—is that phase transitions in networked systems associated with $\mathscr{F}$, τ, their generalizations or their analogs, may be relatively subtle, akin to the Eldredge and Gould [27] pattern of "punctuated equilibrium" in evolutionary transition. Under this model, evolutionary and coevolutionary dynamics undergo relatively long periods of apparent or near-stasis, where changes are small or difficult to see, followed by relatively sudden massive changes leading to fundamentally different coevolutionary configurations. Examples of contemporary debate on punctuated equilibrium, now a central matter in evolutionary biology, can be found in [28–31]. An extended "socioeconomic" case history is given in the next chapter.

The topologically rich network of interaction between the set of information sources $\mathscr{Y} \equiv \{Y_k, k = 1, \ldots, m\}$ representing individual (cognitive) enterprises takes matters beyond into realms of exceedingly complex dynamics.

Nonetheless, we suggest the dynamics of the free energy analog $\mathscr{F}$ at (the perhaps many) critical values of τ can be studied using variants of the standard iterative renormalization techniques as developed for cognitive processes in [32].

The most direct general inference, however, remains that phase transitions associated with changes in τ will be expressed in terms of changes in underlying groupoid symmetries, leading again to some version of the punctuated equilibrium dynamics identified by Eldredge and Gould.

If the set $\mathscr{Y}$ is sufficiently large, we can approximate the sum in the denominator of Eq. (7.2) as an integral across an appropriate distribution, leading to an analysis like that of Fig. 1.5. This is done in the Chapter Appendix for a system with a Gamma distribution in $H(z)$.

The devil, yet again, is in the details, and one of those is again stochastic variation. Schweitzer et al. [33] comment that, for complex socioeconomic networks in particular,

> The problem of network formation changes substantially if the underlying environment is subject to persistent volatility, such as rapid innovation, sociopolitical instability, or environmental change…In such cases, agents cannot hope to attain optimal configurations and, moreover, one finds that the performance of the system can be sharply sensitive to small changes in environmental volatility…[B]ig disruptions on the system's level do not need large perturbations to occur…The ubiquitous tendency to extrapolate new behavior from [the] past…is fundamentally mistaken at such phase transitions, since the new collective organization is in general completely different from the previous one.

Again, see Fig. 7.1.

A more complete description of system dynamics, beyond punctuated phase transitions driven by changes in τ, can be explored, as suggested by Schweitzer et al., and as presented in Sect. 1.4, using stochastic Onsager relations constructed in a standard manner from gradients in an "entropy" defined in terms of a vector of system parameters K as

$$S \equiv \mathscr{F}(K) - K \cdot \nabla_K \mathscr{F} \tag{7.4}$$

The Onsager approximation to system dynamics is then via a linear expansion for the rate of change of the parameter vector K in the gradients of S by the components of K [18, 34].

The general stochastic model is then of the form

$$dK_t = f(K, t)dt + g(K, t)dW_t \tag{7.5}$$

dW_t is a multidimensional Brownian noise, and f is assumed to be locally Lipschitz. It may be inherently stable or unstable, depending on local (or global) conditions. Appleby et al. [35] show that a function g can always be found that stabilizes an inherently unstable function f, i.e., one that diverges to ∞. Conversely, in two or more dimensions, a function g can always be found that destabilizes an inherently stable equilibrium for f. As noted, these dynamics will likely have significant policy implications [36].

7.4 Discussion and Conclusions

Institutions, commercial enterprises, their topologically rich networks and communities, and their larger, embedding social structures, must engage in cognitive process to address rapidly changing patterns of challenge and opportunity, and, more slowly, incorporate the learning necessary for successful adaptation to shifts in larger scale evolutionary selection pressures. Failure of cognition, or of learning/adaptation, can be triggered by environmental challenges that drive such structures into highly persistent ground states where cognitive or learning/adaptational process becomes pathologically fixated, initiating a developmental pathway to failure.

The accelerating political and public health instabilities following the Cold War-induced deindustrialization of the US provide a powerful case history [37, 38]. Details of how US deindustrialization was driven by fifty years of Cold War can be found in Ullmann [39], Melman [40], and related works. Figure 7.2 shows the ratio of industrial to total nonform employment in the US from January 1, 1939 through July 1, 2015 [41]. Two "ecosystem regimes" are evident. The first, from 1946 through 1965, displays a relatively slow rate of decline. Subsequently, the rate of decline underwent a punctuated increase.

Figure 7.3 shows the decade-delayed impact of US deindustrialization on the country's trade deficit, an unsustainable bleeding-out of the nation's economy [42].

Massive deindustrialization and associated economic de-development led, in turn, to punctuated political destabilization as expressed in yet another fundamental regime shift: the 2016 US Presidential election [43].

Figure 7.4 shows counties in the US "Rust Belt" that lost more than 1000 manufacturing jobs between 1972 and 1987. These counties were, in large measure, responsible for the outcome of the 2016 US Presidential election.

These three large-scale events clearly represent a cascade of regime shifts in the sense of Rocha et al. [10].

Wallace and Wallace [44, 45], outlined in Chap. 8 of this work, explore the long-term impacts of a "planned shrinkage" policy aimed against minority voting blocs in New York City and implemented by the targeted withdrawal of fire extinguishment resources from poverty-stricken, high fire incidence, high population density neighborhoods. This deliberate policy triggered rapid downward spirals of social and physical disintegration causing great exacerbation of multiple indicies of morbidity, mortality, and criminal activity.

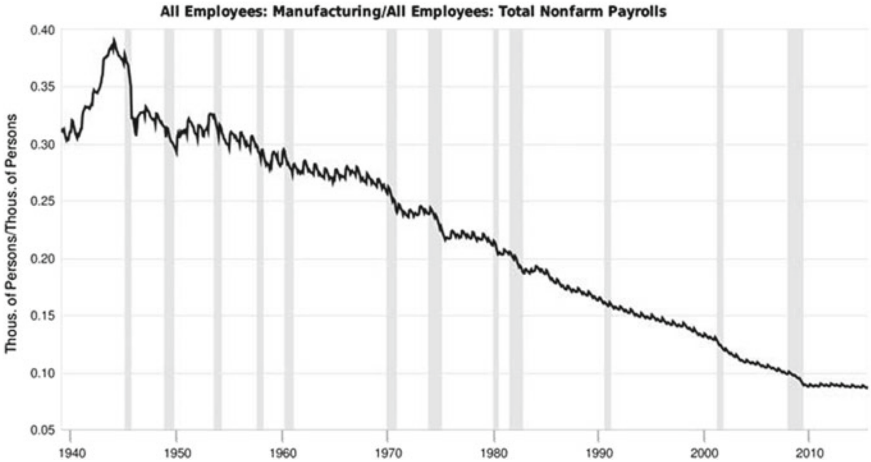


Fig. 7.2 From [41]. For the United States, 1/1/1939-7/7/2015, the ratio of industrial to total nonform employment. The decline between 1946 and 1965 was at a modest rate that underwent a punctuated "eutrophic" increase as a result of the systematic and policy-driven diversion of essential scientific and engineering resources into the Cold War. The punctuated rate of decline led, in turn, to punctuated political destabilization at the time of the 2016 Presidential election [43]. This sequence of large-scale events represents a cascade of regime shifts as studied by Rocha et al. [10]

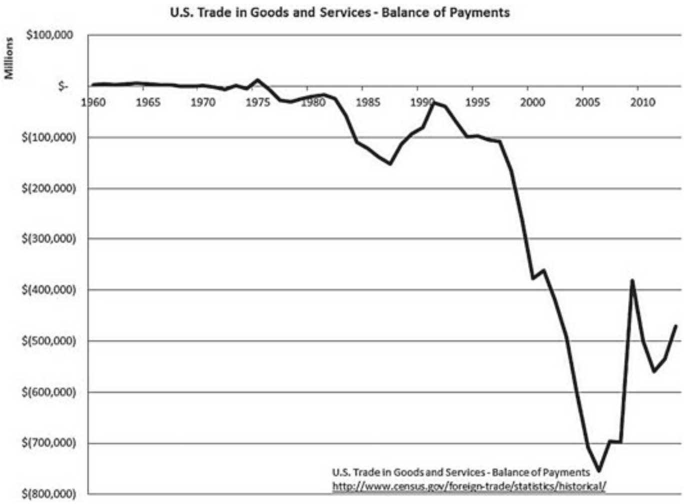


Fig. 7.3 From [42], US trade deficit, 1960–2013. The regime change after 1975 reflects the delayed impact of the 1965 regime change in US industrial structure, again a cascade of regime shifts as studied by Rocha et al. [10]

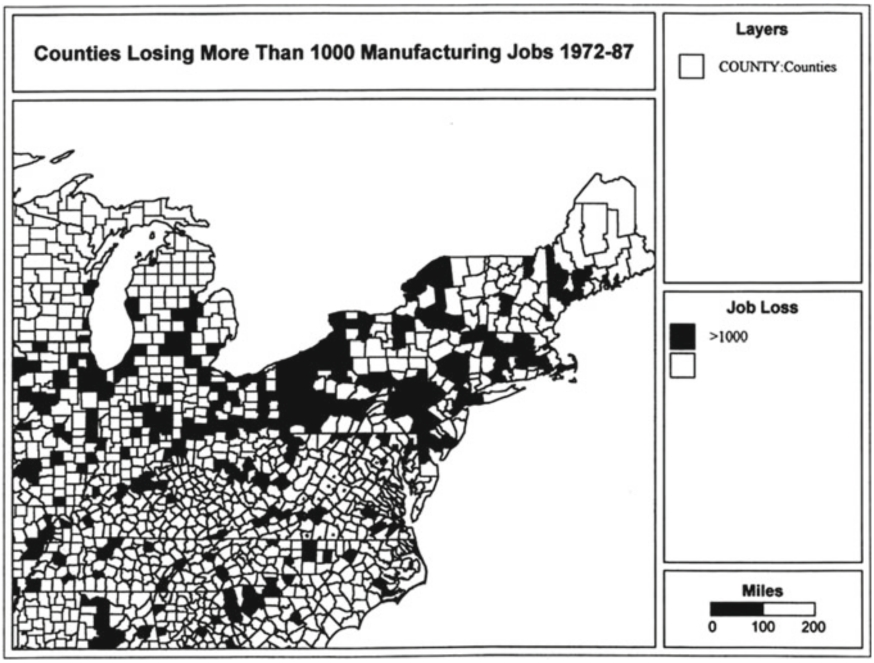


Fig. 7.4 Counties in the Northeastern "Rust Belt" of the US that lost more than 1000 manufacturing jobs between 1972 and 1987. These counties, in a cascade of regime shifts [10], delivered votes essential to the outcome of the 2016 Presidential election

Readers will have their own examples.

None of this resembles the engineering resilience of current US security doctrine as—supposedly—instantiated by the work of Gao et al. [4]. While it may be possible to stabilize relatively simple power and communications systems under moderate perturbation, extension of the approach to topologically complex networks of institutions, enterprises, economic structures, and communities is ludicrous. Ecosystem and institutional evolutionary perspectives that recognize the possibility of persistent eutrophication, in various forms, are more to the point (e.g., [13]), although the implications will most surely not please the present US security establishment, whose many wishful-thinking "Elan"-like doctrines, and the policies based on them, have contributed materially to the nation's decline ([3, 37], and references therein).

Ecosystem resilience and institutional evolutionary theories may, if properly adapted, provide tools actually useful in achieving realistic policy goals.

7.5 Chapter Appendix: A Simplified Model of Phase Transition

If the denominator in Eq. (7.2) can be treated with an integral approximation, then something like the analysis leading to Fig. 1.5 can be carried out.

We assume that the set $\mathcal{Y}$ is very large and dense, and that H is characterized by a Gamma distribution with parameters k, θ having a density function

$$P(H, k, \theta) = \frac{H^{k-1} \exp[-H/\theta]}{\theta^k \Gamma(k)} \tag{7.6}$$

with $\theta, k > 0$ real.

Then $< H >= k\theta$ and

$$\int_0^\infty \exp[-H/g(\tau)] P(H, k, \theta) dH = \exp[-F/g(\tau)] = \left(\frac{g(\tau)k}{g(\tau)k+ < H >} \right)^k \tag{7.7}$$

An entropy-analog is again defined as

$$S(\tau) \equiv F(\tau) - \tau dF(\tau)/d\tau \tag{7.8}$$

leading again to the Onsager nonequilibrium steady-state condition

$$dS/d\tau = 0 \tag{7.9}$$

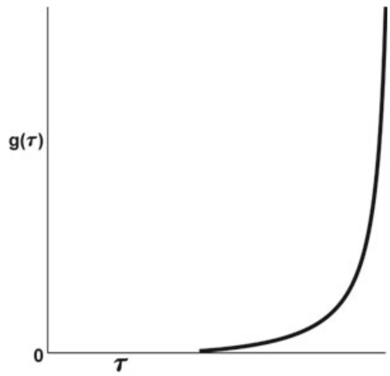


Fig. 7.5 Simplified phase transition treatment of Eq. (7.3), using the methods leading to Fig. 1.5, assuming a Gamma probability distribution for values of H over a dense set of cognitive entities $\mathcal{Y}$. $g(\tau)$ is real and finite only over a limited range, implying the existence of at least three different phases, for this simplistic model. More realistic models will likely have more complicated phase transition dynamics

The resulting differential equation in $g(\tau)$ can be solved to give the relation

$$
\begin{aligned}
z &\equiv -C_2\tau + C_1 \\
g(z) &= \frac{z}{W(-1, (kz/ < H >) \exp[kz/ < H >]) - kz/ < H >}
\end{aligned}
\tag{7.10}
$$

C_1 and C_2 are positive constants and $W(-1, x)$ is the Lambert W-function of order -1, real and finite only in the range

$$
-\exp[-1] < (kz/ < H >) \exp[kz/ < H >] < 0 \tag{7.11}
$$

Thus, $g(\tau)$ can be real and finite only over a limited range of τ-values, as shown in Fig. 7.5, implying at least three different phases in the underlying system.

References

1. US department of homeland security. 2016. https://www.dhs.gov/topic/resilience.
2. Tuchman, B. 1962. *The guns of august*. New York: Dell.
3. Wallace, R. 2015. *An ecosystem approach to economic stabilization: Escaping the neoliberal wilderness*. New York: Routledge Advances in Heterodox Economics.
4. Gao, H., B. Barzel, and A. Barabasi. 2016. Universal resilience patterns in complex networks. *Nature* 530: 307–312. https://doi.org/10.1038/nature18019. See ERRATUM.
5. Watts, B. 2008. US Combat training, operational art, and strategic competence: Problems and opportunities, Center for Strategic and Budgetary Assessments, Washington D.C.
6. Krepinevich, A., and B. Watts. 2009. *Retaining strategic competence*. Washington, D.C.: Center for Strategic and Budgetary Assessment.
7. Bene, C., R. Godfrey Wood, A. Newsham, and M. Davies. 2012. Resilience: New utopia or new tyranny? IDS working paper, 405.
8. Reggiani, A., T. De Graaff, and P. Nijkamp. 2002. Resilience: An evolutionary approach to spatial economic systems. *Networks and Spatial Economics* 2: 211–229.
9. Andergassen, R., F. Nardini, and M. Ricottilli. 2015. Emergence and resilience in a model of innovation and network formation. *Networks and Spatial Economics* 15: 293–311.
10. Rocha, J., G. Peterson, O. Bodin, and S. Levin. 2019. Cascading regime shifts within and across scales. *Science* 362: 1379–1383.
11. Hodgson, G., and T. Knudsen. 2010. *Darwin's conjecture: The search for general principles of social and economic evolution*. Chicago, IL: University of Chicago Press.
12. Wallace, R. 2013. A new formal approach to evolutionary processes in socioeconomic systems. *Journal of Evolutionary Economics* 23: 1–15.
13. Holling, C. 1973. Resilience and stability of ecological systems. *Annual Reviews of Ecological Systematics* 4: 1–23.
14. Rogers, C. 1993. The military revolutions of the hundred years war. *The Journal of Military History* 57: 241–278.
15. Leonard, S. 2001. Inevitable evolutions: Punctuated equilibrium and the revolution in military affairs. A monograph of the school of advanced military studies, United States Army Command and General Staff College, Ft. Leavenworth, KA (online).
16. Atlan, H., and I. Cohen. 1998. Immune information, self-organization, and meaning. *International Immunology* 10: 711–717.
17. Khinchin, A. 1957. *The mathematical foundations of information theory*. New York: Dover Publications.

18. Wallace, R. 2018. New statistical models of nonergodic cognitive systems and their pathologies. *Journal of Theoretical Biology* 436: 72–78.
19. Feynman, R. 2000. *Lectures on computation*. New York: Westview Press.
20. Wallace, R. 2017. *Computational psychiatry: A systems biology approach to the epigenetics of mental diaorders*. Berlin: Springer.
21. Champagnat, N., R. Ferriere, and S. Meleard. 2006. Unifying evolutionary dynamics: From individual stochastic processes to macroscopic models. *Theoretical Population Biology* 69: 297–321.
22. Dembo, A., and O. Zeitouni. 1998. *Large deviations and applications*, 2nd ed. Berlin: Springer.
23. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
24. Pettini, M. 2007. *Geometry and topology in hamiltonian dynamics*. Berlin: Springer.
25. Weinstein, A. 1996. Groupoids: Unifying internal and external symmetry. *Notices of the American Mathematical Society* 43: 744–752.
26. Golubitsky, M., and I. Stewart. 2006. Nonlinear dynamics and networks: The groupoid formalism. *Bulletin of the American Mathematical Society* 43: 305–364.
27. Eldredge, N., and S. Gould. 1972. Punctuated equilibrium: An alternative to phyletic gradualism. In *Models in paleobiology*, ed. T. Schopf, 82–115. San Francisco: Freeman, Cooper and Co.
28. Hunt, G., M. Hopins, and S. Lidgard. 2015. Simple versus complex models of trait evolution and stasis as a response to environmental changes. *PNAS* 112: 4885–4890.
29. Landis, M., and J. Schraiber. 2017. Pulsed evolution shaped modern vertebrate body sizes. *PNAS* 114: 13224–13229.
30. Valverde, S., and R. Sole. 2015. Punctuated equilibrium in the largescale evolution of programming languages. *Journal of the Royal Society Interface* 12: 20150249 (online).
31. Lyne, J., and H. Howe. 2017. ‘Punctuated equilibrium’: Rhetorical dynamics of a scientific controversy, chapter 6. In *Landmark essays on rhetoric of science: case studies*, Second Edition, ed. Harris, R., CRC Press, Boca Raton.
32. Wallace, R. 2005. *Consciousness: A mathematical treatment of the neuronal global workspace model*. Berlin: Springer.
33. Schweitzer, F., G. Fagiolo, D. Sornette, F. Vega-Redondo, and D. White. 2009. Economic networks: What we know and what we need to know. *Advances in Complex Systems* 12: 407–422.
34. de Groot, S., and P. Mazur. 1984. *Nonequilibrium thermodynamics*. New York: Dover.
35. Appleby, J., X. Mao, and A. Rodkina. 2008. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control* 53: 683–691.
36. Wallace, R., L.F. Chaves, L. Bergmann, C. Ayres, L. Hogerwerf, R. Kock, and R.G. Wallace. 2018. *Clear-cutting disease control: Capital-led deforestation, public health austerity, and vector-borne infection*. Berlin: Springer.
37. Wallace, R., and D. Wallace. 2016. *Gene expression and its discontents: The social production of chronic disease*, 2nd ed. Berlin: Springer.
38. Wallace, R. 2015. *An ecosystem approach to economic stabilization: Escaping the neoliberal wilderness*. New York: Routledge.
39. Ullmann, J.E. 1988. *The anatomy of industrial decline*. Westport, CT: Greenwood-Quorum.
40. Melman, S. 1971. *The war economy of the United States: Readings in military industry and economy*. New York: St. Martin’s Press.
41. Federal Reserve Bank of St. Louis data system myf.red/g/mVt1.
42. Johnson, D. 2014. You should be afraid of the trade deficit, not the budget deficit, March 6, OurFuture.org (online) https://ourfuture.org/20140306/you-should-be-afraid-of-the-trade-deficit-not-the-budget-deficit.
43. Wallace, D., and R. Wallace. 2019. *Hierarchy and public health: Voting patterns in the 2016 US presidential election*. New York: Routledge.
44. Wallace, D., and R. Wallace. 1998. *A plague on your houses*. New York: Verso.
45. Wallace, D., and R. Wallace. 2011. Consequences of massive housing destruction: The New York city fire epidemic. *Building Research and Information* 39: 395–411.

Chapter 8
Culture and the Induction of Emotional Dysfunction on a Clausewitz Landscape

8.1 Introduction

Recall Minsky's famous manifesto [1]:

> The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions.

We argue here that no cognitive agent can act in real time intelligently without emotion analogs that permit rapid, patterned response to critical "sensory" input or the impact of sudden internal signals. These agents will range in scale and form across individual humans, small disciplined groups, larger institutions, man/machine cockpit enterprises, and AI systems.

But individual human emotions, as well as institutional and machine emotion-analogs, are always embedded in culture and hence remain—in surprising measure—cultural artifacts. Such constraint on "free will" has implications for the quality of function on Clausewitz landscapes of imprecision, uncertainty, and time constraint. It has been said that "free will" is like living inside a Swiss cheese. One simply does not perceive the holes in a generalized exercise of perceptual completion.

The implication is that sufficient understanding of the cultural blind spots of an adversary—particularly in the context of an opponent's perception of one's own—can serve as a tool for disruption of adversarial decision and deliberation under time constraint. This perspective supplements, in a sense, the considerations of the first essay in this series, moving beyond mechanistic Data Rate Theorem control theory strategy and tactics into deeper matters of semantic manipulation.

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_8

8.2 An Evolutionary Perspective on Emotion

Emotions in animals are widely understood to represent learned or inherited patterns of very rapid response to particular sets of environmental or sensory cues [2]. The evolutionary advantage of rapid patterned responses is obvious, and applies even to what might be considered "nonminded" organisms.

For humans, or at least those tested in an American university setting [3], Cowen and Keltner [4] identify as many as 27 distinct categories of emotion, bridged by continuous gradients. Even the simplest categorizations identify at least two affective dimensions, "valence" and "arousal" [4]. This dimensional multiplicity will prove to be a critical observation.

For humans, matters of emotion have long been known to be deeply intertwined with culture and both individual and community historical trajectory (e.g., [5, 6] Chap. 3, and references therein).

Human institutions, like bee colonies, are cognitive and can learn, but are likewise embedded in a cultural milieu, and their patterns of rapid, intuitive, response to internal ruminations or external information also determine their ability to respond to Lamarckian selection pressures and persist.

Cognitive machines must also sometimes act on incomplete information, via "anytime algorithms" (AA) that provide "good enough" answers under time constraints (e.g., [7]). What is perhaps less obvious is that machines, or synergistic "cockpit" enterprises across man, institution and machine, also operate within a cultural milieu, and are not independent of the sociocultural "riverbanks" confining a de-facto "stream of consciousness".

Machine intelligence is just another cultural artifact, and the design and use of such things—much as woven baskets—is always culturally embedded.

Here, we take an inverse perspective on emotion, using an approach from machine cognition—the Anytime Algorithm (AA)—to examine necessary constraints on rapid patterned response mechanisms in institutions, individual humans, and man/institution/machine composites.

8.3 Generalizing Anytime Algorithms

Recall the argument from Chap. 2, in particular how Zilberstein [7] describes the basic idea of the AA:

> …[Anytime algorithms, AA] give intelligent systems the capability to trade deliberation time for quality of results. This capability is essential for successful operation in domains such as signal interpretation, real-time diagnosis and repair, and mobile robot control. What characterizes these domains is that it is not feasible (computationally) or desirable (economically) to compute the optimal answer…

Recall the the "quality" graphs of a "traveling salesman" problem calculation as it approaches an asymptotic limit with time, Fig. 2.1, and the general calculation of Sect. 2.3. The essential point is that even the most direct AA are not necessarily stable, and that sufficient conditions for stability must be found on a case-by-case basis, often using obscure stochastic Lyapunov function methods that are similar to arguments surrounding the stability of stochastic differential equations (e.g., [8]).

Recall Eqs. (2.4) through (2.8). Taking an exponential model for the approach of calculation quality to nonequilibrium steady state, both externally imposed "shadow price" measures and the magnitude of "noise" determined the stability of the system. Here, we will extend the AA model to include the effects of an embedding cultural milieu on cognition under conditions of imprecision, uncertainty, and time constraint.

Recall that, for a model of a one-dimensional AA, sufficient noise, or the imposition of sufficient "shadow price" by an adversary or environment, will be catastrophic.

The fact that the observed dimensionality of human emotional structure is $n \gg 2$ [4] is itself sufficient to explore something of the considerable complexity of emotional dysfunction in humans, their institutions, their cognitive machines, and the many various composite entities.

Putting aside, for the moment, the question of time constraints, we can suppose that the different cognitive submodules that make up such entities interact according to a deterministic multidimensional matrix relation that is perturbed by noise, leading to a stochastic differential equation like Eq. (2.7) having the form

$$dQ = f(Q)dt + g(Q)dW_t \tag{8.1}$$

where W_t is multidimensional Brownian noise and all quantities are of dimension $n \geq 2$, following observations of the high dimensionality of human emotional response. The multidimensional function $f(Q)$ is assumed to have a set of nonequilibrium steady states as $t \rightarrow \infty$. These may, however, be stable or unstable, depending on context.

In particular, $f(Q)$ includes patterns of crosstalk between modules, and $g(Q)$ quantifies the influence of noise within and across modules, analogous to the role of the term σQ in Eq. (2.7).

Appleby et al. [8] show that, if $n \geq 2$, under very broad conditions, for any $f(Q)$ that satisfies them, a function $g(Q)$ can be found that either stabilizes an unstable nonequilibrium steady state of f or, perhaps most importantly, destabilizes any stable state. A fully worked-out two-dimensional example can be found in Sect. 2.2 of Wallace et al. [9]. There, additive noise simply fuzzes out a two-dimensional deterministic solution, while cross-noise destabilizes a stable nonequilibrium steady state.

But it is possible to markedly extend the argument, and in particular incorporate the effects of embedding cultural milieu on cognitive dynamics.

Much depends on the rate at which complex cognitive processes must proceed under real-world, real-time selection pressures. Emotional responses are predicated on quick patterned response to sensory cues. We will propose here that there can be several levels of such patterning in complex cognitive systems, depending on the rate at which a response is required. We look at cognitive system dynamics as driven by the inverse rate τ at which an answer is required and must be acted on in the real world. High values of τ imply relatively slow response rates, which we will take as corresponding to a high "temperature" in a symmetry-breaking model of cognitive process that invokes groupoids rather than groups.

Cognition involves an agent's choice of an action, or a relatively small number of them, from the complete repertoire of those available [10]. This permits the identification of a "dual" information source to any cognitive enterprise. That is, cognitive choice implies a reduction of uncertainty that, in turn, implies the existence of an underlying information source characterizing that cognition. Wallace [6] provides further details of the argument and a worked-out patterning.

Our central interest is the dynamics of the interacting network of cognitive submodules within an agent. The agent may be an individual person, an institution, an intelligent machine, a cockpit, or more structured composite. The essential point is the need to act under time constraint defined (inversely) by τ.

The submodules will extend across both scale and level of organization.

We thus must study matters across an extensive network of entities, represented by a considerable set of cognitive information sources within a single, larger agent, say $\mathscr{Y} = \{Y_1, Y_2, \ldots\}$.

We consider, then, a dense net of cognitive submodules interacting through information crosstalk and other forms of exchange. Such agents will both cooperate and compete for essential resources, including time.

Sequences of cognitive actions and behaviors of some length $N \rightarrow \infty$ can, in general, be divided into a small high probability "typical" set consonant with a "grammar" and "syntax", and a very large set of paths having vanishingly low probability that is not consistent with the underlying grammar and syntax [11, 12].

However, cognitive dynamics may require invocation of "nonergodic" information sources, extending the arguments presented in ([13] Chap. 6, [14]). "Nonergodic" means that long-term averages do not converge on cross-sectional expectations, roughly analogous to the idea of "nonparametric" statistical models that are not based on the Central Limit Theorem. The source uncertainty of such systems then cannot be described in terms of the classic Shannon "entropy" [12], but regularities arising from Feynman's [15] identification of information as a form of free energy still permit considerable simplification of system dynamics via familiar "phase transition" formalism, based, however, on groupoid rather than group symmetry shifts.

Following [13], the embedding natural or human ecosystem upon and within which the entity of interest must act—the underlying landscape—can be characterized by its own "grammar" and "syntax" defining what are high and low probability sequences of events, permitting representation as an information source, say X.

An additional information source, Z, can be identified, within which the entity of interest is embedded and of which it is an artifact/offspring. This is the culture/historical trajectory that will have its own grammar and syntax.

Yet another information source arises from the dynamics of the "large deviations" to which cognitive agencies or agents are subject. As Champagnat et al. [16] describe, however, sudden transitions between nonequilibrium steady states also involve high probability paths whose description is in terms of the familiar Shannon uncertainty expression $\mathscr{I} = -\sum_k P_k \log(P_k)$, where the P_k form a probability distribution.

Variations of this result appears repeatedly as Sanov's Theorem, Cramer's Theorem, the Gartner–Ellis Theorem, the Shannon–McMillan Theorem, and so on [17]. This result allows the characterization of large deviations in terms of yet another information source, L_D.

Study now focuses on a very large, complex network of (not necessarily ergodic) information sources linked by crosstalk that we write here in terms of the uncertainty of a single-joint information source [11] as

$$H(X, \mathscr{Y}, Z, L_D)$$

This (highly networked) information source uncertainty is defined on a set of jointly typical [11, 12] high probability paths z defined by a sequence of high dimensional states that conform to the "grammars" and "syntaxes" of the interacting information sources X, the set $\mathscr{Y}$, the embedding culture Z, and L_D.

The argument then becomes concerned with the set of "jointly typical" high probability paths $z^n \equiv \{z_0, z_1, \ldots, z_n\} \rightarrow z$.

Following Khinchin [12], it is possible to define a path-dependent source (joint) uncertainty $H(z^n) \rightarrow H(z)$ that can vary across the manifold defined by the full set of paths z. $H(z)$, however, no longer has an "entropy like" mathematical expression, but its dynamics can be inferred using a familiar statistical argument.

The next step is to invoke a Boltzmann pseudoprobability as

$$P[H(z_q)] \equiv \frac{\exp[-H(z_q)/\tau]}{\sum_z \exp[-H(z)/\tau]} \tag{8.2}$$

where the sum, or a generalized integral, of the denominator is over all possible jointly typical paths z, and τ is, again, the inverse rate at which decision must be reached and action taken.

The $H(z_q)$ are uncertainty values associated with an enormously large set of individual paths.

From the pseudoprobability of Eq. (8.2) a "free energy" analog $\mathscr{F}$ can be defined as a Morse function [18], using the denominator of Eq. (8.2) and the relation

$$\exp[-\mathscr{F}/\tau] \equiv \sum_z \exp[-H(z)/\tau] \tag{8.3}$$

Given τ, the inverse rate at which a decision must be reached and acted on, a surprisingly subtle phase transition model emerges by identifying equivalence classes of a system's developmental pathways z, say, "useful" and "pointless". This allows definition of a symmetry groupoid for the cognition/action process [19]. A groupoid is a generalization of an algebraic group in which a product is not necessarily defined between each element. The simplest example is, perhaps, a disjoint union of separate groups, but sets of equivalence classes also define a groupoid. See [19] for details.

Given an information source associated with the system of interest, a full equivalence class algebra can be constructed by choosing different system origin states and defining the equivalence of subsequent states at a later time by the existence of a high probability path connecting them to the same origin state. Disjoint partition by equivalence class, analogous to orbit equivalence classes in dynamical systems, defines a symmetry groupoid associated with the cognitive process.

The equivalence classes across possible origin states define a set of information sources dual to different states available to the systems of interest. These create a large groupoid, with each orbit corresponding to an elementary "transitive" groupoid whose disjoint union is the full groupoid. Each subgroupoid is associated with its own dual information source, and larger groupoids must have richer dual information sources than smaller.

As argued in Chap. 1, the "free energy" Morse Function of Eq. (8.3) is liable to an analog of Landau's classical spontaneous symmetry breaking [18]. Under symmetry breaking, higher "temperatures" τ are associated with more symmetric higher energy states in physical systems.

Under time constraint, decline in the quantity τ from Eq. (8.2) will cause sharply punctuated collapse from higher to lower symmetry states, triggering ultimate system failure, modulated by both the embedding environment represented by the information source X and the culture/historical trajectory represented by Z.

Again, the "free energy" of Eq. (8.2) can be driven into a sudden, highly punctuated "phase transition" reflecting fundamental symmetry shifts by changes in τ, the inverse rate at which decision-and-action must take place. The symmetry shift, however, is between groupoids associated with a synergism across culture, contending/cooperating cognitive institutions, the embedding environment, and the structure of stochastic "large deviations". This is not simply the melting of ice.

In marked and fundamental contrast to the stable/unstable dichotomy associated with the previous treatment of anytime algorithms—Eq. (2.8)—phase change associated with $\mathscr{F}$ and τ may be relatively subtle, akin to the Eldredge and Gould [20] pattern of "punctuated equilibrium" in evolutionary transition described in Chap. 1. Recall that, under their model, evolutionary and coevolutionary dynamics undergo relatively long periods of apparent or near-stasis, where changes are small or difficult to see, followed by relatively sudden massive changes leading to fundamentally different coevolutionary configurations. Again, examples of the contemporary debate on punctuated equilibrium, now a central matter in evolutionary biology, can be found in [21].

The dynamics of the free energy analog $\mathscr{F}$ at critical values of τ can be studied using variants of the standard iterative renormalization techniques as developed for cognitive processes in [22].

The most direct general inference, however, remains that phase transitions associated with changes in τ will be expressed in terms of changes in underlying groupoid symmetries, leading again to some version of the punctuated equilibrium dynamics identified by Eldredge and Gould [20].

An explicit simplified model, assuming an integral approximation and a Gamma distribution across the right-hand side of Eq. (8.2), has been given in the Appendix to Chap. 7.

Note, however, that the underlying approach can be significantly extended through a multidimensional iteration. Suppose there is a set of τ_j across a number of essential modules. These can be characterized as "quality" measures, replacing the quality measure Q_i of the simple AA analysis. Strategy then involves setting up circumstances so that, at a point of need, one's own time frames for action remain long compared to those of an opponent. Then the complexities of Eq. (8.1) can be imposed on an adversary.

8.4 Discussion

We have outlined the beginnings of a general theory of emotional dynamics in the context of culture across individual humans, institutions, machine intelligence, and their many composite and "cockpit" entities. The theory involves the extension of current "anytime algorithm" analyses of machine cognition under time constraint. Emotions are for speed. Emotions are for survival. Accordingly, "emotional dysfunction" leading to inability to fulfill expected norms can always be imposed on a multidimensional cognitive system via a sufficiently crafted pattern of time limitation, "shadow price" and "noise", within the confines of both an embedding environment and a cultural milieu.

Always.

The central constraint on this conclusion is—again—the same as the objection the theoretical ecologist Pielou [23] raises to models in biology: models can propose questions for empirical and observational study. Mathematical models of complex biological and social phenomena do not answer questions.

This being said, the model developed here suggests that emotional dysfunction in humans, their organizations, and their man/machine composites, is necessarily ubiquitous, given the inevitable "noise" of life, both individually and in community. Times of particular stress, in a particular cultural context, will trigger specific patterns of individual social dysfunction that can be interpreted as "emotional" disorders. What is surprising is the possibility of extending the perspective across institutional, machine, and composites operating on Clausewitz landscapes of imprecision, uncertainty, and time constraint.

Just as, in the words of the evolutionary anthropologist Robert Boyd, "Culture is as much a part of human biology as the enamel on our teeth", institutions, intelligent machines, and their many cockpit and other composites, are themselves very particular cultural artifacts that may be subject to culturally characteristic gestalt patterns of behavior and "emotional" pathology. Such pathology may, in fact, be induced by a culturally informed external agent: recall, again, the words of Bin Ladin [24],

> All that we have to do is send two mujahidin to the furthest point east to raise a piece of cloth on which is written al-Qaida, in order to make the generals race there to cause America to suffer human, economic, and political losses without their achieving for it anything of note…[Just] as we…bled Russia for 10 years, until it went bankrupt and was forced to withdraw in defeat…[s]o we are continuing this policy in bleeding America to the point of bankruptcy.

References

1. Minsky, M. 1986. *The society of mind*. New York: Simon and Schuster.
2. Darwin, C. 1872. *The expression of the emotions in man and animals*. London.
3. Henrich, J., S. Heine, and A. Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33: 61–135.
4. Cowen, A., and D. Keltner. 2017. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *PNAS*. www.pnas.org/cgi/doi/10.1073/pnas.1702247114.
5. Kleinman, A. 1991. *Rethinking psychiatry: From cultural category to personal experience*. New York: Free Press.
6. Wallace, R. 2017. *Computational psychiatry: A systems biology approach to the epigenetics of mental disorders*. Berlin: Springer.
7. Zilberstein, S. 1996. Using anytime algorithms in intelligent systems. *AI Magazine* 17: 73–83.
8. Appleby, J., X. Mao, and A. Rodkina. 2008. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control* 53: 126–132.
9. Wallace, R., L.F. Chaves, L. Bergmann, C. Ayres, L. Hogerwerf, R. Kock, and R.G. Wallace. 2018. *Clear-cutting disease control: Capital-led deforestation, public health austerity, and vector-borne infection*. Berlin: Springer.
10. Atlan, H., and I. Cohen. 1998. Immune information, self-organization, and meaning. *International Immunology* 10: 711–717.
11. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
12. Khinchin, A. 1957. *Mathematical foundations of information theory*. New York: Dover Publications.
13. Wallace, R. 2018. *Carl von Clausewitz, the fog-of-war, and the AI revolution: The real world is not a game of go*. Berlin: Springer.
14. Wallace, R. 2018. New statistical models of nonergodic cognitive systems and their pathologies. *Journal of Theoretical Biology* 436: 72–78.
15. Feynman, R. 2000. *Lectures on computation*. New York: Westview Press.
16. Champagnat, N., R. Ferriere, and S. Meleard. 2006. Unifying evolutionary dynamics: From individual stochastic process to macroscopic models. *Theoretical Population Biology* 69: 297–321.
17. Dembo, A., and O. Zeitouni. 1998. *Large deviations and applications*, 2nd ed. Berlin: Springer.
18. Pettini, M. 2007. *Geometry and topology in hamiltonian dynamics and statistical mechanics*. Berlin: Springer.
19. Weinstein, A. 1996. Groupoids: Unifying internal and external symmetry. *Notices of the American Mathematical Association* 43: 744–752.
20. Eldredge, N., and S. Gould. 1972. Punctuated equilibrium: An alternative to phyletic gradualism. In *Models in paleobiology*, ed. T. Schopf, 82–115. San Francisco: Cooper and Co.
21. Lyne, J., H. Howe. 2017. Punctuated equilibrium: Rhetorical dynamics of a scientific controversy, chapter 6. In *Landmark essays on rhetoric of science: Case studies*, Second Edition, ed. Harris, R., CRC Press, Boca Raton.
22. Wallace, R. 2005. *Consciousness: A mathematical treatment of the global neuronal workspace model*. Berlin: Springer.
23. Pielou, E.C. 1977. *Mathematical ecology*. New York: Wiley.
24. Bin Ladin, U. 2004. https://www.aljazeera.com/archive/2004/112008491633364 57223.html.

Chapter 9
Expected Unexpecteds: Cambrian Explosions in Lamarckian Systems

9.1 Introduction

Hinote [1] succinctly characterizes the Second Gulf War and its aftermath:

> With the help of a creative and comprehensive leadership attack strategy, the Coalition accomplished a remarkable feat [in 2003]. It invaded and conquered a country situated halfway around the world while being massively outnumbered on the ground. Indeed, it used the advantages of a small force, including speed and flexibility, coupled with numerous qualitative advantages to disorient and overwhelm the Iraqi regime. Unfortunately, when the collapse of this regime led to chaos, the lack of numbers of ground troops turned into a disadvantage. The result was a lost opportunity to demonstrate how the future would represent real improvement over Saddam.

In this context, by mid-2014, the fundamentalist group ISIL had, with surprisingly little effort, come to control a vast sector of Iraq, including the second-largest city, Mosul. In addition, another eight or so significant anti-government militias had emerged and continued to operate throughout the country.

It wasn't until July 2017 that Iraqi army forces recaptured Mosul, and the last two ISIL strongholds in Iraq were only captured in November, 2017.

Casualties in this outburst were appalling: perhaps a hundred thousand direct deaths, mostly civilians, and five to six million persons displaced, a matter that will produce a further avalanche of morbidity and mortality.

Here, we will reconsider and extend ideas introduced at the very end of Chap. 5, finding such outbreaks to be, in a sense, "expected outliers" in the ongoing routine of punctuated equilibrium evolutionary process.

Suppression of possible outliers should, of course, be a central matter of conflict strategy before the conflict has begun, as, in a certain rough way, Japan learned after Pearl Harbor, both Napoleon and the Nazi State found after invasions of Russia, and the US should have learned—but did not—in Southeast Asia.

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_9

9.2 Some Hard-Core Evolutionary Theory

We expand and reinterpret the results of [2].

Over the years, evolutionary scientists have explored in detail the "Cambrian explosion" of half a billion years ago, a punctuated change in the diversity of life (e.g., [3–6]) sometimes seen in popular and religious literature as challenging evolutionary theory. Here we demonstrate, from a somewhat novel perspective, that a relatively modest formal exercise in that theory accounts neatly for such "explosions" early on, before path-dependent lock-in of essential biochemical, gene regulation, and more generally biological, Bauplans. The approach illuminates as well the current rapid evolution of viral/viroid species and quasi-species [7]. The ISIL outbreak can, perhaps, be viewed as representing such a viral/viroid proliferation.

The underlying mechanisms appear much the same. Similar discussions, under the rubric "punctuated equilibrium", have long been in the literature (e.g., [4]), and references therein). The work here supports that view, but provides a new theoretical line of argument.

The approach follows that of [8], where it is argued multiple punctuated ecosystem regime changes in metabolic free energy broadly similar to the aerobic transition enabled a punctuated sequence of increasingly complex genetic codes and protein translators. Then, in a manner similar to the serial endosymbiosis effecting the eukaryotic transition, codes, and translators coevolved until the ancestor of the present narrow spectrum of protein machineries became locked-in by evolutionary path dependence at a relatively modest level of fitness reflecting a modest embedding metabolic free energy ecology [9]. The search for such preaerobic biochemical Cambrian-like "explosions" is, of course, much hampered by the absence of early chemical evolution from currently studied fossil records.

Population genetics defines evolution by changes in allele frequencies [10, 11]. Evolutionary game dynamics track such shifts under natural selection using the replicator model of Taylor and Jonker [12]. These and related mathematical models purport to be both a necessary and sufficient definition of evolution across disciplines from biology to economics, albeit with sometimes scathing dissent (e.g., [13]).

Wallace [2, 8, 14–17], in contrast, proposes a set of necessary conditions statistical models extending evolutionary theory via the asymptotic limit theorems of communication theory. The method represents genetic heritage, regulated gene expression, and the surrounding environment as interacting information sources. A fundamental insight is that gene expression can be directly seen as a cognitive phenomenon associated with a "dual" information source, while the embedding environment's systematic regularities "remember" imposed changes, resulting in a coevolutionary process in the sense of [18] that is recorded jointly in, genes, gene expression regulation, and the embedding environment. See [2, 8, 14–17, 19] for details.

The focus here is on the effect of "large deviations" representing transitions between the quasi-stable modes of such systems that are analogous to game-theoretic Evolutionary Stable Strategies. Evolutionary path dependence, in general, limits such possible excursions to high probability sequences consistent with, if not originating in, previous evolutionary trajectories: after some three billion years, however, most multicellular organisms evolve, they retain their basic Bauplan, with only relatively small nonfatal variations currently allowed.

Conventional evolutionary theory is interested in matters half a billion years ago, before path dependence solidly locked-in possible large deviations excursions. We will attempt to extend these ideas to the "expected unexpecteds" of Lamarckian dynamics on Clausewitz Landscapes.

In essence, a sufficiently large number of allowed large deviation trajectories lead to many available quasi-equilibrium states. These, in turn, can be treated as an ensemble, i.e., in a manner similar to the statistical mechanics perspective on critical phenomena. This allows a new approach to rapid evolutionary change—in deep time for multicellular organisms, and in real time for current populations of viruses and viroids—and political entities in the context of conflict.

That is, even today, while incorporation of long-term path dependence drastically reduces possible evolutionary dynamics in higher organisms, viral or viroid evolution can be explored in the same way, driven by "noise" defined as much by policy and socioeconomic structure as by reassortment and generation time (e.g., [26]). "Political" outbreaks will be driven in parallel with diseases.

Following [2, 14], assume there are n populations interacting with an embedding environment represented by an information source Z. The genetic and (cognitive) gene expression processes associated with each species i are represented as information sources X_i, Y_i respectively. These information sources undergo a "coevolutionary" interaction in the sense of [18], producing a joint information source uncertainty [20] for the full system as

$$H(X_1, Y_1, \ldots, X_n, Y_n, Z) \tag{9.1}$$

In addition, Feynman's insight that information is a form of free energy allows definition of an entropy-analog as

$$S \equiv H - Q_j \sum_j \partial H / \partial Q_j \tag{9.2}$$

The Q_i are taken as driving parameters that may include, but are not limited to, the Shannon uncertainties of the underlying information sources. See [20] for a basic introduction to information theory.

Again, in the spirit of [18], we can characterize the dynamics of the system in terms of Onsager-like nonequilibrium thermodynamics in the gradients of S as the set of stochastic differential equations [21],

$$\begin{aligned} dQ_t^i = {} & L_i(\partial S/\partial Q^1 \cdots \partial S/\partial Q^m, t)dt + \\ & \sum_k \sigma_k^i(\partial S/\partial Q^1 \cdots \partial S/\partial Q^m, t)dB_k \end{aligned} \tag{9.3}$$

where the B_k represent noise terms having particular forms of quadratic variation. See [22] or other standard references on stochastic differential equations for details.

This can be more simply written as

$$dQ_t^i = L_i(\mathbf{Q}, t)dt + \sum_k \sigma_k^i(\mathbf{Q}, t)dB_k \tag{9.4}$$

where $\mathbf{Q} \equiv (Q^1, \ldots, Q^m)$.

Following the arguments of [18], this is a coevolutionary structure, where fundamental dynamics are determined by component interactions:

1. Setting the expectation of Eq. (9.4) equal to zero and solving for stationary points gives attractor states since the noise terms preclude unstable equilibria. These are analogous to the evolutionarily stable states of evolutionary game theory.
2. This system may, however, converge to limit cycle or pseudorandom "strange attractor" behaviors similar to thrashing in which the system seems to chase its tail endlessly within a limited venue—the "Red Queen".
3. What is "converged" to in any case is not a simple state or limit cycle of states. Rather it is an equivalence class, or set of them, of highly dynamic information sources coupled by mutual interaction through crosstalk and other interactions. Thus "stability" in this structure represents particular patterns of ongoing dynamics rather than some identifiable static configuration.
4. Applying Ito's chain rule for stochastic differential equations to the $(Q_t^j)^2$ and taking expectations allows calculation of variances. These may depend very powerfully on a system's defining structural constants, leading to significant instabilities [23].

As Champagnat et al. [18] note, shifts between the quasi-equilibria of a coevolutionary system can be addressed by the large deviations formalism. The dynamics of drift away from trajectories predicted by the canonical equation can be investigated by considering the asymptotic of the probability of "rare events" for the sample paths of the diffusion.

"Rare events" are the diffusion paths drifting far away from the direct solutions of the canonical equation. The probabilities of such rare events are governed by a large deviation principle, driven by a "rate function" $\mathscr{I}$ that can be expressed in terms of the parameters of the diffusion.

This result can be used to study long-time behavior of the diffusion process when there are multiple attractive singularities. Under proper conditions, the most likely path followed by the diffusion when exiting a basin of attraction is the one minimizing the rate function $\mathscr{I}$ over all the appropriate trajectories.

An essential fact of large deviations theory, however, is that the rate function $\mathscr{I}$ almost always has the canonical form

$$\mathscr{I} = -\sum_j P_j \log(P_j) \tag{9.5}$$

for some probability distribution, i.e., the uncertainty of an information source [20]. The result goes under a number of names; Sanov's Theorem, Cramer's Theorem, the Gartner–Ellis Theorem, the Shannon–McMillan Theorem, etc. [24].

The argument directly complements Eq. (9.4), now seen as subject to large deviations that can themselves be described as the output of an information source L_D defining $\mathscr{I}$, driving or defining Q^j-parameters that can trigger punctuated shifts between quasi-stable system modes.

This is now a common perspective in systems biology (e.g., [25]).

Not all large deviations are possible: only those consistent with the high probability paths defined by the information source L_D will take place.

Recall from the Shannon–McMillan Theorem [26] that the output streams of an information source can be divided into two sets, one very large that represents nonsense statements of vanishingly small probability, and one very small of high probability representing those statements consistent with the inherent "grammar" and "syntax" of the information source. Again, whatever higher order multicellular evolution takes place, some equivalent of backbone and blood remains.

Thus we could now rewrite Eq. (9.1) as

$$H_L(X_1, Y_1, \ldots, X_n, Y_n, Z, L_D) \tag{9.6}$$

where we have explicitly incorporated the "large deviations" information source L_D that defines high probability evolutionary excursions for this system.

Again carrying out the argument leading to Eq. (9.4), we arrive at another set of quasi-stable modes, but possibly very much changed in number; either branched outward in time by a wave of speciation, or decreased through a wave of extinction. Iterating the models backwards in time constitutes a cladistic or coalescent analysis.

A simple extinction model provides one route to significant extension of the theory.

Let $N_t \geq 0$ represent the number of individuals of a particular species at time t. The simplest dynamic model, in this formulation, is then something like

$$dN_t = \alpha N_t(1 - \frac{N_t}{K})dt + \sigma N_t dW_t \tag{9.7}$$

where K is the ecological carrying capacity for the species, α is a characteristic growth rate constant, σ is a "noise" index, and dW_t represents the white noise process.

Applying the Ito chain rule [22] to $\log(N_t)$, one obtains, as a consequence of the added Ito correction factor, the long-time endemic limits

$$\begin{aligned} &N_t \to 0, \alpha < \frac{\sigma^2}{2} \\ &N_t \to K(1 - \frac{\sigma^2}{2\alpha}), \alpha \geq \frac{\sigma^2}{2} \end{aligned} \tag{9.8}$$

That is, from the first part of Eq. (9.8), unless the rate of population growth is sufficiently large, noise-driven fluctuations will inevitably drive the species to extinction.

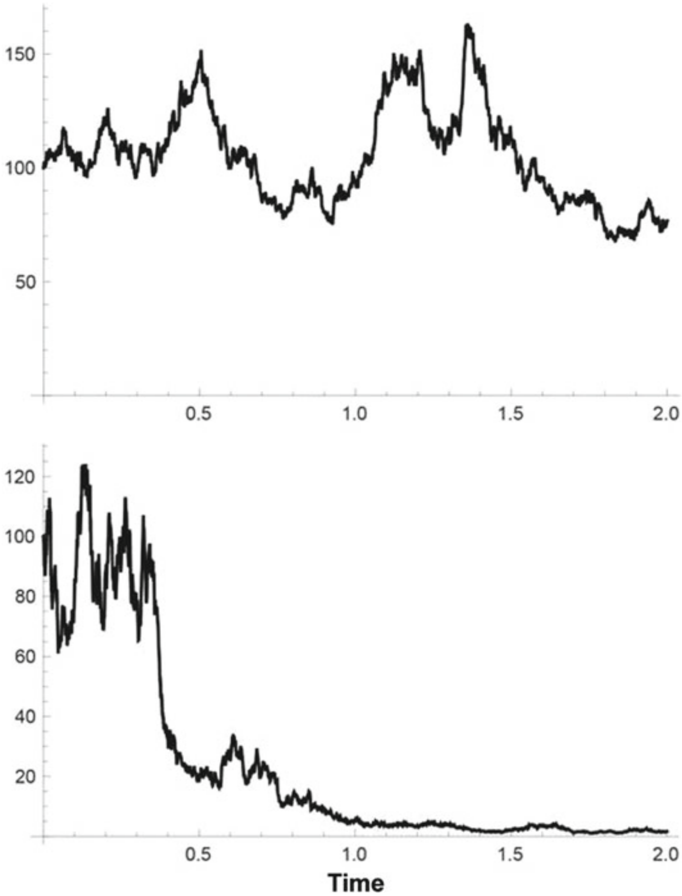


Fig. 9.1 Simulating N_t based on the Ito chain rule expansion of $\log(N_t)$ using Eq. (9.7). The simulations apply the ItoProcess function in the computer algebra program Mathematica for white noise. $N_0 = 100$, $K = 100$, $\alpha = 1$, $\sigma = 0.5, 1.5$. The critical value for σ is $\sqrt{2}$. 2000 time steps. While the upper trace fluctuates at values near K, the lower collapses

Figure 9.1 shows two simulations, with σ below and above criticality. For the first, the population fluctuates about an endemic level near carrying capacity, while the second collapses.

A similar stochastic differential equation approach has been used to model noise-driven criticality in physical systems [27–29], suggesting that a more conventional phase transition methodology may provide particular insight.

In general, for current higher plants and animals, the number of quasi-equilibria available to the system defined by Eq. (9.4), or to its generalization via Eq. (9.6), will be relatively small, a consequence of long-term lock-in by path- dependent evolutionary process. The same cannot be said, however, for virus/viroid species or quasi-species, to which can be applied more general methods (e.g., [7]) that may also represent key processes acting half a billion years in the past.

Under such a relaxation assumption, the speciation/extinction large deviations information source L_D is far less constrained, and there will be very many possible quasi-stable states available for transition, analogous to an ensemble in statistical mechanics.

The noise parameter σ in Eq. (9.7) can then be interpreted as a kind of inverse temperature-analog, and the full set of the N_t as a sort of complexity index representing a "higher symmetry" that vanishes below a critical value of $\tau \equiv 1/\sigma$. This leads to a relatively simple statistical mechanics analog built on the H_L of Eq. (9.6).

Define a pseudoprobability for quasi-stable mode j as

$$P_j = \frac{\exp[-H_L^j/\tau]}{\sum_i \exp[-H_L^i/\tau]} \tag{9.9}$$

Next, define a Morse Function F, in the sense used by Pettini [30], as

$$\exp[-F/\tau] \equiv \sum_i \exp[-H_L^i/\tau] \tag{9.10}$$

Apply Pettini's topological hypothesis to F, taking the set of the N_j, the numbers of members of species (or quasi-species) j as a symmetry index, in Landau's sense [31]. Then τ is seen as a very general temperature-like measure whose changes drive punctuated topological alterations in the underlying ecological and coevolutionary structures associated with the Morse Function F. High τ implies a "rich" population structure—one characterized by many species fragments that undergo further evolution and coevolution, much of it through conflict and other forms of "selection".

Such topological changes, following Pettini's arguments, can be far more general than indexed by the simple Landau-type critical point phase transition.

Indeed, the results of [8], regarding the complexity of the genetic code, could be directly reframed in terms of available metabolic free energy intensity leading to something like Eqs. (9.9) and (9.10). Then the term in τ is replaced by M, a measure of metabolic free energy intensity, and the H_L^j represent the Shannon uncertainties in the transmission of information between codon machinery and amino acid machinery.

Increasing M then leads to more complex genetic codes, i.e., those having higher measures of symmetry, in the Landau sense, until evolutionary lock-in took place at a relatively low level of coding efficiency. Canfield et al. [9], in their Tables 1 and 2, provide a range of possible electron donor and acceptor mechanisms that may have been available to fuel metabolic free energy under preaerobic conditions, in the context of anoxygenic phototrophs. They speculate that the most active early ecosystems were probably driven by the cycling of H_2 and Fe^{2+}, providing relatively low free energy intensities for metabolic process.

The topological changes inherent in a relaxed path-dependence model can represent a great variety of fundamental and highly punctuated coevolutionary alterations, since the underlying species and quasi-species are not so sharply limited by path-dependent evolutionary trajectory in the manner that constrains variation in most current higher organisms.

That is, under an assumption of less lock-in constraint a half-billion years ago, such a model accounts well for the observed Cambrian Bauplan explosion, and possibly as well for much earlier multiple "explosions" in genetic codes postulated by [8]. Implicit also is the possibly explosive coevolution of the environmental information source Z, for example, via the aerobic transition triggered by cyanobacteria.

In general, according to our model, the degree of punctuation in evolutionary punctuated equilibrium [4] will depend strongly on the richness of the distribution of quasi-equilibria to be associated with Eq. (9.4). "Cambrian explosions" therefore require a dense statistical ensemble of them, our analog to the "roughening of the fitness landscape" described as necessary in [5].

The "noise", in the Cambrian case, may have been a kind of species or quasi-species isolation, via separation by geographic or ecological niche, that was (relatively) suddenly lifted. Permitting greater interaction—lowering σ, increasing τ—may then have triggered a long series of rapid coevolutionary transitions, an inverse, as it were, of the observations of [32] on species extinction via habitat fragmentation.

As Marshall put it [5],

> With the advent of ecological interactions between macroscopic adults…especially…predation…, the number of needs each organism had to meet must have increased markedly: Now there were myriad predators to contend with, and a myriad number of ways to avoid them, which in turn led to more specialized ways of predation as different species developed different avoidance strategies, etc…The combinatoric richness already present in the Ediacaran genome was extracted through the richness of biotic interaction as the Cambrian "explosion" unfolded…

Other dynamics may also have contributed, e.g., [33]. In any event, 20–80 million year duration of the Cambrian "explosion" provides ample time for a broad-scale evolutionary divergence in multicellular life forms under conditions of relaxed path dependence. Similar explosive divergences in early evolution of the genetic code may also have occurred.

The general inference is that, in the absence of severe path-dependent lock-in, "Cambrian explosions" can be a common feature of blind evolutionary process, representing expected outliers in the ongoing routine of evolutionary punctuated equilibrium [4, 7].

But evolution on the Clausewitz Landscapes of this monograph's title is not entirely blind.

9.3 The Lamarckian Variant

Doctrine is the genome of armies that undergo, primarily, Lamarckian processes of evolution under a selection pressure that is in no small part determined by the cognitive dynamics, operational capabilities, and resources of a deadly adversary in an extraordinarily unfriendly environment. As described in earlier chapters, cognition requires choice, and choice implies the existence of an information source having restrictions of "grammar" and "syntax" on sequences of actions or behaviors. The genetic and cognitive information sources of Eqs. (9.1) and (9.6) are then those of doctrine and institutional cognition at and across the different scales and levels of organization under the influence of the environment Z.

Rather than "noise", social systems are dominated by Granovetter's "strength of weak ties" [34], representing the degree of nondisjunctive coupling between the different segments of the community. We can, then, take τ in Eqs. (9.9) and (9.10) as measuring that strength, which might be calculated in terms of some scalarization of such things as rates of communication, financial exchange, travel, migration, and so on, between and within geographic and/or social groupings. A high value of the composite τ then implies a highly "symmetric" cognitive groupoid across the different subgroups, implying a fully coherent structure in terms both of doctrine and active cognitive function.

Loss of cohesion—decline in τ below some threshold—then generates a phase transition leading to fragmentation in both doctrine and methods of operation. The different fragments then undergo further evolutionary and coevolutionary processes, diverging in their doctrinal genome, in a large sense.

Wallace and Fullilove [35] explore this dynamic in terms of the evolution of hyperviolence in criminal enterprise. As Wallace et al. [36] argue, on a disintegrating social network, violence becomes a key tool for communication, adding to social disintegration, and fueling even greater violence.

Things like ISIL are sadly predictable, and the possibilities of their occurrence should be routinely factored into strategic doctrine.

Indeed, one can model the first stages of a self-dynamic collapse of the strength of weak ties measure τ using tools introduced at the end of Chap. 5, assuming Eq. (9.10) can be approximated as an integral, and taking something of an inverse perspective.

That is, we assume a rich underlying structure, so that

$$\exp[-F/\tau] = \sum_k \exp[-H_k/\tau] \approx \int_0^\infty \exp[-H/\tau]dH = \tau \tag{9.11}$$

Then

$$F \approx -\tau \log[\tau] \tag{9.12}$$

leading to the definition of an entropy S as

$$S = F(\tau) - \tau dF/d\tau \approx \tau \tag{9.13}$$

As in Chap. 5, imposing an Onsager nonequilibrium thermodynamic model [21], the dynamic equation for τ becomes

$$\begin{aligned}
d\tau/dt &= \mu dS/d\tau \approx \mu \\
\tau &\approx \mu t + \tau_0
\end{aligned} \tag{9.14}$$

Under conditions of social disintegration, μ will be negative, leading to a self-dynamic decline in the strength of weak ties holding a polity together, triggering the Lamarckian Cambrian explosion. Changing the sign of μ—reversing the collapse dynamic—is a matter of very clever culturally informed policy, in combination with an exceedingly skilled use of resources.

We are not good at such programs.

An extension of the model that produces phase transitions in a "highly natural" manner is given in the Chapter Appendix.

Much of this, at least qualitatively, most readers will have realized some time ago, based on hard-won experience and expressed in a variety of vocabularies, but it may prove useful to have a "theoretical" argument that can apply to the future as well as to the past. After all, our Lords and Masters usually assert that, this time around, it won't be a quagmire.

A next stage in the development of theory would be to convert the probability models presented here more fully into statistical tools, roughly analogous to regression equations, for the analysis of real data and limited policy predictions. That work would be arduous and remains to be done.

9.4 Chapter Appendix: "Simple" Lamarckian Phase Transitions

Lamarckian phase change emerges "naturally" if we redefine the Morse Function of Eq. (9.11) in terms of an unknown function $g(\tau)$ so that

$$\exp[-\hat{F}/g(\tau)] = \sum_k \exp[-H_k/g(\tau)] \approx \int_0^\infty \exp[-H/g(\tau)]dH = g(\tau) \tag{9.15}$$

leading to the definition of a free energy as $\hat{F} = -g(\tau)\log[g(\tau)]$, an entropy-analog as $\hat{S} = \hat{F}(\tau) - \tau d\hat{F}/d\tau$ and an Onsager model as

$$d\tau/dt \propto d\hat{S}/d\tau = -\tau \left(-\left(\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} g(\tau)\right) \ln(g(\tau)) - \frac{\left(\frac{\mathrm{d}}{\mathrm{d}\tau} g(\tau)\right)^2}{g(\tau)} - \frac{\mathrm{d}^2}{\mathrm{d}\tau^2} g(\tau) \right) = f(\tau) \tag{9.16}$$

This has the general solution

$$
\begin{aligned}
g(\tau) &= \frac{X}{W(n, X)} \\
X &= -\left(C_1\,\tau - \tau \int \frac{f\,(\tau)}{\tau}\,\mathrm{d}\tau - C_2 + \int f\,(\tau)\,\mathrm{d}\tau\right)
\end{aligned}
\tag{9.17}
$$

where $W(n, X)$ is the Lambert W-function of order n in X.

For $n = -1$, $g(\tau)$ is real only in the limited region $-\exp[-1] < X < 0$. For $n = 0$ the reality range is $-\exp[-1] < X < \infty$.

At nonequilibrium steady state, $f(\tau) = 0$ and the solution reduces to

$$
g(\tau) = \frac{C_1\tau + C_2}{W(C_1\tau + C_2)} \tag{9.18}
$$

Three different parameter representations are shown in Fig. 9.2. Each line shows a different segmentation of the "response surface" depending on τ. The solutions are real only along the lines. The realm-segmentation is itself tunable, depending on the parameterizations of X and τ in terms of C_1 and C_2.

Typically, however, τ reaches its nonequilibrium steady-state condition according to some dynamic process, for example, an exponential so that

$$
\begin{aligned}
&d\tau/dt = f(\tau) = (\beta - \alpha\tau) \rightarrow 0 \\
&\tau \rightarrow \frac{\beta}{\alpha}
\end{aligned}
\tag{9.19}
$$

It is possible, on the basis of such a dynamic process, to model the "strength of weak ties" effect that links groups across a polity [34] in terms of a "noise" that fuzzes Eq. (9.19). Then volatility becomes important, leading to the stochastic differential equation for the dynamics of τ as

$$
d\tau_t = (\beta - \alpha\tau_t)dt + \sigma\tau_t dW_t \tag{9.20}
$$

The last term, in the Brownian noise dW_t, represents "smearing" due to connections across social divisions. Applying the Ito Chain Rule [22] to the stochastic variate τ^2 finds the necessary condition for stability in second order, stability in variance, is

$$
\alpha > \frac{\sigma^2}{2} \tag{9.21}
$$

If that condition is violated, then the punctuations indicated by Fig. 9.2 become impossible, and the "Cambrian explosion" fails. Nondisjunctive "weak ties" between subgroups—again, in the classic and well-studied sense of Granovetter [34]—are an important stabilizing force, in this model.

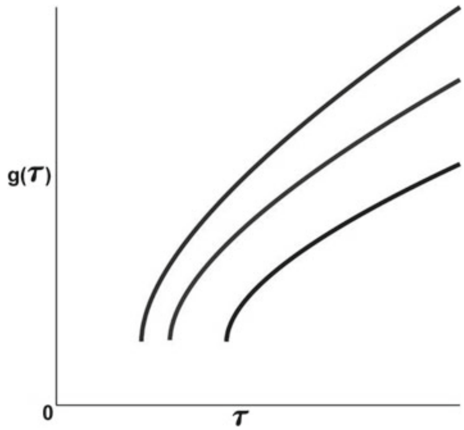


Fig. 9.2 Phases in the nonequilibrium steady-state Lamarckian system of Eq. (9.18) for three different system parameter sets in C_1 and C_2. $g(\tau)$ is real on the lines, but only over a very limited range of τ that is tunable but depends critically on parameterizations of X, $\tau(t)$, and the nature of $f(\tau)$ under stochastic perturbation. Sufficient strength-of-weak-ties "noise" linking subgroups will "fuzz over" the system and prevent "Cambrian explosions"

References

1. Hinote, S.C. 2013. Cutting off the head of the snake: Applying and assessing leadership attack in military conflict. Ph.D. thesis, School of Advanced Air and Space Studies, Air University, Maxwell Air Force Base, Alabama.
2. Wallace, R. 2013. A new formal perspective on 'Cambrian explosions'. *Comptes Rendus Biologies* 337: 1–5.
3. Erwin, D., and J. Valentine. 2013. *The cambrian explosion: The construction of animal biodiversity*. Greenwood Village: Roberts and Company.
4. Gould, S. 2002. *The structure of evolutionary theory*. Cambridge, MA: Harvard University Press.
5. Marshall, C. 2006. Explaining the cambrian 'explosion' of animals. *Annual Reviews of Earth and Planetary Science* 34: 355–384.
6. Whittington, H. 1985. *The burgess shale*. New Haven, CT: Yale University Press.
7. Wallace, R., and R.G. Wallace. 2015. Blowback: New formal perspectives on agriculturally driven pathogen evolution and spread. *Epidemiology and infection* 143 (SE10): 2068–2080.
8. Wallace, R. 2012. Metabolic constraints on the evolution of genetic codes: Did multiple 'preaerobic' ecosystem transitions entrain richer dialects via serial endosymbiosis? *Transactions on Computational Systems Biology XIV, LNBI* 7625: 204–232.
9. Canfield, D., M. Rosing, and C. Bjerrum. 2006. Early anaerobic metabolisms. *Philosophical Transactions of the Royal Society B* 361: 1819–1836.
10. Ewens, W. 2004. *Mathematical population genetics*. Berlin: Springer.
11. Hartl, D., and A. Clark. 2006. *Principles of population genetics*. Sunderland, MA: Sinaur Associates.
12. Taylor, P., and L. Jonker. 1978. Evolutionary stable strategies and game dynamics. *Mathematical Biosciences* 40: 145–156.
13. Roca, C., J. Suesta, and A. Sanchez. 2009. Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Physics of Life Reviews* 6: 208–249.

14. Wallace, R. 2011. A formal approach to evolution as self-referential language. *BioSystems* 106: 36–44.
15. Wallace, R. 2012. Consciousness, crosstalk, and the mereological fallacy: An evolutionary perspective. *Physics of Life Reviews* 9: 426–453.
16. Wallace, R. 2013. A new formal approach to evolutionary processes in socioeconomic systems. *Journal of Evolutionary Economics* 23: 1–15.
17. Wallace, R. 2013. Cognition and biology: Perspectives from information theory. *Cognitive Processing* 15: 1–12.
18. Champagnat, N., R. Ferriere, and S. Meleard. 2006. Unifying evolutionary dynamics: From individual stochastic process to macroscopic models. *Theoretical Population Biology* 69: 297–321.
19. Wallace, R., and D. Wallace. 2008. Punctuated equilibrium in statistical models of generalized coevolutionary resilience: How sudden ecosystem transitions can entrain both phenotype expression and Darwinian selection. *Transactions on Computational Systems Biology IX, LNBI* 5121: 23–85.
20. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
21. de Groot, S., and P. Mazur. 1984. *Non-equilibrium thermodynamics*. New York: Dover.
22. Protter, P. 1990. *Stochastic integration and differential equations*. Berlin: Springer.
23. Khasminskii, R. 2012. *Stochastic stability of differential equations*, 2nd ed. Berlin: Springer.
24. Dembo, A., and O. Zeitouni. 1998. *Large deviations: Techniques and applications*. Berlin: Springer.
25. Kitano, H. 2004. Biological robustness. *Nature Genetics* 5: 826–837.
26. Khinchin, A. 1957. *Mathematical foundations of information theory*. New York: Dover Publications.
27. Horsthemeke, W., and R. Lefever. 2006. Noise-induced Transitions. In *Theory and applications in physics, chemistry, and biology*, vol. 15. Berlin: Springer.
28. Van den Broeck, C., J. Parrondo, and R. Toral. 1994. Noise-induced nonequilibrium phase transition. *Physical Review Letters* 73: 3395–3398.
29. Van den Broeck, C., J. Parrondo, R. Toral, and R. Kawai. 1997. Nonequilibrium phase transitions induced by multiplicative noise. *Physical Review E* 55: 4084–4094.
30. Pettini, M. 2007. *Geometry and topology in hamiltonian dynamics and statistical mechanics*. Berlin: Springer.
31. Landau, L., and E. Lifshitz. 2007. *Statistical physics, part I*, 3rd ed. New York: Elsevier.
32. Hanski, I., G. Zurita, M.I. Bellocq, and J. Rybicki. 2013. Species-fragmented area relationship. *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1311491110.
33. von Bloh, W., C. Bounama, and S. Franck. 2003. Cambrian explosion triggered by geosphere-biosphere feedbacks. *Geophysical Research Letters* 30: 1963–1968.
34. Granovetter, M. 1973. The strength of weak ties. *American Journal of Sociology* 78: 1360–1380.
35. Wallace, R., and R. Fullilove. 2014. State policy and the political economy of criminal enterprise: Mass incarceration and persistent organized hyperviolence in the USA. *Structural Change and Economic Dynamics* 31: 17–31.
36. Wallace, R., M. Fullilove, and A. Flisher. 1996. AIDS, violence, and behavioral coding: information theory, risk behavior, and dynamic process on core-group sociogeographic networks. *Social Science and Medicine* 43: 339–352.

Chapter 10
Reconsidering Clausewitz Landscape Dynamics

ןיסרפו לקת אנמ אנמ

10.1 Introduction

Here, we follow something of the classic Engineering Curve, from the highly complex models of the earlier sections, to a kind of stripped-down reexamination of the bare bones of the thing.

The Clausewitz landscapes that characterize Western concepts of armed conflict are seen as dominated by matters of imprecision and delay in action—friction—and by uncertainty of both one's own effect and deployment and of the intent and disposition of the adversary—the fog-of-war. Previous chapters have examined the dynamics of conflict on such landscapes, using a tool set not ordinarily applied to such problems. One must, after all, look at the elephant-in-the-room from many different angles, over different time scales, and in different ecosystems, to be able to predict its behavior under some specific set of conditions. Here, we will condense much of the earlier work to a kind of irreducible kernel, focusing again on institutional cognition, in a large sense.

The results are not encouraging.

10.2 The Models

Chapter 3 examined a complex "self-dynamic" model of institutional cognition on a Clausewitz Landscape. Recall that the rate of cognition of an institution was calculated in terms of the Rate–Distortion Function (RDF) [1] defined in terms of distortion

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_10

in the transmission of intent to actual effect. This argument follows Feynman [2] by interpreting the RDF as a "free energy" measure, leading to a dynamic model similar to the Arrhenius treatment of the rate of a chemical reaction.

Again, the Rate–Distortion Function, $R(D)$, determines the minimum channel capacity needed to keep the average distortion between what is sent and what is received below a real number limit $0 < D$, by some appropriate measure. The RDF, for any information channel, is always a convex function of the measure D [1].

Taking the Feynman perspective, we again define a Boltzmann-like probability measure defining a normalized, dimensionless, rate of cognition as

$$P[R \geq R_0] = \frac{\int_{R_0}^{\infty} \exp[-R/g(Z)]dR}{\int_0^{\infty} \exp[-R/g(Z)]dR} = \exp[-R_0/g(Z)] \tag{10.1}$$

where $g(Z)$ is a positive, monotonic increasing function of the rate Z at which some composite scalarization of essential resources is delivered. In general, $g(Z)$ must be determined in each case. R_0 is the threshold channel capacity necessary for initiation of a cognitive function, for recognition that a signal must elicit a response.

We again obtain, for the effectiveness—rate of cognition—of an institution the Arrhenius relation

$$F(Z) = \exp[-K/g(Z)] \tag{10.2}$$

for a system-specific parameter K.

We first take $g(Z) = Z$ and define "friction" in terms of an exponential rate of increase of Z with system time toward some asymptotic limit so that

$$\begin{aligned} Z(t) &= \frac{\beta}{\alpha}(1 - \exp[-\alpha t]) \\ dZ/dt &= \beta \exp[-\alpha t] = \beta - \alpha Z(t) \\ Z \rightarrow Z_\infty &= \frac{\beta}{\alpha} \end{aligned} \tag{10.3}$$

Then

$$\begin{aligned} dF/dt = dF/dZ \times dZ/dt = \\ F \times \frac{K}{Z^2}(\beta - \alpha Z) \end{aligned} \tag{10.4}$$

The corresponding stochastic differential equation [3], representing the onset of fog-of-war uncertainties in the context of delay characterized by the rate constant α—friction—is

$$dF_t = F_t \frac{K}{Z_t^2}(\beta - \alpha Z_t)dt + \sigma F_t dB_t \tag{10.5}$$

where dB_t is taken to represent Brownian noise.

Since, in this first model, $F_t = \exp[-K/Z]$, we can replace Z by the corresponding function in F, so that the SDE becomes

$$dF_t = KF_t(\frac{\log(F_t)}{K})^2(\beta + \frac{\alpha K}{\log(F_t)})dt + \sigma F_t dB_t \tag{10.6}$$

Applying the Ito Chain Rule [3] to F_t^2, some calculation finds that a necessary condition for second-order stability in this relation—stability in variance—is

$$\frac{K\alpha}{2Z_\infty} > \sigma^2 \tag{10.7}$$

In the presence of noise indexed by σ, the requirement for stability in variance under a condition of "exponential" friction is that the greater the need for resources, i.e., the higher Z_∞, the greater must be the synergism between the rate of cognition K and the rate at which the supply system can ramp up to meet demand, measured by α.

What happens for a general $g(Z)$?

Then

$$\begin{aligned} dF/dt = F \times \frac{K}{g(Z)^2} dg/dZ \times dZ/dt = \\ F\frac{K}{g(Z)^2} dg/dZ(\beta - \alpha Z) \end{aligned} \tag{10.8}$$

leading to the SDE

$$\begin{aligned} dF_t = F_t \frac{K}{g(Z_t)^2} dg/dZ(\beta - \alpha Z_t)dt + \sigma F_t dB_t \\ dF_t = F_t dY_t \end{aligned} \tag{10.9}$$

where Y_t is a stochastic variable.

The latter relation, which has an unstable nonequilibrium steady state (NSS) at $F_t \equiv 0$, again leads to the Doleans-Dade result [3]:

$$F_t \propto \exp[Y_t - 1/2[Y_t, Yt]] = \exp[Y_t - \frac{\sigma^2}{2}t] \tag{10.10}$$

where $[Y_t, Y_t]$ is again the stochastic variation of Y_t [3].

Again, somewhat heuristically, by the Mean Value Theorem, if $dY_t/dt < \sigma^2/2$, then F_t declines in probability to zero. Appleby et al. [4] make the argument rigorous: sufficiently large σ will stabilize the unstable NSS $F = 0$ under very general conditions.

Details will vary with the precise form of $g(Z)$, which can be found using the methods leading to Eq. (3.6), taking the right hand side of Eq. (3.5) as $\beta - \alpha Z$ for an exponential "friction".

The general result for $g(Z_\infty)$ is then

$$
\begin{aligned}
g(Z_\infty) &= \frac{X}{W(n, X)} \\
X \equiv \beta Z_\infty \log(Z_\infty) - Z_\infty(3/2\beta + C_1) &+ C_2
\end{aligned}
\tag{10.11}
$$

where $W(n, X)$ is the Lambert W-function of orders $n = 0$ or $n = -1$. $W(0, X)$ is real on the interval $-\exp[-1] < X < \infty$ and $W(-1, X)$ is real for $-\exp[-1] < X < 0$.

The author was unable to derive the exact analog to Eq. (10.7) for the extended model. Centrally, however, the appearance of the Lambert W-function implies the existence of a number of phase transitions, depending on n, β, Z_∞, and the initial conditions characterized by C_1 and C_2.

A more general expression is possible for the characterization of cognitive efficacy. Taking

$$
\begin{aligned}
F(Z) &\propto \exp[\int h(Z)dZ] \\
Z(t) &= \frac{\beta}{\alpha}(1 - \exp[-\alpha t])
\end{aligned}
\tag{10.12}
$$

some manipulation gives the SDE

$$
\begin{aligned}
dF_t = F_t h(Z_t)(\beta - \alpha Z_t)dt &+ \sigma F_t dB_t \\
dF_t &= F_t dY_t
\end{aligned}
\tag{10.13}
$$

leading again to the situation of Eq. (10.10).

For example, if, rather than Eqs. (10.1) and (10.2), we take an exponential model of an institutional "anytime algorithm", so that

$$
F = \frac{b}{a}(1 - \exp[-aZ])
\tag{10.14}
$$

then some calculation finds

$$
h(Z) = \frac{a \exp[-aZ]}{1 - \exp[-aZ]}
\tag{10.15}
$$

and the unstable nonequilibrium steady state $F \to 0$ is again "stabilized" by sufficiently large σ.

Slightly more detail can be wrung from this particular version of the general model.

Substituting Eq. (10.15) and the relation

$$Z = \frac{-1}{a} \log(1 - \frac{a}{b} F)$$

into Eq. (10.13) gives the SDE

$$dF_t = (b - aF_t)[\frac{\alpha}{a} \log(1 - \frac{a}{b} F_t) + \beta]dt + \sigma F_t dB_t \tag{10.16}$$

Applying the Ito Chain Rule to F_t^2 finds the condition for stability in variance to depend on the reality of the expression

$$W(n, -\frac{\sigma^2}{2\alpha} \exp[\frac{\beta a}{\alpha} - \frac{\sigma^2}{2\alpha}]) \tag{10.17}$$

where W is the Lambert W-function of order $n = 0, -1$. Boundary conditions suggest $n = 0$, and again, as above, there can be several distinct phases, depending in detail on parameter magnitudes. Recall that $W(X)$ is real-valued only for $-\exp[-1] < X < \infty$. For $\alpha = \beta = a = b = 1$, it is easy to show from Eq. (10.17) that the limit for real values of the variance is $0 \leq \sigma <\approx 0.56$.

We can, in a sense, come full circle by defining efficacy as

$$F(Z) = \exp[\int h(Z)dZ] \equiv \exp[-R/g(Z)] \tag{10.18}$$

Then, as in Eq. (3.3), *if it is actually possible to impose dynamics on the system*, we can again define a "free energy" analog as

$$\exp[-\mathscr{F}/g(Z)] \equiv \int_0^\infty \exp[-R/g(Z)]dR = g(Z)$$
$$\mathscr{F} = -g(Z)\log[g(Z)] \tag{10.19}$$

As in Eq. (3.4), we then can also define an "entropy analog" as

$$S(Z) \equiv \mathscr{F}(Z) - Z\dot{F}(Z)/dZ \tag{10.20}$$

Imposing, as in Eq. (3.5), an Onsager nonequilibrium steady state, i.e., $dZ/dt \propto dS(Z)/dZ \rightarrow 0$, recovers the solution as

$$g(Z) = \frac{X}{W(n, X)}$$
$$X = C_1 Z + C_2 \tag{10.21}$$

where, again, W is the Lambert W-function of order $n = 0, -1$, leading to "phase transitions" defined by the function's restrictions on real value.

Then

$$\int h(Z)dZ = \frac{-R}{g(Z)}$$

$$h(Z) = -Rd(1/g(Z))/dZ =$$

$$\frac{R\,(\mathrm{W}\,(n, C_1\,Z + C_2))^2\,C_1}{(C_1\,Z + C_2)^2\,(1 + \mathrm{W}\,(n, C_1\,Z + C_2))} \tag{10.22}$$

Further algebraic development rapidly becomes intractable.

10.3 Discussion

There are, of course, very significant limitations to this modeling exercise. First, it is predicated on a specific cultural vision of conflict, that of the Western Clausewitz Landscape of delay, imprecision-in-action, and overall uncertainty. And again, in the sense of Pielou [5], these are models about conflict, not statistical or other models of conflict. The word is never the thing.

What do these models suggest? Eq. (10.7), and its presumed extension under a general form of $g(Z)$, imply the possibility of punctuated "phase transition" events at tactical, operational, and strategic time frames, scales, and levels of organization on Clausewitz Landscapes. Eldredge and Gould [6] from Chap. 1 everywhere and everywhen. Avoiding "surprises" of extinction and speciation is, then, predicated on the exercise of conflict outside culturally Western Clausewitz landscapes dominated by friction and the fog-of-war.

As one experienced commentator is reported to have put it, "There are two ways to fight the United States: Asymmetric and stupid."

The East Asian conflict metaphors as described by Jullien [7] also come to mind.

As do the experiences of the armies of Japan and of the Nationalists in China, of the French and Americans in Vietnam, of the British Empire, the USSR, and NATO in Afghanistan.

And yet, and yet…East Asian conflict styles too may not prevail in the face of "third stream" strategies, as has been explored in Chap. 5 using a different formalism. But of course, no fixed doctrinal prescriptions are available for risk-free management of complex sociocultural dynamics among varied contending human institutions or their machine-aided composites. To mangle, somewhat, the East Asian perspective, a flowing river tomorrow will always be different from that river today. Institutions learn and adapt.

Most readers of this monograph will have reached similar conclusions some time ago by much different means.

The implications for conduct of large-scale armed conflict, in the context of the Second Hundred Year's War, and as was summarized long ago for a leader consumed by the hubris of power, will also be familiar: "Mene Mene Tekel Upharsin."

References

1. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
2. Feynman, R. 2000. *Lectures on computation*. New York: Westview Press.
3. Protter, P. 2005. *Stochastic integration and differential equations*, 2nd ed. Berlin: Springer.
4. Appleby, J., X. Mao, and A. Rodkina. 2008. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control* 53: 683–691.
5. Pielou, E. 1977. *Mathematical ecology*. New York: Wiley.
6. Eldredge, N., and S. Gould. 1972. Punctuated equilibrium: An alternative to phyletic gradualism. In *Models in paleobiology*, ed. T. Schopf, 82–115. San Francisco: Cooper and Co.
7. Jullien, F. 2004. *A Treatise on efficacy: Between western and chinese thinking*. Honolulu: University of Hawaii Press.

Chapter 11
Failure of a Paramilitary System: A Case History of Catastrophe

11.1 Introduction

The previous chapters argued that something much like a highly punctuated phase transition must be common at various scales in extended conflict. At the tactical level, this is expressed as sudden collapse consequent on a failure of control, even if the units involved are not simply overrun. At the strategic level, Larmarckian versions of Eldredge/Gould punctuated equilibrium may dominate. The implication is that cognitive networks, which are inherently information networks, may often, if not always, be subject to "symmetry breaking" transitions, involving changes in underlying group or, more likely, groupoid structures.

Network phase transitions are particularly important in cognitive systems.

Here, we will examine the catastrophic failure of an urban fire service system following extensive reconfiguration based on a Rand Corporation operations research model that omitted, among other things, recognition of phase transitions inherent to traffic flow on road networks. Although urban fire service, a hierarchical paramilitary structure, does not routinely confront pushback from an actively cognitive enemy, study of that failure can provide some insight into punctuated equilibrium and other phase transitions affecting contention on Clausewitz landscapes.

11.2 The Rand Fire Models

In the late 1960s. the Rand Corporation, via an entity known as the New York City Rand Institute, and having close connections with the Operations Research Department of the Columbia University Engineering School, was given Carte Blanche to overhaul New York City's fire service delivery system, making it "more efficient", although FDNY only consumed some 2% of the city budget and was under considerable stress from a rising tide of fire and abandonment. See Wallace and Wallace [1]

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_11

for details. That book, written under an Investigator Award in Health Policy Research from the Robert Wood Johnson Foundation, explores a number of critical mismatches between Rand's grotesquely simplistic operations research approach and the underlying realities. Indeed, it is unlikely anything of the abysmal quality of the Rand models would have been allowed at the time for the management of natural populations.

The essential idea of the Rand work was the use of the *model-calculated* travel time of the first responding unit from firehouse to the street alarm box nearest the fireground as a surrogate global measure of fire service quality, and to equalize service across neighborhoods within "hazard regions" by eliminating units, not by adding them to an already highly stressed paramilitary structure.

From the beginning, failure to use actual damage measures as the principal service index doomed the Rand enterprise. In addition, while "response time" of an ambulance is important in taking a sick individual to hospital, a fire extinguishment system must build a hospital around a "patient" getting sicker at a literally exponential rate. Surrogate—or even exact—measures of the "response time" of the first unit is beside the point.

It gets worse.

Since the models had available only map-defined euclidean distances from firehouse to firebox, it was necessary to create a travel distance/travel time model for a highly irregular street network subject to both the vagaries of weather, the public transit system, and the daily commuting cycle.

Figure 11.1, from Rand [2], displays a typical "empirical model fitting" exercise that supposedly establishes such a relationship, here, across the full Trenton, NJ road network in 1975. Rand collapses an obviously diffuse pattern into a simple "square root-linear" model. Wallace and Wallace ([3], pp. 33–34) show two similar NYC examples as done by Rand.

For contrast, Fig. 11.2, from [4], shows measures of traffic flow vs. linear vehicle density for road segments in Rome, Japan, and Flanders. A similar figure for the Minneapolis-St. Paul road network can be found in ([4], Fig. 3.2). The essential point is that, above a critical traffic density of vehicles/unit length, the "fundamental diagram" of traffic flow undergoes a "phase transition" into irregular, clumped, and jammed conditions that are not predictable. At high traffic densities, these changes can be analogous to the sudden freezing of a "supercooled" fluid when triggered by some minor shock. Since the 1930s, many of the finest minds of science, engineering and mathematics have been, and remain, profoundly challenged by the details of traffic phase transitions.

For traffic flow on road networks, as opposed to simple street segments, Geroliminis and Sun [5] caution that there is an essential problem in defining a "macroscopic fundamental diagram" (MFD) like Fig. 11.1:

> ...[F]reeway networks do not have well-defined MFDs between network flow and density, as these networks have topological or control characteristics that are different from arterial networks... [R]esearch is needed in different types of networks to understand how variations in the topology/structure of the networks can affect the shape, the scatter and the existence of an MFD... MFD's should not be universally expected... a careful analysis is necessary before... control strategies/policies are introduced based on monitoring aggregated variables.

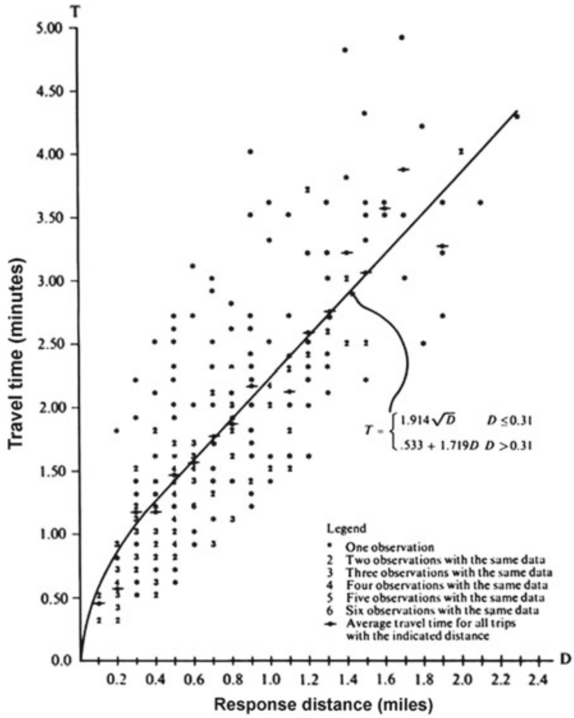


Fig. 11.1 Adapted from Fig. 6.4 of [2]. Relation between fire company travel time and response distance for the full Trenton, NJ road network, 1975. The Rand Fire Project collapsed macroscopic traffic turbulence into a grossly simplistic "square root-linear" model used to design fire service deployment. Real traffic flow is highly dependent on nonlinear "phase transitions" driven by traffic density and road conditions. See Fig. 11.2

In the early 1970s, however, the NYC/Rand Institute "solved" the intractable network traffic flow problem by simply collapsing a nearly random data point cloud onto a "square root-linear" model, and used that model as a basis for closing fire companies in high fire incidence, high population density neighborhoods of New York City that was already undergoing the first stages of a highly nonlinear contagious process of fire and abandonment.

Basic insight into the MFD problem is, in fact, surprisingly easy to obtain. Although by no means a complete analysis, a direct calculation clearly indicates the shoal against which any simple OR model will founder. The approach is via classic network theory. The key is to view a road net as a random graph made up of nodes at which roads intersect and the "edges" which connect them. Edges are identified as "Open" if they are in free flow, and "Closed" if congested. Suppose the road net graph has M vertices and $m = (1/2)aM$ closed edges chosen at random.

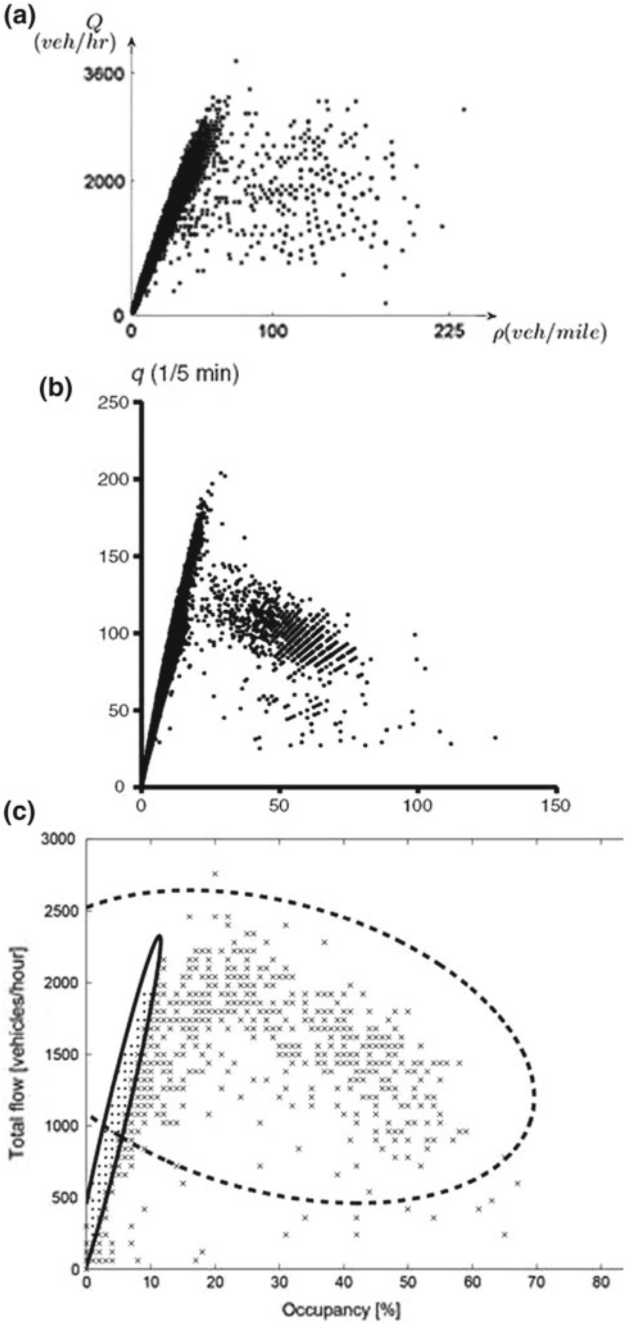


Fig. 11.2 **a** Vehicles per hour as a function of vehicle density per mile for a street in Rome [6]. **b** One month of data at a single point on a Japanese freeway, flow per five minutes versus vehicles per km. [7]. **c** 49 Mondays on a Flanders freeway [8]. All show a highly irregular "phase transition" to traffic jam conditions at a critical value of traffic density

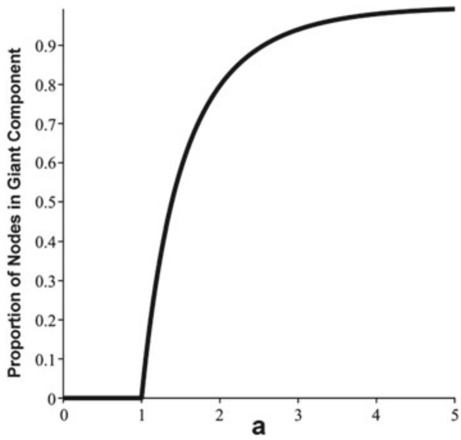


Fig. 11.3 Size of the largest connected component of jammed roads for a random graph. The traffic jam grows very sharply in total extent after the critical density is exceeded

Corless et al. [9] show that, for $a > 1$, the road net graph almost surely has a giant connected component—traffic jam—having $\approx g(a)M$ vertices with

$$g(a) = 1 + W(-a \exp[-a])/a \tag{11.1}$$

where W is the Lambert-W function defined implicitly by the relation $W(x) \exp[W(x)] = x$. See Fig. 11.3.

What happened when the Rand models were used to close or relocate some fifty firefighting units, mainly from high-fire, high population density minority neighborhoods?

11.3 The Catastrophe

Figure 11.4 is a composite index of structural fire—buildings—damage across New York City from 1959 through 1990. It is constructed from annual fire service data representing (i) number of building fires, (ii) number of "all hands" fires requiring a full complement of engines and ladders for extinguishment, and (iii) "extra alarm assignments" (EAA), counting a two-alarm fire as one EAA, and so on. The index is constructed as the largest component in a principal component analysis, counting zero fires as the zero index.

The rise after 1961 was countered, in 1964, by a program of a million fire inspections, a practice that could not be sustained beyond the year.

Parenthetically, when the city of Tokyo was confronted by a similar damage rise, the municipality doubled the fire department size, assigning units one day to fire inspection, and the next day to fire extinguishment. No such program was followed in New York.

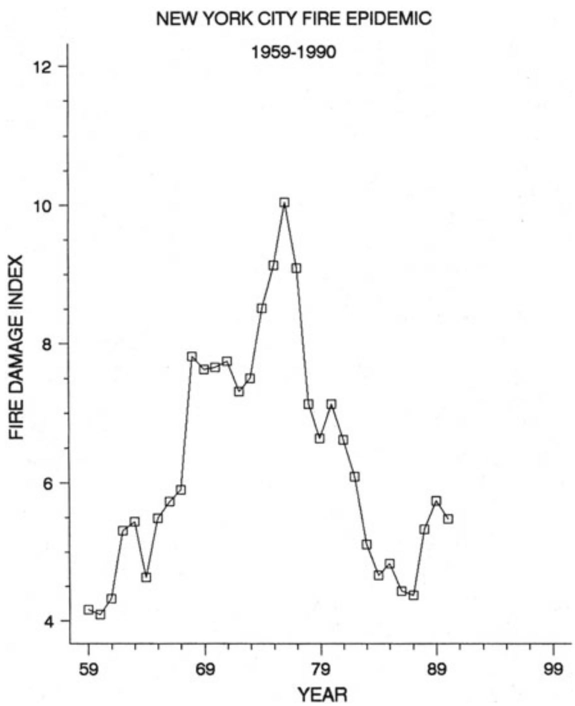


Fig. 11.4 Composite index of structural fire damage, 1959–1990, in New York City. The sudden rise from 1967 to 1968 impelled the fire service unions to successfully demand the opening of "second section" companies in the worst affected communities. Fire damage began declining until the Rand models were used to justify closing companies. The subsequent peak represents a contagious fire/abandonment epidemic that reduced housing stock to a level that could be supported by the reduced level of housing-related municipal services

The sudden rise from 1967 to 1968 was met by a lawsuit from the Uniformed Firefighters Association, resulting in a labor arbitration agreement resulting in the opening of twenty new fire companies, based as "second sections" for firehouses based in high fire incidence, high population density, overcrowded neighborhoods. These were, of course, primarily minority voting blocs.

Beginning in 1972, the Rand models were used as legal justification for closing or relocating some fifty firefighting units within those voting blocs. Under US law, it is not sufficient to show discriminatory impact of a public policy for redress. The plaintiff must show deliberate intent to discriminate. The Rand fire models provided an effective color-of-law for the removal of fire extinguishment services from minority voting blocs at a time of rapidly rising fire and abandonment.

Figure 11.5 shows, for the Bronx section of New York City between 1970 and 1980, housing loss to the contagious fire/abandonment process that was triggered by the Rand fire company closings. The Bronx, with 1.4 million residents, is one of the largest urban centers in the Western world. Vast areas lost between 30 and 80% of housing units to the destruction epidemic. Other parts of the city—almost entirely minority voting blocks—suffered similar losses. This is unprecedented for a Western nation in peacetime.

Figure 11.6 shows, for 1974–75, the resulting pattern of forced migration within New York City, using school transfer data.

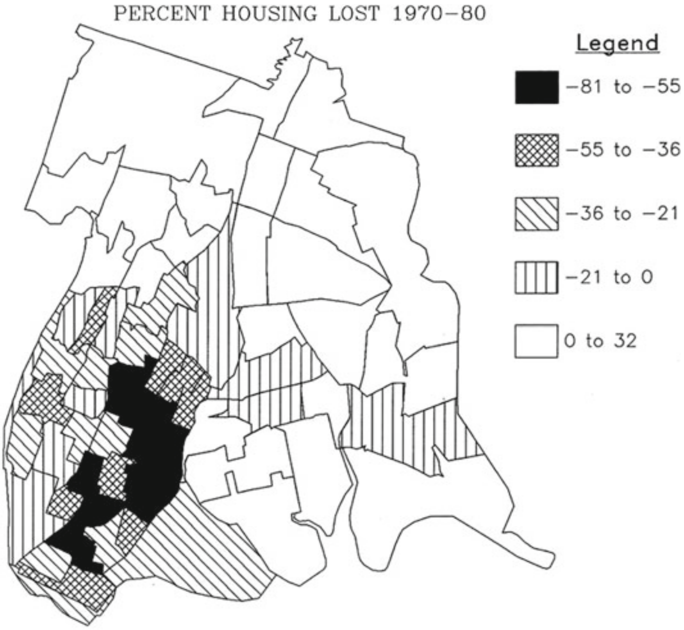


Fig. 11.5 Percent of housing lost in the Bronx by Health Area, 1970–1980. Such damage is unprecedented in a Western nation outside acts of war and, in fact, represents the successful outcome of a clandestine civil war carried out against minority communities

Between 1970 and 1980 whole neighborhoods were depopulated, and adjacent districts became de-facto refugee camps, with the full spectrum of refugee camp behavioral pathologies involving violence, low birthweight, infectious and sexually transmitted disease, and intergenerational demoralization.

Figure 11.7 shows the New York City homicide rate, in comparison to that of the US as a whole. Forced migration created levels of social disintegration for which only the most violent acts could be used as expressions of self-worth, creating further social disintegration in a positive feedback mechanism. In addition, population displacement triggered conflict between established illegal drug operations for access to new markets, causing "drug wars".

It is difficult to properly express the magnitude of this disaster. Wallace and Wallace [1], and the associated references, examine both the extent of the fire/abandonment epidemic and its impacts on public health and public order. Large districts of the city were reduced to Dresden-like rubble, and adjacent neighborhoods became de-facto refugee camps that required a full generation to return to even basic levels of civil order. It is estimated that, over a 30-year period, there were some 100,000 premature mortalities as a result of these catastrophic disruptions, not all concentrated within the city, since New York is both at the apex of the US urban hierarchy and at the center of the commuting field for the NY-NJ-CT conurbation. AIDS, tuberculosis, low weight birth, and violent crime cascaded from the city to

Fig. 11.6 NYC Planning Commission map showing the magnitude and direction of pupil transfers between Community School Districts for the school year 1974–75, time of maximum occupied structural fire worktime for the Bronx. Areas adjacent to burning minority voting blocs became de-facto refugee camps, with the full spectrum of refugee camp health and behavioral pathologies

its suburbs, and along national travel networks to other US conurbations. Indeed, characteristic NYC multiple-drug-resistant tuberculosis strains eventually diffused to countries of the European Union.

Figure 11.8 shows the regional dispersal of AIDS, tuberculosis, and violent crime from New York City into the 24-county New York Metropolitan Region, using a variant of the methodology described in Sect. 2.5. The "probability of contact matrix" was constructed using US Census data on daily job commutes within and between the 24 counties, in combination with a local index of susceptibility, the percent of the population living in poverty. As conditions deteriorated within New York City, behavioral pathologies and infectious disease diffused regionally out along the daily commuting field.

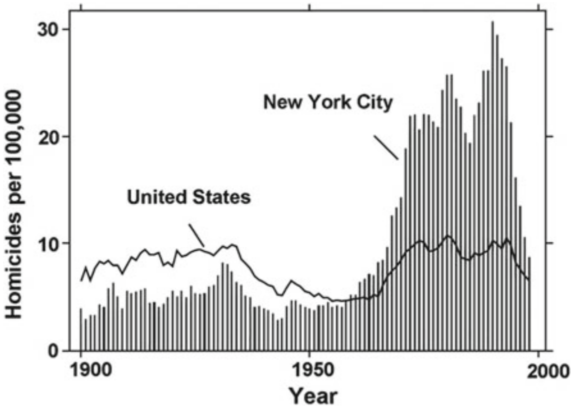


Fig. 11.7 Adapted from Monkonnen [10]. New York City homicide rate per 100,000, 1900–1995, compared to total USA rate. The first peak represents the initial effects of "refugee camp" behaviors triggered by forced migration. The second NYC peak, however, appears to represent the second wave of social disintegration subsequent to the initial burnout of minority communities, i.e., the "crack wars" that followed the dispersion of social networks needed for the socialization of the young and their entrainment into adult patterns of work and family

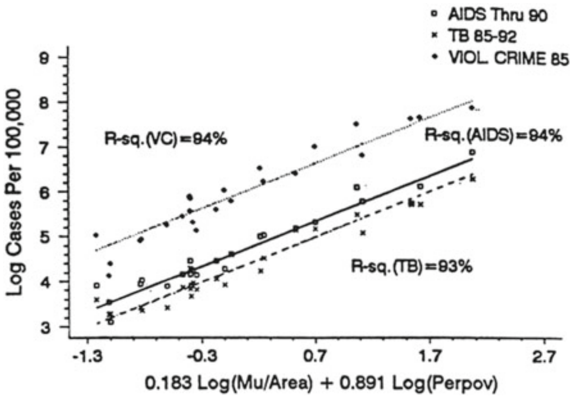


Fig. 11.8 From [1]. Multivariate analysis of covariance of AIDS cases, violent crimes, and tuberculosis cases for the 24-county New York Metropolitan Region as a function of a "global" index of commuting with the central city and the local degree of poverty. The NYMA is a single entity in which the overall patterns of pathology are driven by what happens in the upper right section of the graph: the counties of New York City. Note that low birthweight rates were also regionalized in the same manner, but not parallel to these graphs

Figure 11.9, from [11], displays the diffusion of the "slow plague" of AIDS from the New York Metropolitan Region down along the US urban hierarchy, very strongly dominated by New York. Here, the probability of contact was not homogenized using the Markov method as at the metro regional level, but measured simply by US Census data on the 1990 moving pattern from New York to the other metro

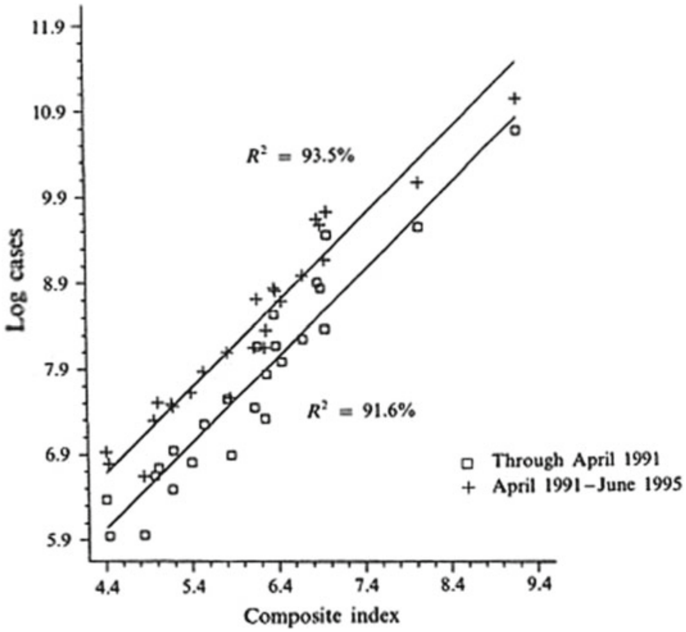


Fig. 11.9 From Wallace et al. [11]. Multivariate analysis of covariance for log AIDS cases, through April 1991, and for the period April 1991–June 1995. The composite index is X = 0.764 log(USVC91) + 0.827 log(USME87/USME72) + 0.299 log(P[l;]). P[1;] is a probability-of-contact index with the New York Metro region constructed from US Census data on migration from New York. USVC91 is the rate of violent crimes per 100,000 for 1991. USMExy is the absolute value of the number of manufacturing jobs in a metro region for the year xy. The two lines are highly parallel but have significantly different intercepts. This suggests the underlying matrix of regression slopes defined across the national urban hierarchy, is indeed fixed on the timescale of applied perturbations. See [11] for details

regions. Susceptibility at the local level was determined by "boomtown-busttown" employment dynamics and the local burden of violent crime.

As New York's AIDS level rose, in no small part as a result of the social disintegration associated with conditions indexed by Figs. 11.4, 11.5 and 11.7, infection avalanched down the US urban hierarchy. One public official commented that the processes of displacement associated with fire service failure "shotgunned AIDS over the Bronx". Similar "shotguns" affected large sections of Central Harlem and the Lower East Side in Manhattan, a vast region across African–American neighborhoods of Northern Brooklyn, and the African–American section of South Jamaica, Queens.

In the waning days of the Carter Administration, an FBI agent assigned to the Civil Rights Division of the Dept. of Justice, and with whom the author and colleagues met to discuss these matters, felt compelled to assert that, by using the Rand fire models, the city was actively attempting to "break up voting blocs".

In that sense, then, the Rand fire models have been a stunning success.

As someone famously put it, “Another such victory and we are undone”...

The Rand Fire Models are still very much in active use by FDNY. In fact, NYC’s Bloomberg Administration attempted to justify the closing of some 20 further fire units using them as late as 2011. Under US law, the Rand Fire Models provide a shield against discrimination lawsuits, as such suits must prove an intent to discriminate rather than an effect of discrimination.

Since NYC’s depopulation after the 1970s, about one million new residents have flowed back into the city, raising the area densities of badly overcrowded housing units above threshold for the outbreak of another round of “South Bronx” contagious fire and abandonment. Curiously, this dynamic has affected even newly White middle-class neighborhoods like Central Harlem. In spite of the growing risk, FDNY’s continuing commitment to the Rand Fire Models—very precisely “doctrinal groupthink” as studied in Chap. 2—prevents reopening of badly needed fire companies. This in the face of both the 9/11 attacks and the numerous conflagrations that accompanied Hurricane Sandy.

The ecological machinery for another catastrophe has been fully rewound.

References

1. Wallace, D., and R. Wallace. 1998. *A plague on your houses*. New York: Verso.
2. The Rand Fire Project. 1979. *Fire deployment analysis: A public policy analysis case study*. New York: North Holland.
3. Wallace, R., and D. Wallace. 1977. *Studies on the Collapse of Fire Service in New York City 1972–1976: The impact of pseudoscience in public policy*. Washington, DC: University Press of America.
4. Wallace, R. 2018. *Canonical Instabilities of Autonomous Vehicle Systems: The unsettling reality behind the dreams of greed*. New York: Springer.
5. Geroliminis, N., and J. Sun. 2011. Hysteresis phenomena of a macroscopic fundamental diagram in freeway networks. *Transportation Research B* 45: 966979.
6. Blandin, S., et al. 2011. A general phase transition model for vehicular traffic. *SIAM Journal of Applied Mathematics* 71: 107127.
7. Sugiyama, Y., M. Fukui, M. Kikuchi, K. Hasebe, A. Nakayama, K. Nishinari, S. Tadaki, and S. Yukawa. 2008. Traffic jams without bottlenecks—experimental evidence for the physical mechanism of the formation of a jam. *New Journal of Physics* 10: 033001.
8. Maerivoet, S., and B. De Moor. 2006. Data quality travel time estimation and reliability. Katholieke University Leuven 06030.
9. Corless, R., G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth. 1996. On the Lambert W-function. *Advances in Computational Mathematics* 4: 329359.
10. Monkonnen, E. 2001. *Murder in New York City*. Berkeley: University of California Press.
11. Wallace, R., D. Wallace, J.E. Ullmann, and H. Andrews. 1999. Deindustrialization, inner-city decay, and the hierarchical diffusion of AIDS in the USA: how neoliberal and cold war policies magnified the ecological niche for emerging infections and created a national security crisis. *Environment and Planning A* 31: 113–139.

Chapter 12
An Emerging Catastrophe: The Weaponization of Emotional Sentience in AI

12.1 A Brief Preface

Here, we reassemble bits and pieces from across the previous essays to conduct a policy analysis of, and make a recommendation regarding, the Next Big Thing in artificial intelligence, the forthcoming creation of "emotionally sentient" systems that can understand—and manipulate—human emotional responses at and across various timescales. Although civilian uses of the technology will clearly be aimed at marketing something—goods, services, compliance with the desires and intents of a purported authority, and so on—there is considerable potential for "dual use" of such machineries on Clausewitz landscapes of conflict and contention. Indeed, emotionally sentient agents would appear to fit well programs of "informationized operations" in the East Asian mode, again over various timescales. Although current PRC military thinking incorrectly believes modern information systems have cleared away the fog-of-war and made the battlefield transparent, their focus on directing the flow-of-events toward long-term strategic advantage could be significantly enabled by weaponized AI emotional sentience.

Conflict is a dauntingly complex gestalt phenomenon, taking place across scales and levels of organizations of time, space, social structure, culture, and material resources. Emotionally sentient AI will essentially be machine psychopaths, distinctly characteristic cultural artifacts having their own deeply obscure emotional analogs, but able to both understand and manipulate human emotions. Such machines will undoubtedly make a significant contribution to open warfare, clandestine conflict, and the dirty ongoing business of suppressing civilian populations.

This is indeed a multifactorial and protean technology. The cascading outfalls surrounding social media provide a salient comparison.

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_12

12.2 Summary

We outline a theory of culturally channeled emotional dynamics across individual human, institutional, machine, and composite entities, based on extension of current "anytime algorithm" analyses of machine cognition under time limitation. Emotions are for speed. Emotions are for survival. Emotions in humans, their social structures, and man/machine composites, are cultural artifacts. In consequence, culturally characteristic "emotional dysfunction" leading to inability to fulfill expected norms can always be imposed on a complex human cognitive system via a sufficiently crafted pattern of "noise" or time constraint, much as was explored by the US military strategist John Boyd. More sophisticated methodology might be aimed, in a characteristic "East Asian" modality, at long-term channeling of emotional responses. In sum, it should be possible—even straightforward—to construct emotionally sentient agents to either subtly channel behavior or to impose explicit patterns of dysfunction on an adversary. The development of emotionally sentient artificial intelligence is, then, very much a dual-use technology that should be subject to considerable oversight.

12.3 Introduction

The Next Big Thing in AI, it appears, will be "emotionally sentient agents". Microsoft's McDuff and Czerwinski [1], for example, put the matter thus:

> The field of affective computing concerns the design and development of computer systems that sense, interpret, adapt, and potentially respond appropriately to human emotions... Systems that respond to social and emotional cues can be more engaging and trusted, enabling computers to perform complex tasks in a more socially acceptable manner.

Indeed, as early as 1988 Ortony, Clore and Collins [2] developed a computational emotion model that, as Bartneck et al. [3] comment,

> "...[H]as established itself as the standard model for emotion synthesis". [The model] specifies 22 emotion categories based on valenced reactions to situations constructed either as being goal relevant events... or as attractive or unattractive objects... It contains a sufficient level of complexity and detail to cover most situations an emotional interface character might have to deal with.

Bartneck et al. [3] open their examination of the relationship between emotion models and artificial intelligence with Minsky's [4] famous statement:

> The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions.

Bartneck et al. claim that this question is still open. Here, we will show that the question was, at least for cognitive real-time critical systems, fully answered decades ago: Such machines cannot be intelligent without emotion. Their emotions, however, are nothing like those of humans-in-culture, although, as in the McDuff and Czerwinski exploration of The Next Big Thing, their intelligence may come to understand—and manipulate—ours.

But emotion—human, institutional, machine, cockpit, or complex composite—will always remain a cultural artifact. And this has certain implications for adversarial interaction.

Sufficient cultural understanding should permit the development of emotionally sentient AI systems that can impose catastrophic emotional dysregulation or, over a longer scale, directive channeling, of an opponent on "Clausewitz landscapes" characterized by uncertainty, imprecision, and delay. This idea can perhaps be viewed as a natural extension of the "OODA loop" perspective of the US military tactician and strategist John Boyd to the realm of AI (e.g., [5]).

Our analysis will suggest that emotional sentience in artificial intelligence is a dual-use technology, and should be regulated as such.

12.4 Revisiting an Evolutionary Perspective on Emotion

Emotional responses in animals are widely understood to represent learned or inherited patterns of very rapid response to particular sets of environmental or sensory cues [6]. The evolutionary advantage of rapid patterned responses is obvious, and applies even to what might be considered "nonminded" organisms.

For humans, or at least those tested in an American university setting [7], Cowen and Keltner [8] identify as many as 27 distinct categories of emotion, bridged by continuous gradients. Even the simplest categorizations identify at least two affective dimensions, "valence" and "arousal" [8]. This dimensional multiplicity will prove to be a critical observation.

For humans, matters of emotion are invariably intertwined with culture and both individual and community historical trajectory ([9, 10] Chap. 3).

Human institutions, like insect colonies, are cognitive and can learn, but are likewise embedded in a cultural milieu, and their patterns of rapid, intuitive, response to internal ruminations or external information also determine their ability to respond to Lamarckian selection pressures and persist.

Cognitive machines must also sometimes act on incomplete information, via "anytime algorithms" (AA) that provide "good enough" answers under time constraints (e.g., [11]). What is perhaps less obvious is that machines, or man/institution/machine "cockpit" enterprises, also operate within a cultural milieu, and are not independent of the sociocultural "riverbanks" confining a de-facto "stream of consciousness".

Machine intelligence is just another cultural artifact, and the design and use of such things—much as woven baskets—is always culturally embedded.

Here, we will take an inverse perspective on human emotion, using an approach from machine cognition—the AA—to examine necessary constraints on rapid patterned response mechanisms in institutions, individual humans, and man/institution/machine composites.

12.5 Anytime Algorithms Redux

We summarize material from Wallace [12].

Zilberstein [11] describes the basic idea in these terms:

> ...[Anytime algorithms] give intelligent systems the capability to trade deliberation time for quality of results. This capability is essential for successful operation in domains such as signal interpretation, real-time diagnosis and repair, and mobile robot control. What characterizes these domains is that it is not feasible (computationally) or desirable (economically) to compute the optimal answer...

Following Zilberstein [11], the "quality" of a "traveling salesman" problem calculation approaches an asymptotic limit with time, as in Fig. 12.1, albeit with some considerable variance.

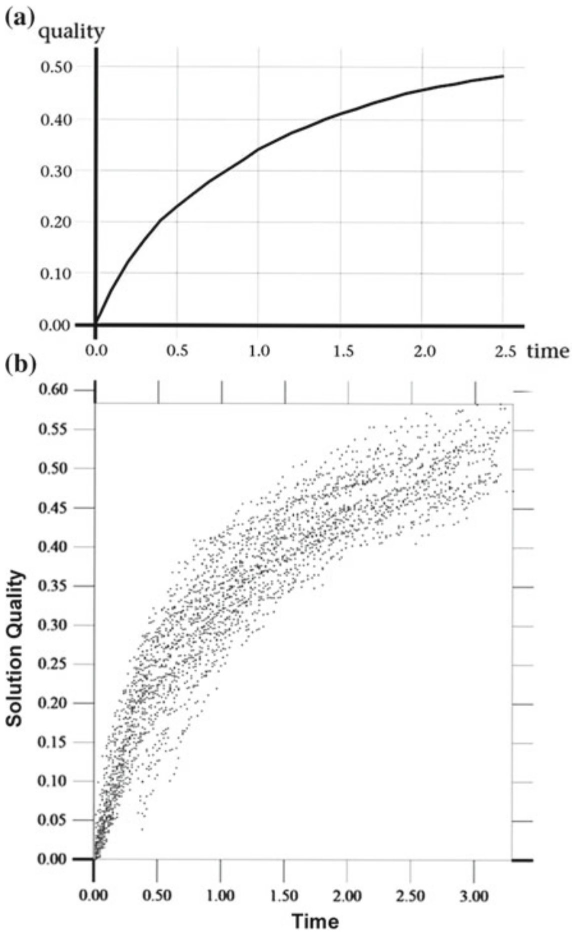


Fig. 12.1 From Zilberstein [11]. **a** Generic form of a quality-of-calculation result as a function of time under a generalized anytime algorithm. **b** Empirical realizations of quality measures at time t of a randomized tour-improvement algorithm for the traveling salesman problem over 50 cities. When does the variance explode?

It is important to realize that, in general, anytime algorithms are not necessarily stable, and sufficient conditions for stability must be established on a case-by-case basis, sometimes using draconian stochastic Lyapunov function methods that are far from transparent but are essentially similar to recent results regarding the stabilization/destabilization of stochastic differential equations (e.g., [13, 14]).

As Greco et al. [15] put it,

> With respect to most anytime algorithms... the fact that anytime controllers interact in feedback with dynamic systems introduces severe difficulties in their synthesis and in the analysis of the resulting closed-loop performance. Indeed, the stochastic switching system ensuing from executing different controllers at different times is even prone to instability.

Most cognitive processes or algorithms will eventually "top out" in quality according to some asymptotic function of time. Taking an exponential model gives a quality measure Q_j as

$$Q_j = \frac{\beta_j}{\alpha_j}(1 - \exp[-\alpha_j T_j]) \to \frac{\beta_j}{\alpha_j} \tag{12.1}$$

We assume, for simplicity, that a number of cognitive submodules within an intelligent agent must compete in sequence for time intervals T_j to complete their tasks under an overall time constraint $T_j < T$

An expression for a standard Lagrangian optimization can then be written as

$$\begin{aligned} L = \sum_j \frac{\beta_j}{\alpha_j}(1 - \exp[-\alpha_j T_j]) + \\ \lambda(T - \sum_j T_j) \\ \partial L/\partial T_j = \beta_j \exp[-\alpha_j T_j] \\ = \lambda \end{aligned} \tag{12.2}$$

λ is often viewed as an irrelevant "undetermined multiplier" used to make the optimization converge, if only via a somewhat elaborate Kuhn–Tucker "reverse-water filling" mechanism [16]. λ, however, is interpreted in economics, not as simply an arbitrarily adjustable "fitting" parameter, or set of them, but as the "shadow price" imposed by the constraint conditions, and this is how we see it here. As Jin et al. [17] show, optimization can fail: the shadow price can be imposed by an adversary as an "environmental demand" that the "firm" simply cannot meet.

In first order and without noise, the constraint on the shadow price becomes $0 \leq \lambda \leq \beta_j$. If λ is beyond this range, the system must fail.

For a chain-of-command structure, in which the failure of one level causes the failure of all, the more natural Lagrangian is $\log(\Pi_j Q_j) = \sum_j \log(Q_j)$, giving

$$\partial L/\partial T_j = \frac{\alpha_j \exp(-\alpha_j T_j)}{1 - \exp(-\alpha_j T_j)} = \lambda$$

Again, sufficient shadow price, as imposed by an adversary, hammers needed response times below what can be achieved in real-world circumstances.

The effects of “noise”, however, must be separately added according to some appropriate scheme, here taken as an Ito stochastic differential equation (SDE) [18].

The fact that

$$dQ/dt = \beta \exp[-\alpha t] = \beta - \alpha Q \tag{12.3}$$

leads to an SDE

$$dQ_t = (\beta - \alpha Q_t)dt + \sigma Q_t dW_t \tag{12.4}$$

where the second term represents the effects of volatility under Brownian noise, given as dW_t.

Here, the expectation $E(Q) \to \beta/\alpha$. However, applying the Ito Chain Rule to Q^2 [18] permits explicit calculation of the nonequilibrium steady state variance as $E(Q^2) - E(Q)^2$, giving

$$Var(Q) \to \left(\frac{\beta}{\alpha - \sigma^2/2}\right)^2 - \left(\frac{\beta}{\alpha}\right)^2 \tag{12.5}$$

This expression explodes as $\sigma^2/2 \to \alpha$, leading to extinction.

That is, for this model of a one-dimensional “anytime algorithm”, sufficient noise will be catastrophic.

A more complete analysis involves full-scale stochastic optimization methods that are beyond the scope of this analysis. Nonetheless, the fact that the observed dimensionality of human emotional structure is $n \geq 2$ is sufficient to explore something of the considerable complexity of emotional dysfunction in humans, their institutions, their cognitive machines, and the many various composite entities.

Putting aside, for the moment, the question of time constraints, we suppose that the different cognitive submodules interact according to a deterministic multidimensional matrix relation that is perturbed by noise, leading to a stochastic differential equation like Eq. (12.4) having the form

$$dQ = f(Q)dt + g(Q)dW_t \tag{12.6}$$

where W_t is multidimensional Brownian noise and all quantities are of dimension $n \geq 2$, following observations of the high dimensionality of human emotional response. The multidimensional function $f(Q)$ is assumed to have a set of nonequilibrium steady states as $t \to \infty$. These may, however, be stable or unstable, depending on the context.

In particular, $f(Q)$ includes patterns of crosstalk between modules, and $g(Q)$ quantifies the influence of noise within and across modules, analogous to the role of σQ in Eq. (12.4).

Appleby et al. [19] show that, if $n \geq 2$, under very broad conditions, for any $f(Q)$ that satisfies them, a function $g(Q)$ can be found that either stabilizes an unstable nonequilibrium steady state of f or, perhaps most importantly, destabilizes any stable state. A fully worked-out two-dimensional example can be found in Sect. 2.2 of Wallace et al. [20]. There, additive noise simply fuzzes out the deterministic solution, while cross-noise destabilizes a stable nonequilibrium steady state.

And thereupon will hang something of our tale.

But there is more.

12.6 Generalized Anytime Algorithms I

Much depends on the rate at which complex cognitive processes must proceed under real-world, real-time selection pressures. Emotional responses are predicated on quick patterned response to sensory cues. We will propose here that there can be several levels of such patterning in complex cognitive systems, depending on the rate at which a response is required. We look at cognitive system dynamics as driven by the inverse rate τ at which an answer is required and must be acted on in the real world. High values of τ imply relatively slow response rates, which we will take as corresponding to a high "temperature" in a symmetry-breaking model of cognitive process that invokes groupoids rather than groups.

Cognition involves an agent's choice of an action, or a relatively small number of them, from the complete repertoire of those available [21]. This permits the identification of two "dual" information sources representing a pair of contending agents. That is, cognitive choice implies a reduction of uncertainty which, in turn, implies the existence of an underlying information source characterizing that cognition. Wallace [10] provides further details of the argument.

Our central interest is the dynamics of the interacting network of cognitive submodules within a single agent. The agent may be an individual person, an institution, an intelligent machine, a cockpit, or more structured composite. The essential point is the need to act under time constraint defined (inversely) by τ.

The submodules will extend across both scale and level of organization.

We thus must study matters across an extensive network of entities, represented by a considerable set of cognitive information sources within a single, larger agent, say $\mathscr{Y} = \{Y_1, Y_2, ...\}$.

We consider, then, a dense net of cognitive submodules interacting through crosstalk and other forms of exchange. Such agents will both cooperate and compete for essential resources, including time.

Sequences of actions and behaviors of some length $N \to \infty$ can, in general, be divided into a small high probability "typical" set consonant with a "grammar" and "syntax", and a very large set of paths having vanishingly low probability that is not consistent with the underlying grammar and syntax [16, 22].

However, cognitive dynamics may require invocation of "nonergodic" information sources, extending the arguments presented in Wallace ([23], Chap. 3) and [24]. "Nonergodic" means that long-term averages do not converge on cross-sectional expectations, roughly analogous to the idea of "nonparametric" statistical models that are not based on the Central Limit Theorem. The source uncertainty of such systems then cannot be described in terms of the classic Shannon "entropy" [22], but regularities arising from Feynman's [25] identification of information as a form of free energy still permit considerable simplification of system dynamics via familiar "phase transition" formalism, based, however, on groupoid rather than group symmetry shifts.

Following Wallace [23], the embedding natural or human ecosystem upon and within which the entity of interest must act—the underlying landscape—can be characterized by its own "grammar" and "syntax" defining what are high and low probability sequences of events, permitting representation as an information source, say X.

An additional information source within which the entity of interest is embedded is the culture/historical trajectory within which it is an artifact/offspring, say Z. Culture-of-origin will have its own grammar and syntax.

Yet another information source arises from the dynamics of the "large deviations" to which cognitive agencies or agents are subject. As Champagnat et al. [26] describe, however, sudden transitions between nonequilibrium steady states also involve high probability paths whose description is in terms of the familiar Shannon uncertainty expression $\mathscr{I} = -\sum_k P_k \log(P_k)$, where the P_k form a probability distribution. Variations of this result appears repeatedly as Sanov's Theorem, Cramer's Theorem, the Gartner–Ellis Theorem, the Shannon–McMillan Theorem, and so on [27]. This result allows the characterization of large deviations in terms of yet another information source, L_D.

Study now focuses on a very large, complex network of (not necessarily ergodic) information sources linked by crosstalk that we write here in terms of the uncertainty of a single joint information source [16] as

$$H(X, \mathscr{Y}, Z, L_D)$$

This (highly networked) information source uncertainty is defined on a set of jointly typical [16] high probability paths z defined by a sequence of high dimensional states that conform to the "grammars" and "syntaxes" of the interacting information sources X, the set $\mathscr{Y}$, the embedding culture Z, and L_D.

The argument then becomes concerned with the set of "jointly typical" high probability paths $z^n \equiv \{z_0, z_1, ..., z_n\} \to z$.

Following Khinchin [22], it is possible to define a path-dependent source (joint) uncertainty $H(z^n) \to H(z)$ that can vary across the manifold defined by the full set of paths z. $H(z)$, however, no longer has an "entropy like" mathematical expression, but its dynamics can be inferred using a familiar statistical argument.

The next step is to invoke a Boltzmann pseudoprobability as

$$P[H(z_q)] \equiv \frac{\exp[-H(z_q)/\tau]}{\sum_z \exp[-H(z)/\tau]} \tag{12.7}$$

where the sum, or, as we will explore below, a generalized integral, of the denominator is over all possible jointly typical paths z, and τ is, again, the inverse rate at which decision must be reached and action taken.

It is important to recognize that the $H(z_q)$ are uncertainty values associated with an enormously large set of individual paths.

From the pseudoprobability of Eq. (12.7) a "free energy" analog $\mathscr{F}$ can be defined as a Morse function [28], using the denominator of Eq. (12.7) and the relation

$$\exp[-\mathscr{F}/\tau] \equiv \sum_{z} \exp[-H(z)/\tau] \tag{12.8}$$

Given τ, the inverse rate at which a decision must be reached and acted on, a surprisingly subtle phase transition model emerges on the strategic scale by identifying equivalence classes of a system's developmental pathways z, say, "functional" and "doomed". This allows definition of a symmetry groupoid for the cognition/action process [29]. A groupoid is a generalization of an algebraic group in which a product is not necessarily defined between each element. The simplest example is, perhaps, a disjoint union of separate groups, but sets of equivalence classes also define a groupoid. See Weinstein [29] for details.

Given an information source associated with the system of interest, a full equivalence class algebra can be constructed by choosing different system origin states and defining the equivalence of subsequent states at a later time by the existence of a high probability path connecting them to the same origin state. Disjoint partition by equivalence class, analogous to orbit equivalence classes in dynamical systems, defines a symmetry groupoid associated with the cognitive process.

The equivalence classes across possible origin states define a set of information sources dual to different states available to the systems of interest. These create a large groupoid, with each orbit corresponding to an elementary "transitive" groupoid whose disjoint union is the full groupoid. Each subgroupoid is associated with its own dual information source, and larger groupoids must have richer dual information sources than smaller.

The "free energy" Morse Function of Eq. (12.8) is liable to an analog of Landau's classical spontaneous symmetry breaking [28]. Under symmetry breaking, higher "temperatures" τ are associated with more symmetric higher energy states in physical systems.

Under time constraint, decline in the quantity τ from Eq. (12.8) will cause sharply punctuated collapse from higher to lower symmetry states, triggering ultimate system failure, modulated by both the embedding environment represented by the information source X and the culture/historical trajectory represented by Z.

Again, the "free energy" of Eq. (12.8) can be driven into a sudden, highly punctuated "phase transition" reflecting fundamental symmetry shifts by changes in τ, the inverse rate at which decision-and-action must take place. The symmetry shift, however, is between groupoids associated with a synergism across culture, contending/cooperating cognitive institutions, the embedding environment, and the structure of stochastic "large deviations". This is not simply the melting of ice.

In marked and fundamental contrast to the stable/unstable dichotomy associated with the previous treatment of anytime algorithms, phase change associated with $\mathscr{F}$ and τ may be relatively subtle, akin to the Eldredge and Gould [30] pattern of "punctuated equilibrium" in evolutionary transition. Under their model, evolutionary and coevolutionary dynamics undergo relatively long periods of apparent or near-stasis, where changes are small or difficult to see, followed by relatively sudden massive changes leading to fundamentally different coevolutionary configurations. Examples of contemporary debate on punctuated equilibrium, now a central matter in evolutionary biology, can be found in [30].

The dynamics of the free energy analog $\mathscr{F}$ at critical values of τ can be studied using variants of the standard iterative renormalization techniques as developed for cognitive processes in Wallace [31].

The most direct general inference, however, remains that phase transitions associated with changes in τ will be expressed in terms of changes in underlying groupoid symmetries, leading again to some version of the punctuated equilibrium dynamics identified by Eldredge and Gould [30]. The details will hold many devils.

An explicit simplified model, assuming an integral approximation and a Gamma distribution across the right hand side of Eq. (12.8), is given in the Chapter Mathematical Appendix.

Note, however, that the underlying approach can be significantly extended through a multidimensional iteration. Suppose there is a set of τ_j across a number of essential modules. These can be characterized as "quality" measures, replacing the Q_i of the previous section. Strategy then involves setting up circumstances so that, at a point of need, one's own time frames for action remain long compared to those of an opponent. Then the complexities of Eq. (12.6) can be imposed on an adversary.

12.7 Generalized Anytime Algorithms II

What happens when time is not the limiting factor, but time dynamics still count? We envision an institutional setting in which the quality of institutional cognition depends on, for example, the degree to which a full "sensory" picture of external conditions has been built up. Then Eq. (12.1) becomes

$$Q = \frac{\beta}{\alpha}(1 - \exp[-\alpha Z]) \tag{12.9}$$

where Z represents the quality of the sensory picture that itself approaches an asymptotic level as time unwinds:

$$Z = \frac{b}{a}(1 - \exp[-at]) \tag{12.10}$$

so that

$$Q = \frac{\alpha}{\beta}\left(1 - \mathrm{e}^{-\frac{\alpha a\left(1-\mathrm{e}^{-at}\right)}{b}}\right) \tag{12.11}$$

Then $dQ/dt = dQ/dZ \times dZ/dt$, leading to the stochastic differential equation

$$dQ_t = [dQ_t/dZ \times dZ_t/dt]dt + \sigma Q_t dW_t \tag{12.12}$$

where, again dW_t is taken as Brownian noise.

Some manipulation finds

$$\begin{aligned} dQ/dZ &= (\beta - \alpha Q) \\ dZ/dt &= (b - aZ) \\ Z &= \frac{-1}{\alpha}\log[1 - \frac{\alpha}{\beta}Q] \end{aligned} \tag{12.13}$$

Equation (12.12) then becomes

$$\begin{aligned} dQ_t = (\beta - \alpha Q_t)b\exp[-at]dt + \sigma Q_t dW_t = \\ (\beta - \alpha Q_t)(b - aZ_t)dt + \sigma Q_t dW_t = \\ (\beta - \alpha Q_t)(b + \frac{a}{\alpha}\log[1 - \frac{\alpha}{\beta}Q_t])dt + \sigma Q_t dW_t \end{aligned} \tag{12.14}$$

Using the Ito chain rule on Q_t^2 for the third form of Eq. (12.14), the limiting value of the variance $Var(Q) = E(Q^2) - E(Q)^2$ is then

$$\begin{aligned} Var \to -\frac{\beta^2}{\alpha^2}\left(\mathrm{e}^{-\frac{b\alpha}{a}} - 1\right)^2 + 1/4\,\frac{\beta^2}{a^2\alpha^2} \\ \times\left(2\,a\,\mathrm{W}\left(-1/2\,\frac{\sigma^2}{a}\mathrm{e}^{1/2\,\frac{2\,b\,\alpha-\sigma^2}{a}}\right) + \sigma^2\right)^2 \times \\ \left(\mathrm{W}\left(-1/2\,\frac{\sigma^2}{a}\mathrm{e}^{1/2\,\frac{2\,b\,\alpha-\sigma^2}{a}}\right)\right)^{-2} \end{aligned} \tag{12.15}$$

$W(X)$ is the Lambert W-function that is real only for $-\exp[-1] < X < \infty$. Thus stability is more subtle than the result of Eq. (12.5), implying the possibility of real and complex phases in the iterated system, depending on a complicated manner on the relative values of $b\alpha/a$ and $\sigma^2/2a$. Figure 12.2 shows three sets of twenty simulations of the first form of Eq. (12.14), using the ItoProcess function in the Financial package of the computer algebra program Maple 2019. For that calculation, $a = b = \alpha = \beta = 1$, and $\sigma = 0.01, 0.1, 0.56$. Above the largest value for σ, the simulation blows up, as expected from the appearance of the Lambert W-function.

Taking matters one step further, we can envision the multicomponent convoluted system of Eqs. (12.9) and (12.10) leading to Eq. (12.11) as not so much limited by

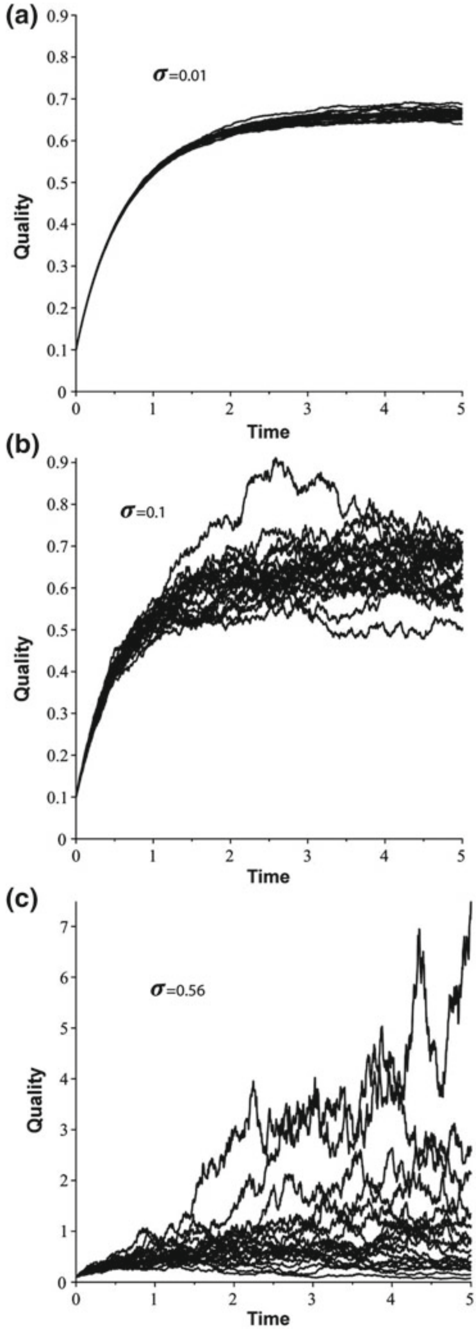


Fig. 12.2 Twenty simulations of the first form of Eq. (12.14), using the ItoProcess function of the computer algebra program Maple 2019, taking $a = b = \alpha = \beta = 1$, $\sigma = 0.01, 0.1, 0.56$. Small values of σ converge on the expected mean value $1 - \exp[-1] \approx 0.632$. Above $\sigma \approx 0.56$ the simulation fails, consistent with appearance of the Lambert W-function in the exact calculation

total availability of resources, but limited by time. That is, every subcomponent still has the resource limit $Z_j \to b_j/a_j$, but the total sequential response time allowed to the whole is constrained and must be parceled out in an optimized manner, i.e., $T = \sum_j T_j$, $0 < T_j < T$, *in the context of limitations on the rates at which essential resources can be delivered*, determined by the individual rate constants a_j. This circumstance is perhaps analogous to U.S. experiences in Vietnam, Iraq, Afghanistan, and so on. That is, even having plenty of essential resources, one is still nibbled to death by ducks, as it were.

Then it becomes possible to conduct a standard optimization in the T_j like Eq. (12.2), but for a Lagrangian ultimately defined in terms of the *product* of the Q_i, since the chain of command can fail at any single link. Finessing the argument, an appropriate Lagrangian can be constructed from the log of the product of the composite $Q_j(T_j)$, leading to another sum of terms expression for L and the shadow price relation

$$\partial L/\partial T_j = \frac{\alpha_j \, a_j{}^2 e^{-a_j T_j}}{b_j} e^{-\frac{\alpha_j a_j \left(1-e^{-a_j T_j}\right)}{b_j}} \left(1 - e^{-\frac{\alpha_j a_j \left(1-e^{-a_j T_j}\right)}{b_j}}\right)^{-1} = \lambda \qquad (12.16)$$

as shown in Fig. 12.3. If λ is sufficiently large, T_j is driven toward zero, and the composite structure must fail.

High enough shadow price λ—as imposed by an adversary—then forces the system to impossibly small response times resulting in collapse. This is again in line with the OODA loop arguments, but under circumstances not only constrained by materiel or information, in the sense that that $Z_j \to b_j/a_j$, but also by the rates a_j at which they can be delivered, making the T_j even more critical.

The approach could be extended to a stochastic optimization.

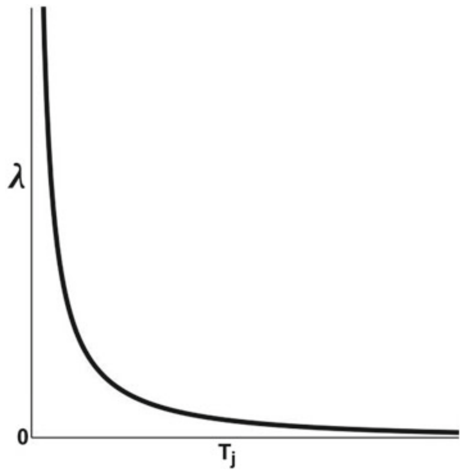


Fig. 12.3 Shadow price relation for the sequential composite chain of command of Eq. (12.16). Sufficiently great λ drives the needed response time to impossibly small values

12.8 Discussion

We present a theory of emotional dynamics in the context of culture across individuals, institutions, machine intelligence, and their many composites that extends the well-established technology of “anytime algorithms” in a highly natural manner.

Emotions are for speed and survival, and, in human systems, are always and inevitably cultural artifacts, thereby having particularities that can be exploited. That is, according to the models given here, “emotional dysfunction” leading to inability to fulfill expected norms can always be imposed on a multidimensional cognitive human system via a sufficiently crafted pattern of time limitation, “shadow price” and “noise”, within the confines of both an embedding environment and a cultural milieu. The perspective represents both a confirmation and a generalization of John Boyd’s OODA loop insights [5], i.e., that, at least on the tactical level, it is always possible to “get inside the decision loop” and crash out an adversary’s command and control structures.

On the longer “East Asian” timescales of extended conflict described by Jullien [32], it should be possible to subtly channel behavior of an adversary via broadly “emotional” manipulations, along with the imposition of more direct constraints on an enemy’s possible actions.

According to our models, consequent on the inevitable “cultural back door”, there will almost always be modes of AI emotional sentience that can impose both short-term catastrophic failure and longer-term channeling on a necessarily culturally constrained adversary’s cognitive structures at, and across, various scales and levels of organization. Sometimes, of course, the requirements for real-world implementation of such strategies cannot be fulfilled, a circumstance that may indeed have been carefully arranged by the opponent.

The limit on this conclusion is the same as the objection the theoretical ecologist Pielou [33] raises to models in biology: models can propose questions for empirical and observational study. Mathematical models of complex biological and social phenomena do not answer questions.

This being said, the models developed here suggest that emotional dysfunction in humans is necessarily ubiquitous, given the inevitable “noise” of life, both individually and in community. Times of particular stress, in a particular cultural context, will trigger specific patterns of individual social dysfunction that can be interpreted as “emotional” disorders. What is surprising is the possibility of extending the perspective across institutional, machine, and composite structures.

Similar arguments can be made for the “East Asian” channeling of behaviors on longer timescales [32].

Just as, in the words of the evolutionary anthropologist Robert Boyd, “Culture is as much a part of human biology as the enamel on our teeth”, our institutions, intelligent machines, and their many cockpit and other composites, are themselves very particular cultural artifacts that may be subject to culturally characteristic gestalt patterns of behavior and “emotional” pathology.

It is reasonable to expect that weaponized emotionally sentient AI will fully pursue such inherent culturally specific vulnerabilities, making this a dual-use technology that should be subject to stringent oversight.

A more complete formal analysis, using stochastic optimization tools, remains to be done, as does the iteration suggested at the end of the previous section.

12.9 Mathematical Appendix

If the denominator in Eq. (12.7) can be treated with an integral approximation, then the analysis can be simplified.

The essential idea is that, for a set of simultaneous complex situations that must be addressed at a given time, or, in an "ergodic" sense, for a single situation that changes over time, the value of the H is itself subject to stochastic variation. Fairly generally, then, we assign a probability distribution for values of H as a Gamma distribution with parameters k, θ having a density function

$$P(H, k, \theta) = \frac{H^{k-1}\exp[-H/\theta]}{\theta^k \Gamma(k)} \tag{12.17}$$

with $\theta, k > 0$ real.

Then the average of H across the distribution is $\hat{H} = k\theta$ and we redefine a "free energy" F as

$$\begin{aligned}\int_0^\infty \exp[-H/G(\tau)]P(H, k, \theta)dH = \\ \exp[-F/G(\tau)] = \\ \left(\frac{G(\tau)k}{G(\tau)k + \hat{H}}\right)^k\end{aligned} \tag{12.18}$$

An entropy-analog, in the sense of the Onsager treatment of nonequilibrium thermodynamics [34] is defined as

$$S(\tau) \equiv F(\tau) - \tau dF(\tau)/d\tau \tag{12.19}$$

leading to the Onsager nonequilibrium steady state condition

$$d\tau/dt \propto dS/d\tau = 0 \tag{12.20}$$

The resulting differential equation in $G(\tau)$ can be solved to give the relation

$$Z \equiv -C_2\tau + C_1$$
$$G(Z) = \frac{Z}{W(-1, (kZ/\hat{H})\exp[kZ/\hat{H}]) - kZ/\hat{H}} \tag{12.21}$$

C_1 and C_2 are positive constants and $W(-1, x)$ is the Lambert W-function of order -1, real and finite only in the range

$$-\exp[-1] < (kZ/\hat{H})\exp[kZ/\hat{H}] < 0 \tag{12.22}$$

Thus, $G(\tau)$ can be real and finite only over a limited range of positive τ-values, implying at least three different phases for the underlying system.

Extending Eq. (12.20) to a more general expression in terms of τ, for example,

$$d\tau/dt \approx dS/d\tau = q(\tau) \tag{12.23}$$

gives a relation for $G(\tau)$ that can be explicitly solved in terms of q and the Lambert W-function of appropriate order. This can be done, for example, by the computer algebra program Maple 2019. The resulting expression overwhelms the LaTex typesetting compiler and is omitted.

References

1. McDuff, D., and M. Czerwinski. 2018. Designing emotionally sentient agents. *Communications of the ACM* 61 (12): 74–83.
2. Ortony, A., G. Clore, and A. Collins. 1988. *The cognitive structure of emotions*. New York: Cambridge University Press.
3. Bartneck, C., M. Lyons, and M. Saerbeck. 2017. The relationship between emotion models and artificial intelligence, https://arXiv.org/pdf/1706.09554
4. Minsky, M. 1986. *The society of mind*. New York: Simon and Schuster.
5. Osinga, F. 2007. *Science, strategy and war: The strategic theory of John Boyd*. London: Routledge.
6. Darwin, C. 1872. *The expression of the emotions in man and animals*. London.
7. Henrich, J., S. Heine, and A. Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33: 61–135.
8. Cowen, A., and D. Keltner. 2017. Self-report captures 27 distinct categories of emotion bridged by continuous gradients, PNAS. www.pnas.org/cgi/doi/10.1073/pnas.1702247114.
9. Kleinman, A. 1991. *Rethinking Psychiatry: From cultural category to personal experience*. New York: Free Press.
10. Wallace, R. 2017. *Computational Psychiatry: A systems biology approach to the epigenetics of mental disorders*. New York: Springer.
11. Zilberstein, S. 1996. Using anytime algorithms in intelligent systems. *AI Magazine* 17: 73–83.
12. Wallace, R. 2019. Cognition, conflict, and doctrine: how groupthink fails on a Clausewitz landscape. In press, *Journal of Defense Modeling and Simulation*.
13. Dang, T., K. Ling, and D. Quevedo. 2018. Stability analysis of event-triggered anytime control with multiple control laws. arXiv:1804.07876v1 [math.OC] April 2018.

14. Quevedo, D., and V. Gupta. 2012. Sequence-based anytime control. *IEEE Transactions on Automatic Control* 58: 377–390.
15. Greco, L., D. Fontanelli, and A. Bicchi. 2007. Almost sure stability of anytime controllers via stochastic scheduling. In *46th IEEE conference on decision and control*, pp. 5640–5645.
16. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
17. Jin, H., Z. Hu, and X. Zhou. 2008. A convex stochastic optimization problem arising from portfolio selection. *Mathematical Finance* 18: 171–183.
18. Protter, P. 2005. *Stochastic integration and differential equations*, 2nd ed. New York: Springer.
19. Appleby, J., X. Mao, and A. Rodkina. 2008. Stabilization and destabilization of nonlinear differential equations by noise. *IEEE Transactions on Automatic Control* 53: 126–132.
20. Wallace, R., L.F. Chaves, L. Bergmann, C. Ayres, L. Hogerwerf, R. Kock, and R.G. Wallace. 2018. *Clear-Cutting Disease Control: Capital-led deforestation, public health austerity, and vector-borne infection*. New York: Springer.
21. Atlan, H., and I. Cohen. 1998. Immune information, self-organization, and meaning. *International Immunology* 10: 711–717.
22. Khinchin, A. 1957. *Mathematical foundations of information theory*. New York: Dover Publications.
23. Wallace, R. 2018. *Carl von Clausewitz, the Fog-of-War, and the AI Revolution: The real world is not a game of Go*. New York: Springer.
24. Wallace, R. 2018. New statistical models of nonergodic cognitive systems and their pathologies. *Journal of Theoretical Biology* 436: 72–78.
25. Feynman, R. 2000. *Lectures on computation*. New York: Westview Press.
26. Champagnat, N., R. Ferriere, and S. Meleard. 2006. Unifying evolutionary dynamics: From individual stochastic process to macroscopic models. *Theoretical Population Biology* 69: 297–321.
27. Dembo, A., and O. Zeitouni. 1998. *Large deviations and applications*, 2nd ed. New York: Springer.
28. Pettini, M. 2007. *Geometry and topology in hamiltonian dynamics and statistical mechanics*. New York: Springer.
29. Weinstein, A. 1996. Groupoids: Unifying internal and external symmetry. *Notices of the American Mathematical Association* 43: 744–752.
30. Lyne, J., and H. Howe. 2017. Punctuated equilibrium: Rhetorical dynamics of a scientific controversy, Chapter 6 in Harris, R. (ed.), *Landmark essays on Rhetoric of science: Case studies*, Second Edition, CRC Press, Boca Raton.
31. Wallace, R. 2005. *Consciousness: A mathematical treatment of the global neuronal workspace model*. New York: Springer.
32. Jullien, F. 2004. *A Treatise on Efficacy: Between Western and Chinese thinking*. Honolulu: University of Hawaii Press.
33. Pielou, E.C. 1977. *Mathematical ecology*. New York: Wiley.
34. de Groot, S., and P. Mazur. 1984. *Nonequilibrium thermodynamics*. New York: Dover.

Chapter 13
Final Remarks

Conflict between cognitive institutions on a Clausewitz landscape dominated by friction and the fog-of-war enters the same realm of difficulty as confronts and afflicts the study of consciousness in higher animals (e.g., [1–4], and references therein). Cognitive submodules of institutions and higher animals become linked, at different scales and levels of organization, in tunable, dynamically shifting coalitions in response to signals arising from both internal crosstalk and external "sensory" sources. The formal description of consciousness remains highly contentious and appears to require mathematical tools familiar to the far outlands of string theory. This is no exaggeration: the art and science of war is, perhaps, even more confounding than our most arcane scientific edifices.

That being said, well-trained, well-equipped forces are themselves highly cognitive and can often successfully confront tactical problems from an engineering viewpoint, skillfully using the available tools-of-the-trade. Watts [5] puts the contrast between tactical, operational, and strategic matters as follows:

> Tactical problems are "tame" in that they generally have definite solutions in an engineering sense. So-called "wicked" problems are fundamentally social ones. They are ill-structured, open-ended, and not amenable to closed, engineering solutions. Operational and strategic problems appear to lie within the realm of wicked or messy problems. Such problems give every indication of requiring hard thinking and careful analysis rather than quick, intuitive responses.

At the strategic level, affairs are particularly—but not entirely—dominated by selection pressures and Lamarckian evolutionary dynamics that are no less difficult than, but are significantly different from, phenomena of rapid-fire cognition, "cockpit consciousness", and control. Periods of apparent stasis can mask "upstream" activities leading to sudden punctuated change. Conflict becomes a question of a long-term "farming" of a developing situation via selection pressures, recognizably akin to the directed evolution that led to our current spectrum of domesticated plants

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_13

and animals. The formal analysis of evolutionary process is no less subtle or mathematically challenging than the description of consciousness (e.g., [6–8]).

Although we have not done so here, it seems possible to examine the operational level that links tactics to strategy using a "developmental" model similar to gene expression in which "signals", on an intermediate time scale between tactical and strategic, are imposed on a developing trajectory to guide its final expression (e.g., [9, 10]).

The quotation from Bracken [11] in the Preface describes something of the fundamental disjunction between theory and praxis characteristic of US (and other) national security studies. Clausewitz [12] expresses similar thoughts (Fig. 13.1):

> The first common error is an awkward and quite impermissible use of certain narrow systems as formal bodies of laws...
>
> A far more serious menace is the retinue of *jargon, technicalities, and metaphors* that attends these systems. They swarm everywhere – a lawless rabble of camp followers. Any critic who has not seen fit to adopt a system – either because he has not found one that he likes or because he has not yet got that far – will still apply an occasional scrap of one as if it were a ruler, to show the crookedness of a commander's course. Few of them can proceed without the occasional scraps of scientific military theory. The most insignificant of them – mere technical expressions and metaphors – are sometimes nothing more than ornamental flourishes of the critical narrative. But it is inevitable that all the terminology and technical expressions of a given system will lose what meaning they have, if any, once they are torn from their context and used as general axioms or nuggets of truth that are supposed to be more potent than a simple statement.
>
> Thus it has come about that our theoretical and critical literature, instead of giving plain, straightforward arguments in which the author at least always knows what he is saying and the reader what he is reading, is crammed with jargon, ending at obscure crossroads where the author loses his readers. Sometimes these books are even worse: they are just hollow shells. The author himself no longer knows just what he is thinking and soothes himself with obscure ideas which would not satisfy him if expressed in plain speech...

Contemporary examples include Lykke's [13] S = E + W + M "model" of strategy, and the many undisciplined "nonlinearity" metaphors (e.g., [14]). Indeed, this monograph itself is open to the Clausewitz criticism of overcomplication.

We reiterate the comment by Tse-Tung ([15], p. 195):

> Epistemologically speaking, the source of all erroneous views on war lies in idealist and mechanistic tendencies on the question of war. People with such tendencies have a subjective and one-sided approach to problems. They either indulge in groundless and purely subjective talk, or, basing themselves upon a single aspect or a temporary manifestation, magnify it with similar subjectivity into the whole of the problem... [O]nly by... taking an objective and all-sided view in making a study of war can we draw correct conclusions...

Sometimes—following Mao Tse-Tung—for complicated institutional dynamics akin to phase change or punctuated equilibrium evolutionary transitions, "plain speech" will not suffice and may disastrously mislead, hence our attempts here at a more comprehensive "all-sided view".

The brutal history of the twentieth century, the "second hundred year's war", and its amplification factors of "collateral damage" and recurrence, suggest that we

Fig. 13.1 The Doctrine. Photo: Rodrick Wallace

misunderstand and hence fail to contain institutional conflict on Clausewitz landscapes. What we do not comprehend, we cannot remedy. Although it is certainly true that "wars will cease when men refuse to fight", the entrainment of individuals into social structure, and the interpenetration of culture, society, and human self, make that outcome a daunting challenge.

Nonetheless, in the aftermath of Cold War-driven deindustrialization ([16], Chap. 7), Vietnam, the "war on terror", and the associated relentless avalanches of debilitating expenditures and other failures, and as Krepinevich and Watts [17, 18]—and many others – argue at length, the US national security team badly needs to up its game. In the spirit of Richardson, perhaps, one can hope this monograph contributes toward better understanding of the underlying constraints limiting the usefulness of current approaches to conflict between nations.

Regarding these constraints, the Armed Conflict Working Group of the International Work Group on Death, Dying, and Bereavement concludes [19]

> Armed conflict is no longer an acceptable way to resolve disputes. The destructive potential of modern weapons, the suffering they cause, the interdependence of states (globalization), massive civilian casualties, large disruptions of populations, and enormous cost of modern warfare support this view...
>
> Armed attacks affect all whom they involve, not only those whose lose people they love, but all those who witness or dread the losses that threaten them, the people whose safe world has become unsafe...
>
> In these circumstances, the powerful emotions that accompany fear and grief may feed into the cycles of violence that perpetuate the fear and grief...
>
> From time immemorial unscrupulous leaders, advisers, propagandists, lobbyists, and members of the media have taken advantage of these human responses to frighten, dominate, exploit, and profit...

The experience of the last hundred years suggests that routine resort to armed conflict has become obsolete. Large-scale Lamarckian evolutionary processes of institutional contention under conditions of uncertainty, imprecision and delay can, however, still be "farmed", if one or more of the contending agents is sufficiently clever and persistent.

Perhaps, in a reversal of the Clausewitz dictum, future political negotiation and skilled maneuver can come to be seen as warfare carried out by other means. One fervently hopes for diminished death, destruction, and recurrence.

References

1. Baars, B. 1988. *A cognitive theory of consciousness*. New York: Cambridge University Press.
2. Wallace, R. 2005. *Consciousness: A mathematical treatment of the neuronal global workspace model*. New York: Springer.
3. Wallace, R. 2012. Consciousness, crosstalk, and the mereological fallacy: An evolutionary perspective. *Physics of Life Reviews* 9: 426–453.
4. Wallace, R. 2017. *Computational Psychiatry: A systems biology approach to the epigenetics of mental disorder*. New York: Springer.
5. Watts, B. 2008. *US Combat training, operational art, and strategic competence: Problems and opportunities*. Washington, DC: Center for Strategic and Budgetary Assessments.
6. Wallace, R. 2010. Expanding the modern synthesis. *Comptes Rendus Biologies* 333: 701–709.
7. Wallace, R. 2011. A formal approach to evolution as a self-referential language. *BioSystems* 106: 36–44.
8. Wallace, R. 2014. A new formal perspective on 'Cambrian Explosions'. *Comptes Rendus Biologies* 337: 1–5.
9. Wallace, R., and D. Wallace. 2008. Code, context and epigenetic catalysis in gene expression. *Transactions on Computational Systems Biology XI, LNBI* 5750: 283–334.

10. Wallace, R., and D. Wallace. 2016. *Gene expression and its discontents: The social production of chronic disease*, 2nd ed. New York: Springer.
11. Bracken, P. 2006. *Net assessment, a practical guide, Parameters*. Springer, 90–100.
12. von Clausewitz, C. 1989. On War, Howard, M., P. Paret (eds.), Princeton University Press, Princeton, NJ.
13. Lykke, A. 1989. Defining military strategy. *Military Review* 2–8.
14. Beyerchen, A. 1992. Clausewitz, nonlinearity, and the unpredictability of war. *International Security* 17: 59–90.
15. Tse-Tung, M. 1963. *Selected military writings of Mao Tse-Tung*. PRC: Foreign Language Press.
16. Wallace, R. 2015. *An ecosystem approach to economic stabilization: Escaping the neoliberal wilderness*. New York: Routledge.
17. Krepinevich, A., and B. Watts. 2009. *Retaining strategic competence*. Washington, D.C.: Center for Strategic and Budgetary Assessment.
18. Krepinevich, A., and B. Watts. 2009. Lost at the NSC, The National Interest 99: 63–72.
19. Armed Conflict Working Group. 2013. Armed Conflict: A model for understanding and intervention. *Death Studies* 37: 61088.

Chapter 14
Mathematical Appendix

14.1 Groupoids

Following Weinstein [1] closely, a groupoid, G, is defined by a base set A upon which some mapping—a morphism—can be defined. Note that not all possible pairs of states (a_j, a_k) in the base set A can be connected by such a morphism. Those that can define the groupoid element, a morphism $g = (a_j, a_k)$ having the natural inverse $g^{-1} = (a_k, a_j)$. Given such a pairing, it is possible to define "natural" end-point maps $\alpha(g) = a_j$, $\beta(g) = a_k$ from the set of morphisms G into A, and a formally associative product in the groupoid $g_1 g_2$ provided $\alpha(g_1 g_2) = \alpha(g_1)$, $\beta(g_1 g_2) = \beta(g_2)$, and $\beta(g_1) = \alpha(g_2)$. Then the product is defined, and associative, $(g_1 g_2) g_3 = g_1(g_2 g_3)$.

In addition, there are natural left and right identity elements λ_g, ρ_g such that $\lambda_g g = g = g \rho_g$ [1].

An orbit of the groupoid G over A is an equivalence class for the relation $a_j \sim G a_k$ if and only if there is a groupoid element g with $\alpha(g) = a_j$ and $\beta(g) = a_k$. Following Cannas da Silva and Weinstein [2], we note that a groupoid is called transitive if it has just one orbit. The transitive groupoids are the building blocks of groupoids in that there is a natural decomposition of the base space of a general groupoid into orbits. Over each orbit, there is a transitive groupoid, and the disjoint union of these transitive groupoids is the original groupoid. Conversely, the disjoint union of groupoids is itself a groupoid.

The isotropy group of $a \in X$ consists of those g in G with $\alpha(g) = a = \beta(g)$. These groups prove fundamental to classifying groupoids.

If G is any groupoid over A, the map $(\alpha, \beta) : G \to A \times A$ is a morphism from G to the pair groupoid of A. The image of (α, β) is the orbit equivalence relation $\sim G$, and the functional kernel is the union of the isotropy groups. If $f : X \to Y$ is a function, then the kernel of f, $ker(f) = [(x_1, x_2) \in X \times X : f(x_1) = f(x_2)]$ defines an equivalence relation.

Groupoids may have additional structure. As Weinstein [1] explains, a groupoid G is a topological groupoid over a base space X if G and X are topological spaces

R. Wallace, *Cognitive Dynamics on Clausewitz Landscapes*,
https://doi.org/10.1007/978-3-030-26424-6_14

and α, β and multiplication are continuous maps. A criticism sometimes applied to groupoid theory is that their classification up to isomorphism is nothing other than the classification of equivalence relations via the orbit equivalence relation and groups via the isotropy groups. The imposition of a compatible topological structure produces a nontrivial interaction between the two structures.

In essence, a groupoid is a category in which all morphisms have an inverse, here defined in terms of connection to a base point by a meaningful path of an information source dual to a cognitive process.

As Weinstein [1] points out, the morphism (α, β) suggests another way of looking at groupoids. A groupoid over A identifies not only which elements of A are equivalent to one another (isomorphic), but *it also parameterizes the different ways (isomorphisms) in which two elements can be equivalent*, i.e., in our context, all possible information sources dual to some cognitive process. Given the information-theoretic characterization of cognition presented above, this produces a full modular cognitive network in a highly natural manner.

Brown [3] describes the fundamental structure as follows:

> A groupoid should be thought of as a group with many objects, or with many identities... A groupoid with one object is essentially just a group. So the notion of groupoid is an extension of that of groups. It gives an additional convenience, flexibility and range of applications...
>
> EXAMPLE 1. A disjoint union [of groups] $G = \cup_\lambda G_\lambda$, $\lambda \in \Lambda$, is a groupoid: the product ab is defined if and only if a, b belong to the same G_λ, and ab is then just the product in the group G_λ. There is an identity 1_λ for each $\lambda \in \Lambda$. The maps α, β coincide and map G_λ to λ, $\lambda \in \Lambda$.
>
> EXAMPLE 2. An equivalence relation R on [a set] X becomes a groupoid with $\alpha, \beta : R \to X$ the two projections, and product $(x, y)(y, z) = (x, z)$ whenever $(x, y), (y, z) \in R$. There is an identity, namely (x, x), for each $x \in X$...

Weinstein [1] makes the following fundamental point:

> Almost every interesting equivalence relation on a space B arises in a natural way as the orbit equivalence relation of some groupoid G over B. Instead of dealing directly with the orbit space B/G as an object in the category S_{map} of sets and mappings, one should consider instead the groupoid G itself as an object in the category G_{htp} of groupoids and homotopy classes of morphisms.

The groupoid approach has become quite popular in the study of networks of coupled dynamical systems which can be defined by differential equation models, (e.g., [4]).

14.2 Morse Theory

Morse theory examines relations between analytic behavior of a function—the location and character of its critical points—and the underlying topology of the manifold on which the function is defined. We are interested in a number of such functions, for example a "free energy" constructed from information source uncertainties on a parameter space and "second order" iterations involving parameter manifolds determining critical behavior. These can be reformulated from a Morse theory perspective. Here we follow closely Pettini [5].

The essential idea of Morse theory is to examine an n-dimensional manifold M as decomposed into level sets of some function $f : M \to \mathbf{R}$ where $\mathbf{R}$ is the set of real numbers. The a-level set of f is defined as

$$f^{-1}(a) = \{x \in M : f(x) = a\},$$

the set of all points in M with $f(x) = a$. If M is compact, then the whole manifold can be decomposed into such slices in a canonical fashion between two limits, defined by the minimum and maximum of f on M. Let the part of M below a be defined as

$$M_a = f^{-1}(-\infty, a] = \{x \in M : f(x) \leq a\}.$$

These sets describe the whole manifold as a varies between the minimum and maximum of f.

Morse functions are defined as a particular set of smooth functions $f : M \to \mathbf{R}$ as follows. Suppose a function f has a critical point x_c, so that the derivative $df(x_c) = 0$, with critical value $f(x_c)$. Then f is a Morse function if its critical points are nondegenerate in the sense that the Hessian matrix of second derivatives at x_c, whose elements, in terms of local coordinates are

$$H_{i,j} = \partial^2 f / \partial x^i \partial x^j,$$

has rank n, which means that it has only nonzero eigenvalues, so that there are no lines or surfaces of critical points and, ultimately, critical points are isolated.

The index of the critical point is the number of negative eigenvalues of H at x_c.

A level set $f^{-1}(a)$ of f is called a critical level if a is a critical value of f, that is, if there is at least one critical point $x_c \in f^{-1}(a)$.

Again following [5], the essential results of Morse theory are:

1. If an interval $[a, b]$ contains no critical values of f, then the topology of $f^{-1}[a, v]$ does not change for any $v \in (a, b]$. Importantly, the result is valid even if f is not a Morse function, but only a smooth function.

2. If the interval $[a, b]$ contains critical values, the topology of $f^{-1}[a, v]$ changes in a manner determined by the properties of the matrix H at the critical points.

3. If $f : M \to \mathbf{R}$ is a Morse function, the set of all the critical points of f is a discrete subset of M, i.e., critical points are isolated. This is Sard's Theorem.

4. If $f : M \to \mathbf{R}$ is a Morse function, with M compact, then on a finite interval $[a, b] \subset \mathbf{R}$, there is only a finite number of critical points p of f such that $f(p) \in [a, b]$. The set of critical values of f is a discrete set of $\mathbf{R}$.

5. For any differentiable manifold M, the set of Morse functions on M is an open dense set in the set of real functions of M of differentiability class r for $0 \leq r \leq \infty$.

6. Some topological invariants of M, that is, quantities that are the same for all the manifolds that have the same topology as M, can be estimated and sometimes computed exactly once all the critical points of f are known: Let the Morse numbers $\mu_i (i = 1, ..., m)$ of a function f on M be the number of critical points of f of index i, (the number of negative eigenvalues of H). The Euler characteristic of the complicated manifold M can be expressed as the alternating sum of the Morse numbers of any Morse function on M,

$$\chi = \sum_{i=0}^{m} (-1)^i \mu_i .$$

The Euler characteristic reduces, in the case of a simple polyhedron, to

$$\chi = V - E + F$$

where V, E, and F are the numbers of vertices, edges, and faces in the polyhedron.

7. Another important theorem states that, if the interval $[a, b]$ contains a critical value of f with a single critical point x_c, then the topology of the set M_b defined above differs from that of M_a in a way which is determined by the index, i, of the critical point. Then M_b is homeomorphic to the manifold obtained from attaching to M_a an i-handle, i.e., the direct product of an i-disk and an $(m - i)$-disk.

Again, Pettini [5] contains both mathematical details and further references. See, for example, Matusmoto [6].

14.3 An RDT Proof of the DRT

The Rate–Distortion Theorem of information theory asks how much a signal can be compressed and have average distortion, according to an appropriate measure, less than some predetermined limit $D > 0$. The result is an expression for the minimum necessary channel capacity, R, as a function of D. See Cover and Thomas [7] for details. Different channels have different expressions. For the Gaussian channel under the squared distortion measure,

$$\begin{aligned} R(D) &= \frac{1}{2} \log[\frac{\sigma^2}{D}] \; D < \sigma^2 \\ R(D) &= 0 \; D \geq \sigma^2 \end{aligned} \tag{14.1}$$

where σ^2 is the variance of channel noise having zero mean.

Our concern is how a control signal u_t is expressed in the system response x_{t+1}. We suppose it possible to deterministically retranslate an observed sequence of system outputs $x_1, x_2, x_3, ...$ into a sequence of possible control signals $\hat{u}_0, \hat{u}_1, ...$ and to compare that sequence with the original control sequence $u_0, u_1, ...$, with the difference between them having a particular value under the chosen distortion measure, and hence an observed average distortion.

The correspondence expansion is as follows.

Feynman [8], following ideas of Bennett, identifies information as a form of free energy. Thus $R(D)$, the minimum channel capacity necessary for average distortion D, is also a free energy measure, and we may define an entropy S as

$$S \equiv R(D) - DdR/dD \tag{14.2}$$

For a Gaussian channel under the squared distortion measure,

$$S = 1/2 \log[\sigma^2/D] + 1/2 \tag{14.3}$$

Other channels will have different expressions.

The simplest dynamics of such a system are given by a nonequilibrium Onsager equation in the gradient of S, [9] so that

$$dD/dt = -\mu dS/dD = \frac{\mu}{2D} \tag{14.4}$$

By inspection,

$$D(t) = \sqrt{\mu t} \tag{14.5}$$

which is the classic outcome of the diffusion equation. For the "natural" channel having $R(D) \propto 1/D$, $D(t) \propto$ the cube root of t.

This correspondence reduction allows an expansion to more complicated systems, in particular, to the control system of Fig. 1.2.

Let $\mathcal{H}$ be the rate at which control information is fed into an inherently unstable control system, in the presence of a further source of control system noise β, in addition to the channel noise defined by σ^2. The simplest generalization of Eq. (14.4), for a Gaussian channel, is the stochastic differential equation

$$dD_t = [\frac{\mu}{2D_t} - M(\mathcal{H})]dt + \beta D_t dW_t \tag{14.6}$$

where dW_t represents white noise and $M(\mathcal{H}) \geq 0$ is a monotonically increasing function.

This equation has the nonequilibrium steady-state expectation

$$D_{nss} = \frac{\mu}{2M(\mathcal{H})} \tag{14.7}$$

measuring the average distortion between what the control system wants and what it gets. In a sense, this is a kind of converse to the famous radar equation which states that a returned signal will be proportional to the inverse fourth power of the distance between the transmitter and the target. But there is a deeper result, leading to the DRT.

Applying the Ito chain rule to Eq. (14.6) [10, 11], it is possible to calculate the expected variance in the distortion as $E(D_t^2) - (E(D_t))^2$. But application of the Ito rule to D_t^2 shows that *no real number solution for its expectation is possible unless the discriminant of the resulting quadratic equation is* ≥ 0, so that a necessary condition for stability is

$$\begin{aligned} M(\mathscr{H}) &\geq \beta\sqrt{\mu} \\ \mathscr{H} &\geq M^{-1}(\beta\sqrt{\mu}) \end{aligned} \tag{14.8}$$

where the second expression follows from the monotonicity of M.

As a consequence of the correspondence reduction leading to Eq. (14.4), we have generalized the DRT of Eq. (1.2). Different "control channels", with different forms of $R(D)$, will give different detailed expressions for the rate of generation of "topological information" by an inherently unstable system.

14.4 An Information Black–Scholes Model

We look at $\mathscr{H}(\rho)$ as the control information rate "cost" of stability at the integrated environmental insult ρ. To determine the mathematical form of $\mathscr{H}(\rho)$ under conditions of volatility, i.e., variability proportional to a signal, we must first model the variability of ρ, most simply taken as

$$d\rho_t = g(t, \rho_t)dt + b\rho_t dW_t \tag{14.9}$$

Here, dW_t is white noise and—counterintuitively—the function $g(t, \rho)$ will fall out of the calculation on the assumption of certain regularities.

$\mathscr{H}(\rho_t, t)$ is the minimum needed incoming rate of control information under the Data Rate Theorem. Expand $\mathscr{H}$ in ρ *using the Ito chain rule* [10]:

$$\begin{aligned} d\mathscr{H}_t = [\partial\mathscr{H}/\partial t + g(\rho_t, t)\partial\mathscr{H}/\partial\rho + \frac{1}{2}b^2\rho_t^2\partial^2\mathscr{H}/\partial\rho^2]dt \\ +[b\rho_t\partial\mathscr{H}/\partial\rho]dW_t \end{aligned} \tag{14.10}$$

It is now possible to define a Legendre transform, L, of the rate $\mathscr{H}$, by convention having the form

$$L = -\mathscr{H} + \rho\partial\mathscr{H}/\partial\rho \tag{14.11}$$

$\mathscr{H}$ is an information index, a free energy measure in the sense of Feynman [8], so that L is a classic entropy measure.

We make an approximation, replacing dX with ΔX and applying Eq. (14.10), so that

$$\Delta L = (-\partial \mathscr{H}/\partial t - \frac{1}{2} b^2 \rho^2 \partial^2 \mathscr{H}/\partial \rho^2)\Delta t \quad (14.12)$$

According to the classical Black–Scholes model [12], the terms in g and dW_t "cancel out", and white noise has been subsumed into the Ito correction factor, a regularity assumption making this an exactly solvable but highly approximate model.

The conventional Black-Scholes calculation takes $\Delta L/\Delta T \propto L$. At nonequilibrium steady state, by some contrast, we can assume $\Delta L/\Delta t = \partial \mathscr{H}/\partial t = 0$, giving

$$-\frac{1}{2} b^2 \rho^2 \partial^2 \mathscr{H}/\partial \rho^2 = 0 \quad (14.13)$$

so that

$$\mathscr{H} \approx \kappa_1 \rho + \kappa_2 \quad (14.14)$$

The κ_i will be nonnegative constants.

A more general approach to the Black-Scholes approximation is possible.

Indeed, using Eq. (14.11) and adapting the classic Black-Scholes condition to $dL/dt = -\alpha L$, it is possible to solve the full PDE in terms of a product $\mathscr{H} \approx F_1(\rho)F_2(t)$, for example using the "pdsolve" function of the computer algebra program Maple. Imposing the stationary condition $dF_2/dt = 0$ then gives

$$\mathscr{H} \approx \kappa_1 \rho + \kappa_2 \rho^{2\alpha/b^2} \quad (14.15)$$

where the boundary condition defining the parameters is that $\mathscr{H}$ is both monotonic increasing in ρ and ≥ 0.

For small α and large b^2, Eq. (14.15) reduces to the simple linear approximation of Eq. (14.14).

Adopting the "exponential" relation $dL/dt = \beta - \alpha L$, and again assuming a stationary solution, leads to a more complicated expression including the result of Eq. (14.15). Escalation to generalities of the form $dL/dt = f(L)$ rapidly outruns available computer algebra programs.

14.5 Symmetry and Information

As repeatedly emphasized above, a central difficulty faces the work presented here. Although the safest route is to simply stop at Pielou's characterization of mathematical models as useful for raising questions, but not answering them, one seeks possible new tools for data analysis and policy purposes. Pielou herself recommends autoregressive moving average models as potentially of use in the management of real ecosystems. However, conflict, unlike ecological or evolutionary processes, involves active cognition, and must therefore be addressed using tools that arise from, and accurately reflect, the special nature of cognitive phenomena. Lamarckian evolutionary dynamics, and the shorter-time frame of "Lamarckian" control theory, are very special indeed.

As Atlan and Cohen [13] emphasize, the central feature of cognition is the choice of one possible action from a larger set of those available. Choice implies a reduction in uncertainty, and such reduction implies the existence of an information source dual to the cognitive process under study. The argument is both direct and sufficient.

From the strategic perspective, a particularly essential matter is the structure of the "phase transitions" available to the contending agents. These are determined, in physical systems, by the underlying group symmetries characterizing the system. Cognitive phenomena are, of course, far more subtle and, we argue, are to be studied using groupoid generalizations of group theory.

There is support for this approach in the literature.

Yeung [14], in particular, has explored the unexpected relations between information theory and the theory of finite groups. These mirror the relation between synchronicity in networks and their permutation symmetries as shown in Golubitsky and Stewart [4].

More explicitly, given random variables X_1 and X_2 having Shannon uncertainties $H(X_1)$ and $H(X_2)$, the information theory chain rule [4] states that, for their joint uncertainty, which we write as $H(X_1, X_2)$,

$$H(X_1) + H(X_2) \geq H(X_1, X_2) \tag{14.16}$$

Similarly, let G be any finite group, and G_1, G_2 be subgroups of G. Take $|G|$ as the order of a group—the number of elements. Then the intersection $G_1 \cap G_2$ is also a subgroup, and a "group inequality" can be derived that is the exact analog of Eq. (5.1):

$$\log[\frac{|G|}{|G_1|}] + \log[\frac{|G|}{|G_2|}] \geq \log[\frac{|G|}{|G_1 \cap G_2|}] \tag{14.17}$$

Yeung defines a probability for a pseudorandom variate associated with a group G as $Pr\{X = a\} = 1/|G|$. This allows construction of a group-characterized information source, noting that, in general, the joint uncertainty of a set of random variables is not necessarily the logarithm of a rational number. Yeung [14] ultimately finds a surprising one-to-one correspondence between unconstrained information inequalities—generalizations of Eq. (14.16)—and finite group inequalities: unconstrained inequalities can be proved by techniques in group theory, and many group-theoretic inequalities can be proven by techniques of information theory. Yeung uses an obscure unconstrained information inequality to derive, in his Eq. (16.116), a complex group inequality for which, as he puts it, the "...implications in group theory are yet to be understood".

An argument in this direction follows from the fundamental dogma of algebraic topology [15]: forming algebraic images of topological spaces. The most basic of these images is the fundamental group, leading to van Kampen's Theorem allowing the computation of the fundamental group of spaces that can be decomposed into simpler spaces whose fundamental group is already known. As Hatcher ([15], p. 40) puts it, "By systematic use of this theorem one can compute the fundamental groups of a very large number of spaces... [F]or every group G there is a space X_G whose fundamental group is isomorphic to G". As Golubitsky and Stewart forcefully argue, network structures and dynamics are imaged by fundamental groupoids, for which there also exists a version of the Siefert–van Kampen theorem [16]. Yeung's [14] results suggest information theory-based "cognitive" generalizations that may include essential dynamics of conflict and its regulation.

Extending this work to the relations between groupoid structures and rather subtle information theory properties may provide one route toward the development of useful statistical models. More direct routes may exist.

This being said, the development of useful statistical tools from what are essentially probability models is difficult. It is one thing to raise questions using mathematical models of complex phenomena. It is quite another to answer, or at least address, those questions using statistical tools on real data sets.

References

1. Weinstein, A. 1996. Groupoids: Unifying internal and external symmetry. *Notices of the American Mathematical Association* 43: 744–752.
2. Cannas da Silva, A., and A. Weinstein. 1999. *Geometric models for noncommutative algebras*. Providence, RI: American Mathematical Society.
3. Brown, R. 1987. From groups to groupoids: A brief survey. *Bulletin of the London Mathematical Society* 19: 113–134.
4. Golubitsky, M., and I. Stewart. 2006. Nonlinear dynamics of networks: The groupoid formalism. *Bulletin of the American Mathematical Society* 43: 305–364.
5. Pettini, M. 2007. *Geometry and topology in hamiltonian dynamics and statistical mechanics*. New York: Springer.
6. Matsumoto, Y. 2002. *An introduction to morse theory*. Providence, RI: American Mathematical Socity.
7. Cover, T., and J. Thomas. 2006. *Elements of information theory*, 2nd ed. New York: Wiley.
8. Feynman, R. 2000. *Lectures on computation*. Boulder CO: Westview Press.
9. de Groot, S., and P. Mazur. 1984. *Nonequilibrium thermodynamics*. New York: Dover.
10. Protter, P. 1990. *Stochastic integration and differential equations*. New York: Springer.
11. Khashminskii, R. 2012. *Stochastic stability of differential equations*, 2nd ed. New York: Springer.
12. Black, F., and M. Scholes. 1973. The pricing of options and corporate liabilities. *Journal of Political Economy* 81: 637–654.
13. Atlan, H., and I. Cohen. 1998. Immune information, self-organization, and meaning. *International Immunology* 10: 711–717.
14. Yeung, H. 2008. *Information theory and network coding*. Springer, New York. www.bangor.ac.uk/r.brown/nonab-a-t.html.
15. Hatcher, A. 2001. *Algebraic topology*. New York: Cambridge University Press.
16. Brown, R., P. Higgins, and R. Sivera. 2011. *Nonabelian algebraic topology: Filtered spaces, crossed complexes, cubical Homotopy Groupoids, EMS tracts in mathematics*, Vol. 15