AT&T Bell Laboratories Technical Journal
Vol. 63, No. 7, September 1984
Printed in U.S.A.

On the Capacity of Sticky Storage Devices

By H. S. WITSENHAUSEN*

(Manuscript received December 26, 1983)

When required to make a transition to a new state, a memory cell may, with a probability dependent on the state, refuse to do so (i.e., "stick"). Assuming that error correcting codes can be used at each write-read cycle, one seeks the maximum error-free (in the Shannon sense), long-term average capacity per cell and cycle. This problem is solved here for binary cells with either unilateral or symmetric stickiness. The methods used apply to more general cases as well. In the Appendix, some essential inequalities of dynamic programming are demonstrated.

I. INTRODUCTION

Information theoretic studies of storage devices have been mostly concerned with overcoming the existence of subsets of permanently defective cells.[1] We are concerned instead with the case of identical cells with the deficiency that use of a cell in one write-read cycle affects the cell's behavior in the next cycle. An extreme example of this is "write-once" memory,[2,5,9] such as punched paper tape or optical disks, where the long-term average throughput per cell and cycle is, of course, zero. A previous paper[8] considers a deterministic cell model in which the aftereffect of usage is of limited duration, permitting positive average rates. The present paper similarly considers the perhaps more realistic case of stochastic cells.

In this model, the store has N cells, each of which can be viewed as an input-output automaton. The store is used for T successive cycles.

* AT&T Bell Laboratories.

At each cycle, fresh data from a source S_t, independent of all earlier sources, are encoded into an N-vector $\mathbf{X}_t$ of inputs applied to the cells. The resulting N-vector $\mathbf{Y}_t$ of cell outputs goes to a decoder that must reproduce the source output with a probability of error approaching 0 as N increases. Note that if cell states can only change on inputs, not spontaneously, then the reading operation may be repeated any number of times in the same cycle with the same results. At each cycle only the probability distribution of the state of the store is known. This distribution, together with the cell model, defines a channel relating $\mathbf{Y}_t$ to $\mathbf{X}_t$. The difficulty of the problem is that operating at or near the capacity of that channel at one cycle may leave an unfavorable state distribution for the next cycle. Thus, the largest possible number of bits per cell and cycle that are accurately recoverable can only be determined by dynamic programming. To make such a program manageable, a theorem is first proven that permits one to obtain the limit for large N at each cycle by considerations involving a single cell. This theorem applies to any cell for which state and output are one and the same.

Results are obtained for two types of binary cells. For the first type, only one of the two states is "sticky": when the cell is in that state and the input requires transition to the other state, there is a probability ϵ that the transition does not take place. For the second type, both states are assumed sticky in the same sense. The computations reveal that for both models the long-term policy is steady rather than periodic. For the symmetric case, this implies that the maximum throughput per cell and cycle is just the capacity of the binary symmetric channel with crossover probability one-half the sticking probability.

II. INFORMATION-THEORETIC BOUNDS

Let $\mathscr{X}$, $\mathscr{Y}$, and $\mathscr{W}$ be finite alphabets, and let $X_t^i \in \mathscr{X}$ denote the input and $Y_t^i \in \mathscr{Y}$ the output of cell i at cycle t. Let $W_t^i \in \mathscr{W}$ be independent random variables representing the internal random effects of cell i at cycle t.

The initial state of cell i is represented by Y_0^i in $\mathscr{Y}$. We use the notation $\mathbf{X}_t = (X_t^1, \cdots, X_t^N)$, $\mathbf{Y}_t = (Y_t^1, \cdots, Y_t^N)$, $\mathbf{W}_t = (W_t^1, \cdots, W_t^N)$. The quantities $\mathbf{X}_1, \cdots, \mathbf{X}_T$, $Y_0^1, \cdots, Y_0^N$, $W_1^1, \cdots, W_T^N$ (a total of T vectors in $\mathscr{X}^N$, N variables in $\mathscr{Y}$, and NT variables in $\mathscr{W}$) are jointly independent.

The distributions of the Y_0^i and W_t^i are given, together with the equation describing the operation of the cells

$$Y_t^i = f_t^i(X_t^i, Y_{t-1}^i, W_t^i). \tag{1}$$

The distribution of the $\mathbf{X}_t$ is dependent upon the choice of encoding, and the source statistics.

The case of interest is the "homogeneous" one where:

1. f_t^i is the same function f for all i and t.
2. The Y_0^i have the same distribution as a generic Y_0 for all i.
3. The W_t^i have the same distribution as a generic W for all i and t.

The more general case is allowed in the model because the main theorem below is valid without the homogeneity assumptions.

At each cycle t, the encoder and decoder are designed with knowledge of the distribution of past events but not of their realizations. Thus, neither encoder nor decoder knows the exact state of the memory (i.e., the vector $\mathbf{Y}_{t-1}$) at the end of the previous cycle. As $\mathbf{W}_t$ is also known only in distribution, the memory appears as a noisy channel with input $\mathbf{X}_t$ and output $\mathbf{Y}_t$. Then the entropy of the message that can be reconstructed with negligible error at cycle t is bounded, in view of the data-processing theorem,* by

$$I(\mathbf{X}_t;\ \mathbf{Y}_t),$$

where I is the mutual information. For the total throughput, one thus has the upper bound

$$\sum_{t=1}^{T} I(\mathbf{X}_t;\ \mathbf{Y}_t). \tag{2}$$

Theorem 1:

$$\sum_{t=1}^{T} I(\mathbf{X}_t;\ \mathbf{Y}_t) \le \sum_{i=1}^{N} \sum_{t=1}^{T} I(X_t^i;\ Y_t^i). \tag{3}$$

Proof: By the independence of the initial states, the internal disturbances, and the sources, one has (H denoting entropy)

$$H(\mathbf{Y}_0) = \sum_{i=1}^{N} H(Y_0^i), \tag{4}$$

$$H(\mathbf{W}_t) = \sum_{i=1}^{N} H(W_t^i) \qquad t = 1, \cdots, T, \tag{5}$$

and also

$$H(X_t^i Y_t^i W_t^i Y_{t-1}^i) = H(X_t^i) + H(Y_{t-1}^i) + H(W_t^i), \tag{6}$$

since Y_t^i is determined by (1), and Y_{t-1}^i depends only on the initial state, the earlier source encodings, and the earlier disturbances, all independent of $\mathbf{X}_t$, hence of X_t^i. And for the same reason

* If X and Z are conditionally independent given Y, then $I(X; Z) \le I(X; Y)$.

$$H(\mathbf{X}_t\mathbf{Y}_t\mathbf{W}_t\mathbf{Y}_{t-1}) = H(\mathbf{X}_t) + H(\mathbf{Y}_{t-1}) + H(\mathbf{W}_t). \tag{7}$$

The following four relations are standard properties of entropy:

$$H(\mathbf{X}_t\mathbf{Y}_t\mathbf{W}_t\mathbf{Y}_{t-1}) - H(\mathbf{X}_t\mathbf{Y}_t) = H(\mathbf{Y}_{t-1}\mathbf{W}_t \,|\, \mathbf{X}_t\mathbf{Y}_t) \tag{8}$$

$$H(X_t^iY_t^iW_t^iY_{t-1}^i) - H(X_t^iY_t^i) = H(Y_{t-1}^iW_t^i \,|\, X_t^iY_t^i) \tag{9}$$

$$H(\mathbf{Y}_{t-1}\mathbf{W}_t \,|\, \mathbf{X}_t\mathbf{Y}_t) \le \sum_{i=1}^{N} H(Y_{t-1}^iW_t^i \,|\, X_t^iY_t^i) \tag{10}$$

$$H(\mathbf{Y}_T) \le \sum_{i=1}^{N} H(Y_T^i). \tag{11}$$

Using these facts one has

$$\sum_{t=1}^{T} I(\mathbf{X}_t, \mathbf{Y}_t)$$

$$= \sum_{t=1}^{T} H(\mathbf{X}_t) + H(\mathbf{Y}_t) - H(\mathbf{X}_t\mathbf{Y}_t) \text{ (by definition)}$$

$$= \sum_{t=1}^{T} H(\mathbf{Y}_t) + H(\mathbf{X}_t\mathbf{Y}_t\mathbf{W}_t\mathbf{Y}_{t-1} - H(\mathbf{X}_t\mathbf{Y}_t) - H(\mathbf{Y}_{t-1})$$

$$- H(\mathbf{W}_t) \text{ [by (7)]}$$

$$= H(\mathbf{Y}_T) - H(\mathbf{Y}_0) + \sum_{t=1}^{T} [H(\mathbf{Y}_{t-1}\mathbf{W}_t \,|\, \mathbf{X}_t\mathbf{Y}_t) - H(\mathbf{W}_t)]$$

[by summation and (8)]

$$\le \sum_{i=1}^{N} -[H(Y_T^i) - H(Y_0^i) + \sum_{t=1}^{T} H(Y_{t-1}^iW_t^i \,|\, X_t^iY_t^i) - H(W_t^i)]$$

[by (11), (4), (10), and (5)]

$$= \sum_{i=1}^{N} \sum_{t=1}^{T} (H(Y_t^i) - H(Y_{t-1}^i) - H(W_t^i) + H(Y_{t-1}^iW_t^iX_t^iY_t^i) - H(X_t^iY_t^i)$$

[by summation and (9)]

$$= \sum_{i=1}^{N} \sum_{t=1}^{T} H(Y_t^i) + H(X_t^i) - H(X_t^iY_t^i) \text{ [by (6)]}$$

$$= \sum_{i=1}^{N} \sum_{t=1}^{T} I(X_t^i; Y_t^i) \text{ (by definition)},$$

which was to be proved.

In the homogeneous case, the problem of finding the maximum M of $\sum_{t=1}^{T} I(X_t^i; Y_t^i)$ over all distributions of the independent variables

$X_1^i, \cdots, X_T^i$ subject to (1) is the same for all i. Having found M by solving this generic single-cell problem, a bound of NM is established for the total throughput. On the other hand, a throughput per cell, arbitrarily close to M, can be achieved for large enough N by choosing encoders with code words picked at random with each X_t^i independently having the maximizing distribution of the generic problem at stage t. Thus, the bound is sharp, asymptotically in N.

The bound is obtained, for any given cell model, by solving a T-stage, finite horizon, dynamic program. For large T it is the maximum long-term average per cycle (and cell) that is of prime interest. Some of the dynamic programming issues are discussed in the Appendix.

III. UNILATERAL STICKINESS

It is assumed that a binary cell acquires its input as new state when its previous state was 0, but when the previous state was 1 and the input is 0 the cell remains stuck at 1 with probability ϵ.

This is modeled, with X_t, Y_t, $W_t \in \{0, 1\}$ by letting $W_t = 1$ with probability ϵ and 0 with $1 - \epsilon$, and in ordinary arithmetic,

$$Y_t = X_t + (1 - X_t)Y_{t-1}W_t \qquad t = 1, \cdots, T. \tag{12}$$

Let

$$p_t = \Pr\{X_t = 0\}, \tag{12a}$$

$$s_t = \Pr\{Y_t = 0\}, \tag{12b}$$

and

$$\epsilon = \Pr\{W_t = 1\}. \tag{12c}$$

Then (12) implies

$$s_t = p_t(1 - \epsilon(1 - s_{t-1})), \tag{13}$$

and one obtains, with h the binary entropy function,

$$\begin{aligned} I(X_t; Y_t) &= h(p_t(1 - \epsilon(1 - s_{t-1})) - p_t h(\epsilon(1 - s_{t-1})) \\ &= h(s_t) - p_t h\left(\frac{s_t}{p_t}\right). \end{aligned} \tag{14}$$

This leads to the dynamic program (where τ denotes the number of cycles to go)

$$V_\tau(s) = \max_p h(p(1 - \epsilon + \epsilon s)) - ph(\epsilon(1 - s)) + V_{\tau-1}(p(1 - \epsilon + \epsilon s)) \tag{15}$$

with $V_0(s) = 0$.

This is an easy task for a computer, and the results show that p and s soon stabilize around steady-state values. Furthermore, as the Appendix shows, using the value functions found in the finite horizon solution, one can derive both upper and lower bounds on the optimal long-term average per cycle. In this problem these bounds soon agree to many decimals. The computer results are used only to conclude that the long-term optimum is steady. (This cannot be taken for granted, as often such problems have periodic solutions, crop rotation being the most ancient example of this.)

In a steady state, the constant values of p and s must satisfy, by (13), the relation

$$s = p(1 - \epsilon(1 - s)) \tag{16}$$

and by (14) the optimal throughput per cell and cycle is the maximum of

$$h(s) - ph\left(\frac{s}{p}\right) \tag{17}$$

subject to (16).

This amounts to maximizing a transcendental function on the unit interval. While no closed-form solution is known, the maximum, and the corresponding p and s are easily computed, and they are given as a function of ϵ in Table I. (The maximum is given in base 2, i.e., bits per cell and cycle.) It is important to note that for fixed T, the dynamic program (15) defines a total throughput which can only be approached from below as N increases. On the other hand, the long-range average is approached from above, because for finite T the transient effect of an initially clear memory will permit a slightly higher total.

As the case of small ϵ is of greatest interest we note the expansions [following from (16) and (17)]: at the maximum

$$s = \frac{1}{2} - \frac{\epsilon}{8} + o(\epsilon), \tag{18}$$

and

$$p = \frac{1}{2} + \frac{\epsilon}{8} + o(\epsilon), \tag{19}$$

and the value of the maximum is

$$\log 2 - \frac{1}{2} h\left(\frac{\epsilon}{2}\right) + o(\epsilon). \tag{20}$$

Table I—Results for unilateral sticking

Epsilon	Maximum	Input p	State s
0.00	1.00000000	0.50000000	0.50000000
0.01	0.97718297	0.50125629	0.49874371
0.02	0.95921284	0.50252532	0.49747469
0.03	0.94300294	0.50380733	0.49619267
0.04	0.92790483	0.50510257	0.49489743
0.05	0.91361310	0.50641131	0.49358869
0.06	0.89994925	0.50773381	0.49226619
0.07	0.88679616	0.50907034	0.49092966
0.08	0.87407111	0.51042119	0.48957881
0.09	0.86171265	0.51178665	0.48821335
0.10	0.84967335	0.51316702	0.48683298
0.15	0.79315764	0.52030370	0.47969631
0.20	0.74099706	0.52786405	0.47213596
0.25	0.69176025	0.53589839	0.46410162
0.30	0.64460887	0.54446658	0.45553343
0.35	0.59898428	0.55364065	0.44635936
0.40	0.55447912	0.56350833	0.43649168
0.45	0.51077418	0.57417812	0.42582189
0.50	0.46760281	0.58578644	0.41421357
0.55	0.42472834	0.59850838	0.40149163
0.60	0.38192756	0.61257412	0.38742589
0.65	0.33897650	0.62829543	0.37170459
0.70	0.29563534	0.64611064	0.35388938
0.75	0.25162917	0.66666668	0.33333334
0.80	0.20661826	0.69098302	0.30901701
0.85	0.16014419	0.72082550	0.27917453
0.90	0.11151237	0.75974695	0.24025309
0.95	0.05945393	0.81725604	0.18274403
0.96	0.04841568	0.83333337	0.16666671
0.97	0.03707783	0.85236594	0.14763415
0.98	0.02536387	0.87610072	0.12389940
0.99	0.01313736	0.90909100	0.09090918

IV. SYMMETRIC STICKINESS

Suppose that, when its input would require a change of state, a cell remains, with probability ϵ, in its former state.

This is modeled, with X_t, Y_t, $W_t \in \{0, 1\}$ by letting $W_t = 1$ with probability ϵ and

$$Y_t = (1 - W_t)X_t + W_t Y_{t-1}. \tag{21}$$

Let

$$p_t = \Pr\{X_t = 0\}, \tag{22a}$$

$$s_t = \Pr\{Y_t = 0\}, \tag{22b}$$

$$\epsilon = \Pr\{W_t = 1\}. \tag{22c}$$

Then (21) implies

$$s_t = (1 - \epsilon)p_t + \epsilon s_{t-1} \tag{23}$$

and

$$I(X_t; Y_t) = h((1 - \epsilon)p_t + \epsilon s_{t-1}) - p_t h(\epsilon(1 - s_{t-1})) + (1 - p_t)h(\epsilon s_{t-1}). \tag{24}$$

Note that, if both states are equally likely, then at the next cycle one faces a binary symmetric channel with crossover probability $\epsilon/2$. The capacity of this channel

$$\log 2 - h(\epsilon/2) \tag{25}$$

is attained by choosing symmetrically distributed input and this will lead to a symmetric distribution of the next state, so that this situation perpetuates itself. So (25) is an achievable long-term average, and it will turn out to be the best possible.

Assuming only that the optimum is time-invariant, one has, by (23), the condition

$$s = p, \tag{26}$$

and this reduces the problem to showing that the maximum, over $0 \le s \le 1$, of

$$h(s) - sh(\epsilon(1 - s)) - (1 - s)h(\epsilon s) \tag{27}$$

is at $s = 1/2$, where its value is $\log 2 - h(\epsilon/2)$. When $s = 0$ or 1, (27) vanishes. For fixed s in (0, 1) let

$$F(\epsilon) = \log 2 - h(\epsilon/2) - h(s) + (1 - s)h(\epsilon s) + sh(\epsilon(1 - s)). \tag{28}$$

It suffices to show that $F(\epsilon) \ge 0$ for $0 < \epsilon < 1$. One has $F(1) = F'(1) = 0$ and

$$F''(\epsilon) = \frac{(1 - \epsilon)(1 - 4s(1 - s))}{\epsilon(2 - \epsilon)(1 - \epsilon s)(1 - \epsilon(1 - s))} \ge 0. \tag{29}$$

So F is convex and tangent to the ϵ axis at $\epsilon = 1$, hence nonnegative, as required.

To confirm the steady-state solution, one runs the finite horizon dynamic program, with τ stages to go:

$$V_\tau(s) = \max_p h((1 - \epsilon)p + \epsilon s) - ph(\epsilon(1 - s)) - (1 - p)h(\epsilon s) + V_{\tau-1}((1 - \epsilon)p + \epsilon s) \tag{30}$$

with $V_0(s) \equiv 0$.

Using a grid of probabilities including $p = s = 1/2$, the upper and lower bounds on the long-term average (see Appendix), derived from the value functions, converge towards each other. In Table II the

Table II—Upper and lower bounds on optimal long-term average per cycle

Epsilon	Bounds	Formula
0.05	[0.83133907, 0.83133907]	0.8313390685
0.10	[0.71360304, 0.71360304]	0.7136030429
0.15	[0.61568846, 0.61568846]	0.6156884559
0.20	[0.53100441, 0.53100441]	0.5310044064
0.25	[0.45643556, 0.45643556]	0.4564355568
0.30	[0.39015970, 0.39015970]	0.3901596953
0.35	[0.33098416, 0.33098416]	0.3309841649
0.40	[0.27807191, 0.27807191]	0.2780719051
0.45	[0.23080717, 0.23080717]	0.2308071710
0.50	[0.18872188, 0.18872188]	0.1887218755
0.55	[0.15145182, 0.15145812]	0.1514518217
0.60	[0.11870910, 0.11870910]	0.1187091008
0.65	[0.09026388, 0.09026388]	0.0902638775
0.70	[0.06593194, 0.06593194]	0.0659319446
0.75	[0.04556600, 0.04556600]	0.0455659971
0.80	[0.02904941, 0.02904941]	0.0290494055
0.85	[0.01629174, 0.01629174]	0.0162917374
0.90	[0.00722555, 0.00722555]	0.0072255460
0.95	[0.00180412, 0.00180412]	0.0018041210

bounds shown have been obtained by iterating until the difference is below $5 \cdot 10^{-9}$ and they are compared with formula (25), all logarithms are in base 2.

V. CONCLUSIONS

While complex coding may be impractical, the value of the above results is that they provide precise bounds on what is achievable. One notes, comparing (20) with (25), that two-sided stickiness is, for the same small sticking probability ϵ, twice as damaging as one-sided stickiness. The case of two different positive sticking probabilities can be handled by the same techniques.

Note that the problem can be considered in another light by interchanging time and space. It then becomes a special case of the general interference channel, as shown in an earlier paper.[8]

REFERENCES

1. C. Heegard and A. A. El Gamal, "On the Capacity of Computer Memory With Defects" I.E.E.E. Trans. Inform. Theory, *IT-29*, No. 5 (September 1983), pp. 731–9.
2. C. Heegard, "On the Capacity of Permanent Memory," 1983 Conf. on Inform. Sciences and Systems, Johns Hopkins University (March 1983).
3. A. R. Odoni, "On Finding the Maximal Gain for Markov Decision Processes," Oper. Res., *17*, No. 5 (September–October 1969), pp. 857–60.
4. M. L. Puterman, ed., *Dynamic Programming and Its Applications*, New York: Academic Press, 1978.
5. R. L. Rivest and A. Shamir, "How to Reuse a 'Write-Once' Memory," Inform. Contr., *55*, No. 1–3 (October–November–December 1982), pp. 1–19.

6. D. J. White, *Finite Dynamic Programming*, New York: John Wiley and Sons, 1978.
7. P. Whittle, *Optimization Over Time*, Vol. I., New York: John Wiley and Sons, 1982.
8. H. S. Witsenhausen and A. D. Wyner, "On Storage Media With Aftereffects," Inform. Contr., *56*, No. 3 (March 1983), pp. 199–211.
9. J. K. Wolf et al., "Coding for a Write-Once Memory," AT&T Bell. Lab. Tech. J., *63*, No. 6, Part 2 (July–August 1984), pp. 1089–112.
10. N. A. J. Hastings, "Bounds on the Gain of a Markov Decision Process," Oper. Res., *19*, No. 1 (January–February 1971), pp. 240–4.

APPENDIX

Remarks on Dynamic Programming

A.1 Introduction

The literature on dynamic programming (e.g., Refs. 4, 6, 7) is primarily devoted to the important case of stochastic systems (controlled Markov chains). Deterministic dynamic programs are covered as a special case of degenerate probabilities. As a result, the theorems proven in most texts and papers make assumptions that are not necessary for the most important deterministic results we need. Among such assumptions the following two are particularly bothersome:

1. The state space must be finite.
2. All policies must be ergodic.

The first assumption rules out even the case of a real interval as state space (which is our case). The second is essentially never true in a deterministic setup. This is our justification for giving the very simple proof of the results we need in this paper.

We are concerned with the optimal long-term average return per stage. It has been claimed that this case is only academic because either the number of stages is small, and then a finite horizon treatment is appropriate, or it is large, and then the time span involved is such that the time value of utility must be included by using the discounted model instead. In this paper, as millions of memory cycles can take place per second, we have a strong counterexample to this viewpoint: the undiscounted, long-term average is the appropriate quantity to study.

A.2 The deterministic case

The deterministic programs are defined by two sets, S, U, and two functions: r: $S \times U \rightarrow R$, f: $S \times U \rightarrow S$. Here S is the state space, U the control variable set, $r(s, u)$ the return using $s \in S$, $u \in U$ and $f(s, u)$ is the next state. The finite horizon program is thus written

$$V_\tau(s) = \sup_{u \in U} r(s, u) + V_{\tau-1}(f(s, u)), \tag{31}$$

where τ is the number of stages to go and $V_0(s) \equiv 0$.

It has long been known that one can eliminate U, r, and f in favor of a relation on states and a function defined on this relation. Specifically, let $\rho(s) \subset S$ be defined by

$$\rho(s) = f(s, U), \tag{32}$$

then the pairs $\{(s, s') \mid s' \in \rho(s)\}$ form a subset ρ of $S \times S$. Also, for $s' \in \rho(s)$, let

$$k(s, s') \equiv \sup_u \{r(s, u) \mid f(s, u) = s'\}. \tag{33}$$

In words, one need only know (*i*) which states one can go to next, and (*ii*) what is the optimal value of going there. The finite horizon problem then becomes

$$V_\tau(s) = \sup_{s' \in \rho(s)} k(s, s') + V_{\tau-1}(s'). \tag{34}$$

This reformulation (which is not possible for stochastic problems) has two advantages:

(*i*) Theoretical analysis is simplified.

(*ii*) In computation, when an infinite S is approximated by a finite subset S_1, no requantization is needed, avoiding this cause of error accumulation. One solves the problem for the subset, with ρ and k restricted to $S_1 \times S_1$.

A.3 Bellman's equation for the long-term average

If λ denotes the optimal long-term average return per stage, then it should satisfy the equation, already given by Bellman

$$\lambda + W(s) = \sup_{s' \in \rho(s)} k(s, s') + W(s'). \tag{35}$$

for all $s \in S$.

The unknown constant λ and unknown function W (which matters only modulo the addition of an arbitrary constant) are reminiscent of an eigenvalue equation. However, the maximum operator, while nonlinear, has a more favorable numerical behavior. Of course (35) is stated on the assumption that the required limit exists: it assumes that a policy for which the lim sup is as large as possible actually gives a limit. This depends on the structure of relation ρ, but we will not pursue this question, as our need is rather for inequalities.

A.4 Performance inequalities

Let V be any bounded real function on the state space S. For instance, V could be a value function obtained from a finite horizon program, an average of several such functions (in periodic cases), or just a plain wild guess. Then define $\bar{V}$ by

$$\bar{V}(s) \equiv \sup_{s' \in \rho(s)} k(s, s') + V(s'). \tag{36}$$

Let

$$M \equiv \sup_{s \in S} \bar{V}(s) - V(s), \tag{37}$$

$$m \equiv \inf_{s \in S} \bar{V}(s) - V(s). \tag{38}$$

Theorem 2: For any bounded V, with M and m defined by (36, 37, 38): (*i*) *The supremum, over all admissible policies, of the limit inferior of the average return is at least m.* (*ii*) *The limit superior of the average return, under any admissible policy, cannot exceed M.*

Proof: (*i*) The supremum on the right of (36) is attained, exactly or within ϵ, by choice of $s' \in \rho(s)$ for each s. Thus there exists, for any $\epsilon > 0$, a function $\phi_\epsilon: S \rightarrow S$ such that, for all s in S,

$$\bar{V}(s) \geq k(s, \phi_\epsilon(s)) + V(\phi_\epsilon(s)) \geq \bar{V}(s) - \epsilon, \tag{39}$$

and $\phi_\epsilon(s) \in \rho(s)$. As V is bounded, let

$$c \equiv \sup_{s \in S} V(s) - \inf_{s \in S} V(s). \tag{40}$$

From (38) one has $\bar{V}(s) - V(s) \geq m$, hence by (39),

$$k(s, \phi_\epsilon(s)) + V(\phi_\epsilon(s)) - V(s) \geq m - \epsilon. \tag{41}$$

Generate a feasible sequence of states from some initial s_0 by

$$s_t = \phi_\epsilon(s_{t-1}), \qquad t = 1, 2, \cdots. \tag{42}$$

For this sequence

$$\begin{aligned} k(s_0, s_1) + V(s_1) - V(s_0) &\geq m - \epsilon \\ k(s_1, s_2) + V(s_2) - V(s_1) &\geq m - \epsilon \\ &\cdots\cdots \\ k(s_{t-1}, s_t) + V(s_t) - V(s_{t-1}) &\geq m - \epsilon. \end{aligned} \tag{43}$$

Adding these relations gives

$$\sum_{i=1}^{t} k(s_{i-1}, s_i) + V(s_t) - V(s_0) \geq t(m - \epsilon) \tag{44}$$

and by (40), dividing by t

$$\frac{1}{t} \sum_{i=1}^{t} k(s_{i-1}, s_i) \geq m - \epsilon - \frac{c}{t}, \tag{45}$$

so that

$$\liminf_{t \to \infty} \frac{1}{t} \sum_{i=1}^{t} k(s_{i-1}, s_i) \geq m - \epsilon \tag{46}$$

and as ϵ is arbitrarily small, claim (*i*) is proved.

(*ii*) By (36), for any s,

$$\bar{V}(s) - V(s) \le M \tag{47}$$

which, by (36), gives, for all s and $s' \in \rho(s)$,

$$k(s, s') + V(s') - V(s) \le M. \tag{48}$$

Thus if $s_0, s_1, s_2, \cdots$ is any admissible sequences [i.e., $s_{t+1} \in \rho(s_t)$ for all t], then

$$\begin{array}{c} k(s_0, s_1) + V(s_1) - V(s_0) \le M, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ k(s_{t-1}, s_t) + V(s_t) - V(s_{t-1}) \le M. \end{array} \tag{49}$$

By addition

$$\sum_{i=1}^{t} k(s_{i-1}, s_i) + V(s_t) - V(s_0) \le Mt. \tag{50}$$

Thus by (40), dividing by t

$$\frac{1}{t} \sum_{i=1}^{t} k(s_{i-1}, s_i) \le M + \frac{c}{t} \tag{51}$$

so that

$$\limsup_{t \to \infty} \frac{1}{t} \sum_{i=1}^{t} k(s_{i-1}, s_i) \le M, \tag{52}$$

which was to be proved.

These inequalities have been known at least since the publication of Refs. 3 and 10, although their proof is submerged under unnecessary assumptions. The inequalities are crucial to any computer treatment of the long-run average. In addition, the above theorem proves the sufficiency of Bellman's equation (35), for if this equation holds with a bounded V and some λ, then for this V one obtains $m = M = \lambda$, and thus λ is optimal and achievable.

AUTHOR

Hans S. Witsenhausen, Ph.D. (Electrical Engineering), 1966, The Massachusetts Institute of Technology; AT&T Bell Laboratories, 1966—. Mr. Witsenhausen has been associated with the Université Libre de Bruxelles, with the European Computation Center, Brussels, and with the Princeton Research Division of Electronic Associates, Inc. At MIT he was with the Electronic Systems Laboratory as a Lincoln Laboratory Associate and a Hertz Fellow. At AT&T Bell Laboratories, he is associated with the Mathematics and Statistics Research Center. He was a Senior Fellow at the Imperial College of Science and Technology, London, in 1972, a Visiting Professor at MIT in 1973, and Vinton Hayes Senior Fellow at Harvard University in 1975–76. He has worked on problems of hybrid computation, control theory, optimization, geometric inequalities, and other applied mathematical fields.