Appendix A

(1) Let $i = 1, \ldots, n$ be an index variable (that is, a quantity assuming one and only one value at once) running over a set of elements, $s_i \in S$, and pointing them out uniquely (there is a one to one correspondence between the values of i and n elements of S).

Let $p_1, \ldots, p_n$ be probabilities so that $1 \geqslant p_i \geqslant 0$ and so that $\sum_i p_i = 1$. The uncertainty/information measure over those elements of S indexed by i is

$$H(s) = -\sum_{i=1}^{i=n} p_i \log p_i; \quad \text{for} \quad s_i \in S \qquad (1)$$

which gives the *information*; the negative of this quantity is an *uncertainty*.

If it happens that a set S is a set of alternatives (for example, of initially exclusive and exhaustive events) then i runs over all (not just n) elements.

(2) Let x and y be index variables and let $p(x), p(y)$ stand for general designators of p values associated with their distinct values $x_i, y_j; i = 1, \ldots, n; j = 1, \ldots, m$. Using a shorthand due to Garner (1962).

$$H(x) = -\sum_x p(x) \log p(x) \equiv H(x) = -\sum_{i=1}^{i=n} p(x = x_i) \log p(x = x_i)$$

or

$$H(y) = -\sum_y p(y) \log p(y) \equiv H(y) = -\sum_{j=1}^{j=m} p(y = y_j) \log p(y = y_j)$$

These quantities are maximised if the p values are all equal and minimised if one is unity, and the others (by definition) zero. Thus in the case of x (for y, similarly) the maximum value, for n unchanging equally likely alternatives, is

$$H^*(x) = -\sum \frac{1}{n} \log \frac{1}{n} = \log n$$

Further $\quad H^*(x) \geqslant H(x) \geqslant 0$

It is certainly true that the values of x and y circumscribe a possible universe; it may be the universe U_α, U_β; in which case these quantities are identical with $H(\alpha)$ and $H(\beta)$. But this need not always be so; for example, if x and y are state variables of a system, then X (the set of values of x) and Y (the set of values of y)

are regarded as indices over the same universe U_α. Either interpretation is legitimate; in the latter case, however, the observer is explicitly talking about relation $R \subset X \times Y \sqsubseteq U_\alpha$ which may be due to undetermined or other than deterministic connections. Thus, either by juxtaposing two distinct universes U_α, U_β or by considering X and Y as alternative sets over the same universe, it is possible to assign probability numbers to joint events designated by pairs $\langle x,y \rangle \in X \times Y$; namely, the joint event probabilities $p(x,y)$. The joint uncertainty/information index, relevant to the selection of $\langle x,y \rangle$ in U_α is

$$H_\alpha(x,y) = -\sum_x \sum_y p(x,y) \log p(x,y)$$

and if x, y index separate universes U_α, U_β then this is identical with $H(\alpha,\beta)$. Any correlation between occurrences indexed by x and occurrences indexed by y reduces the value of this quantity. Its maximum value occurs in the absence of correlation and is

$$H^*_\alpha(x,y) = H_\alpha(x) + H_\alpha(y)$$

$$\text{or} \quad H^*(\alpha,\beta) = H(\alpha) + H(\beta) = H^*(U_\alpha \cup U_\beta)$$

The degree of coupling is obtained as a difference.

$$T_\alpha(x,y) = H^*_\alpha(x,y) - H_\alpha(x,y) = H_\alpha(x) + H_\alpha(y) - H_\alpha(x,y)$$

or

$$T(\alpha,\beta) = H(\alpha) + H(\beta) - H(\alpha,\beta) \qquad (2)$$

(3) The former interpretation, $T_\alpha(x, y)$, is best regarded as a *coupling* and the latter, $T(\alpha, \beta)$, as a *transmission*. Though the forms are identical, 'coupling' exists between features of a *system*; regarded on other grounds, as coherent or unitary. 'Transmission' takes place between systems regarded, on other grounds, as distinct. In either case, the term in question is positive or zero.

$$H^*_\alpha(x,y) \geqslant T_\alpha(x,y) \geqslant 0$$

$$H^*(\alpha,\beta) \geqslant T(\alpha,\beta) \geqslant 0$$

If the value of $T_\alpha(x,y)$ is greater than zero then x and y are *not* independent (as supposed, in speaking of an *existent* relation between them); if $T(\alpha,\beta) > 0$ then U_α and U_β are *not* independent. Conversely, if $T_\alpha(x,y) = 0$ or $T(\alpha,\beta) = 0$ then, up to this moment in an observation, no dependency has been demonstrated.

The *redundancy* is simply the degree of 'internal transmission' between possible universes regarded as 'the same' (for example, state variables or states of the same system) and is a measure of *constraint*.

$$\text{Redundancy} \triangleq T(x,y) = H^*(x,y) - H(x,y)$$

or if x, y are unique indices on 'the same' universe then

$$H^*_\alpha = H_\alpha(x) + H_\alpha(y)$$

$$H_\alpha = H_\alpha(x,y)$$

$$\text{Redundancy} \triangleq H^*_\alpha - H_\alpha = T(\alpha, \alpha) \text{ (not } T(x,x) \text{ or } T(y,y)\text{).}$$

which is usually normalised to yield the specific measure

$$Z = \frac{H^*_\alpha - H_\alpha}{H_\alpha} = 1 - \frac{H_\alpha^*}{H_\alpha} \tag{3}$$

(Until further notice, the subscripts α, β, are omitted.)

(4) Conditional uncertainty/information/indices are derived from conditional probabilities of the form $p(y|x = x_i)$ meaning 'the probability of each value of y (stated separately for all m values) if one specific value of $x = x_i$ is given'. The selective uncertainty/information, given $x_i = x$, is

$$H(y|x = x_i) = -\sum_y p(y|x = x_i) \log p(y|x = x_i)$$

and the general conditional uncertainty/information is obtained as a weighted mean.

$$H(y|x) = -\sum_x p(x)\, H(y|x = x_i)$$
$$= -\sum_x p(x) \sum_y p(y|x) \log p(y|x)$$

The converse term $H(x|y)$ is obtained in the same manner and the following identities are demonstrable

$$H(x,y) = H(x) + H(y|x) = H(y) + H(x|y)$$
$$T(x,y) = T(y,x) = H(y) - H(y|x) = H(x) - H(x|y)$$

For more than two variables (or universes) the total constraint is more complex. It is designated $T^*(v,x,y)$ for variable v, x, y and is specified for any finite number of variables.

$$T^*(v,x,y) = H^*(v,x,y) - H(v,x,y)$$
$$= H(v) + H(x) + H(y) - H(v,x,y)$$

The specific couplings (transmissions) are terms such as

$$T(v, \langle x,y \rangle)$$

representing the coupling between values of v and pairs of values $\langle x,y \rangle$ of x and y. On decomposing the total constraint pairwise (recalling that $T(x,y) = T(y,x)$) the following equalities are obtained:

$$T^*(v,x,y) = T(v, \langle x,y \rangle) + T(x,y)$$
$$= T(x, \langle v,y \rangle) + T(v,y)$$
$$= T(y, \langle v,x \rangle) + T(v,x)$$

so that, in summary

$$T^*(v,x,y) = T(v,x) + T(v,y) + T(x,y) + Q(v,x,y)$$

where the T terms are couplings between pairs of variables and $Q(v,x,y)$ is an interaction (which may be positive, negative or zero) representing the trivariate

coupling, that cannot be expressed pair-wise (i.e. without considering the 'paired' coupling between 'one variable' and 'a pair of other variables').

$$Q(v,x,y) = T(v,x) + T(v,y) + T(x,y) - T^*(v,x,y) \tag{4}$$

For an m variable system, the expression has m-tuple, $(m-1)$-tuple, . . . terms. If the variables are indices on universes U_α, U_β, U_γ that are held to be potentially distinct, then $Q(v,x,y) = Q(\alpha,\beta,\gamma)$.

(5) Von Foerster notes that redundancy (Z of equation 3) is one measure of organisation, and that a system is rightly regarded as self organising only if the rate of change of redundancy is positive (that is $dZ/dt > 0$). Considering the statistical quantities H^* and H as variables in a statistical metalinguistic model of the system, differentiate the expression for Z with respect to time, thus obtaining

$$\frac{dZ}{dt} = -\frac{H^*\left(\frac{dH}{dt}\right) - H\left(\frac{dH^*}{dt}\right)}{H^{*2}}$$

For $H^* > 0$ (that is, fully ordered systems are excluded) the condition that $dZ/dt > 0$ is satisfied if and only if

$$H\frac{dH^*}{dt} > H^*\frac{dH}{dt}$$

of which there are two special cases (for one or other variable unchanging) namely

(a) If $\frac{dH^*}{dt} = 0$, then $0 > \frac{dH}{dt}$

(b) If $\frac{dH}{dt} = 0$, then $\frac{dH^*}{dt} > 0$

Case (a) corresponds to adaptation in a fixed universe. For example, in the universe of input/output behaviour (x indexing input, y indexing output) in a fixed field of contingencies or relations U_α, where

$$H^*_\alpha = H^*_\alpha(x) + H^*_\alpha(y)$$
$$H_\alpha = H_\alpha(x,y)$$

The rate of change is approximated by considering as distinct the universes $U_\alpha(t)$, $U_\alpha(t + \Delta t)$. The self-organising condition is secured if, for any Δt, chosen the behavioural indices satisfy

$$H_{\alpha,t}(x,y) > H_{\alpha,t+\Delta t}(x,y)$$

Case (b) is of greater interest, since it entails the idea that U_α does change, $U_\alpha(t) \to U_\beta(t + \Delta t)$ as, for example, by changing the field of attention. The required condition is secured if, for any value of Δt, the maximum uncertainty/information/index values (signifying possible behavioural relations) satisfy

$$H^*_{\beta,t+\Delta t} > H^*_{\alpha,t}$$

For example, this transformation of H^* may be performed by changing the *learnable* relations between x and y. In some cases but not all (see Chapter 5 onwards), it is permissible to express this operation as changing the structure of *one* universe $U = U_\alpha \equiv U_\beta$ and thus transforming only $H^*_{t+\Delta t} > H^*_t$. This is so, for instance, if $U_\alpha \subset U_\beta$. Conversely, however, the transformation always *can* be expressed as an operation upon the universe: $U_\alpha \rightarrow U_\beta$. If H is to remain constant as required by condition (b) the transformation must usually be instantaneous (i.e. very rapid).

The general case is of greater interest and has two paradigms:

(a) For all $\alpha, \beta, t, t + \Delta t$: $U_\alpha \subset U_\beta$ so that $U = U_\beta$ and H^* is altered by reducing the constraint between x and y in U, as a result of which more relations can be learned. It is, as above, *possible* to express the *same* operation as a transformation $\alpha \rightarrow \beta$. Since there is usually no way of metricising the effect of relaxing constraint (though the *sense* of the *change* in H^* is determined) this operation is carried out in a steady-state system (see Chapter 7) *until* $0 > dH/dt$ and *if* dH/dt approaches the value 0.

(b) For some $\alpha, \beta, t, t + \Delta t$, U_α and U_β are not directly comparable, in the sense that an external observer's images of U_α and U_β do not necessarily agree with the *system's* image. In this case (see Chapter 9 onwards) the transformation $\alpha \rightarrow \beta$ that regulates H^* must be under the system's control though it may still be instigated by the same conditions on the rate of change of H.

Appendix B

Consider a test item with two alternative answers $\{a,b\}$. The student wishes to maximise his expected test score and the tester wishes to extract all available information about the student's present state of knowledge. Instead of requiring a simple 'yes—no' response to each alternative answer, the tester can seek more information about the student by asking him to assign probability numbers to the alternatives so as to represent his current state of doubt or degree of belief that a given alternative is the correct answer. In the case of two alternatives, the student is required to assign values $P(a)$ and $P(b)$ such that $P(a) = 1 - P(b)$. Clearly such responses are potentially more informative than simple 'yes—no' responses, since in the 'yes—no' case the student is restricted to only two possible response forms *either* $P(a) = 1$ and $P(b) = 0$ *or* $P(a) = 1$ and $P(b) = 0$.

Given a set of probability numbers as responses, the tester has to compute a score. Assuming the probability numbers *do* represent a student's degree of belief, he may, for example, sum the values the student has assigned to the correct alternatives, perhaps also weighting items for their relative importance. This procedure is analogous to that of counting the number of correct responses in the 'yes—no' case, but suffers from a serious deficiency; the tester cannot validly assume the assigned probability numbers do represent the student's degree of belief, since a student who is a 'good statistician' will realise that when such a scoring scheme is employed, he can maximise his expected score by assigning probability numbers that do *not* represent his degree of belief. This is illustrated by the following example.

If the student's actual degree of belief that alternative '*a*' is the correct answer is 0·75 and he follows instructions, setting $P(a) = 0{\cdot}75$ and $P(b) = 0{\cdot}25$, he can expect a score of 0·75 with probability 0·75 and a score of 0·25 with probability 0·25. This gives an *expected* score of $(0{\cdot}75 \times 0{\cdot}75) + (0{\cdot}25 \times 0{\cdot}25) = 0{\cdot}625$. However, the student can make his expected score larger by assigning a larger probability number to alternative '*a*'. In fact, he maximises his expected score by assigning $P(a) = 1{\cdot}00$. This gives an expected score of $(0{\cdot}75 \times 1{\cdot}00) + (0{\cdot}25 \times 0{\cdot}00) = 0{\cdot}75$. In general, the 'good statistician' can maximise his expected score by assigning a value of 1·00 if his degree of belief that an alternative is correct is greater than 0·50, and by assigning a value of 0·00 if his degree of belief that an alternative is incorrect is less than 0·50. If his degree of belief is 0·50, it makes no difference what value he assigns since all values yield equal expected scores.

What is required is a scoring system where the student's optimal strategy is to assign probability numbers that do reflect his degree of belief. Shuford, Massengill and Albert (1966) derive the necessary and sufficient conditions for a scoring system that satisfies this criterion for the case of two alternatives and show how to create a virtually inexhaustible number of such systems. In addition, they give partial results for the case of more than two alternatives.

An example scoring system for the case of two alternatives is one which specifies:

(1) that the student receives a score of $1 - (1 - P(a))^2$ points if he assigns $P(a)$ as his degree of belief that alternative '*a*' is correct when it is in fact correct, and

(2) that he receives $1 - (P(a))^2$ if he assigns $P(a)$ to alternative '*a*' when it is, in fact, incorrect.

Thus, if as before, the student has a degree of belief of 0·75 that alternative '*a*' is correct and assigns $P(a) = 0{\cdot}75$, his expected score is $0{\cdot}75 \times (1 - (1 - 0{\cdot}75)^2) + 0{\cdot}25(1 - (1 - 0{\cdot}25)^2) = 0{\cdot}811$. Assigning $P(a)$ less than or greater than 0·75 leads to a lower expected score.

The following scheme (mentioned in the text) can be employed for the case of more than two alternatives and has the advantage over other such schemes that the student's score depends only on the value he assigns to the correct alternative.

Given the correct answer from a set of exclusive alternatives is the *i*th alternative and the student's assigned value to the *i*th alternative is P_i, this scheme specifies that:

(1) if P_i is greater than or equal to 0·1, he receives $1 + \log P_i$ points, and
(2) if P_i is less than 0·1 he receives -1 points.

Condition (2) provides a cut off point that gives a point scale running from -1 to $+1$ (rather than $-\infty$ to $+1$). Thus, for extreme values of P_i, some information is lost. In any practical applications, this loss in accuracy is insignificant.

Appendix C

1 Fuzzy Operations (Zadeh)

For a thorough description the reader is referred to Goguen (1968 and Zadeh (1973). The following notes and definitions are chiefly intended to indicate the kind of operation we have in mind and to convince the diffident reader that the concept of a fuzzy operation is not, itself, fuzzy (or vague).

A fuzzy subset X of a universe U (here a set of objects, see footnote on the first page of Chapter 1) is defined by a function $C_X(u)$ or $C_X: U \rightarrow$ (Interval 1, 0) that assigns a 'grade of membership' to each element of U. Those elements of U that are associated with positive values (C) of $C_X(u)$ are called the *support* of X. A fuzzy unit set $\{x\}$ has one support $u \in U$ and a positive grade of membership $C_{\{x\}}(U)$. Using a slash notation to designate grades of membership, i.e. C/u. A non-fuzzy unit set is $1/u$. X itself is the union (written $\underline{UN}$, Zadeh uses the integral to comprehend indefinitely large unions) of certain fuzzy unit sets. That is

$$X \triangleq \underset{u \in \text{support}}{\underline{UN}} (C_x(u)/u)$$

The complement of X is

$$Not\ X \triangleq \underset{u \in U}{\underline{UN}} (1 - C(u)/u)$$

The union of fuzzy sets X, Y, is

$$Or\ (X,Y) \triangleq \underset{u \in U}{(\underline{UN})} [C_X(u)\ \underline{V}\ C_Y(u)]/u$$

where $\underline{V}$ is the maximal operator;

$$r\ \underline{V}\ s = Max\ (r,s) \triangleq \begin{Bmatrix} r \text{ if } r \geqslant s \\ s \text{ if } s > r \end{Bmatrix}$$

The intersection of fuzzy sets X, Y is

$$and\ (X,Y) \triangleq \underset{u \in U}{\underline{UN}} [C_X(u)\ \underline{\wedge}\ C_Y(u)]/u$$

where $\underline{\wedge}$ is the minimal operator

$$r\ \underline{\wedge}\ s = Min\ (r,s) \triangleq \begin{Bmatrix} r \text{ if } s > r \\ s \text{ if } r \geqslant s \end{Bmatrix}.$$

A fuzzy relation, R, is a fuzzy subset of a cartesian product; if R relates X to Y then

$$R \triangleq \underset{x \times y}{\underline{UN}}\, C_R(x,y)/\langle x,y\rangle : x \in X, y \in Y.$$

If R and S are fuzzy relations between X, Y and Y, Z their composition is

$$R \bigcirc S \triangleq \underset{x \times z}{\underline{UN}}\, \mathit{Supremum}\, (C_R(x,y) \wedge C_S(y,z))/\langle x,z\rangle$$

where 'supremum' is used in the usual sense of 'collection of maximal elements' of a relation and $C_R(x,y)$ is a grade of membership function of two or more variables.

One special case of a fuzzy relation is the fuzzy cartesian product of universes U, V with elements u, v namely,

$$X \otimes Y \triangleq \underset{U \times V}{\underline{UN}}\, c_X(u) \wedge C_Y(v)/\langle u,v\rangle$$

Any non fuzzy (standard) conditional: 'If α, then β else η' is equivalent (since 'if . . . then' is an implication $\rightarrow$), to the Boolean expression $(\alpha \rightarrow \beta) \wedge (\alpha^{\sim} \rightarrow \eta)$ where the '$\sim$' stands for Boolean negation. By extension, the *fuzzy conditional* is defined (for the same or different universes, U, V and *fuzzy* sets W, X, Y), as

'If X then Y else W' $\triangleq X \otimes Y$ *or* (*not* $(x) \otimes W$)

which expresses, in executable or intensional form, a fuzzy relation, R, from U to V and
which is generalised to image the (standard) extended conditional form

'If X_1 then Y_1 else if X_2 then Y_2 else . . . if X_n then Y_n'

A *fuzzy algorithm* is a non-fuzzily indexed (i.e. well ordered) sequence of instructions of this form (or the degenerate forms' unconditional imperative') and 'assignments' using, in execution, the major inference rule (for a fuzzy relation R from domain X to codomain Y) which is

$$Y = X \bigcirc R$$

The execution is *numericised* if each supremum is resolved by ranking (using any consistent ranking rule) to select a unique outcome.

2 Finite-Function Machines (Von Foerster)

These comments establish the salient distinction between finite-state machines and finite-function machines; the serious reader is (as usual) referred to the original papers.

The characterising equations of a finite-state machine are

$$u_{\tau+1} = f(u_\tau, z_\tau)$$
$$v_\tau = g(u_\tau, z_\tau)$$

and may be expanded to eliminate all but a finite history of n inputs and an initial internal state or output state. By this construction, it is possible to avoid direct

reference to any internal state function, f. Thus, supposing g has an inverse $z = \phi(u,v)$, it follows that

$$v_{\tau+1} = g(u_{\tau+1}, f(u_\tau, \phi(v_\tau, u_\tau))$$

or that

$$v_\tau = g^1(u_\tau, u_{\tau-1}, v_{\tau-1})$$

but (from the last equation)

$$v_{\tau-1} = g^1(u_{\tau-1}, u_{\tau-2}, v_{\tau-2});$$

which, if substituted in the last equation, gives

$$v_\tau = g^2(u_\tau, u_{\tau-1}, u_{\tau-2}, v_{\tau-2})$$

or, for n stages

$$v_\tau = g^n(u_\tau, u_{\tau-1}, \ldots, u_{\tau-n}, v_{\tau-n})$$

which is a *recursive function* (the behavioural history) containing only one reference to an operation or computation: namely $\langle u_{\tau-n}, v_{\tau-n}\rangle$ (or, equivalently, the internal state at $\tau - n$: namely $z_{\tau-n} = \phi(u_{\tau-n}, v_{\tau-n})$).

Turning to a finite-function machine, its characterising equations are

$$f_{\tau+1} = F(u_\tau, f_\tau)$$
$$g_\tau = G(u_\tau, f_\tau)$$

and are thus isomorphic with the finite-state machine equations under the correspondences given below.

Next internal state function $\longleftrightarrow$ Internal state

Output state function $\longleftrightarrow$ Output state

If we apply the steps which lead to the expression for a finite-state machine, stripped of direct reference to internal states (in terms of a recursive function g^n), we obtain

$$g = G^n((u_\tau, u_{\tau-1}, \ldots, u_{\tau-n}), (g_{\tau-1}, g_{\tau-2}, \ldots, g_{\tau-n}))$$

which is a *recursive functional*.

The important difference is that the recursive functional contains explicit reference to a history of operations (the several $g_{\tau-\tau}$) whereas the recursive function does not (just one $v_{\tau-n}$). This is the salient distinction between a finite-state machine and a finite-function machine. Only the latter contains the computations needed to reconstruct an image of its history.

Appendix D

Specification 1 Model I Simulation

Nodes in the environment. Sum = 256. Nodes are indexed by a number k or by coordinates α, β, equivalent to k (as in Fig. 22). Each node designates a food store.

Food store. The amount of food at node k at trial n is $u_k(n)$, maximum value 128.

Food increment. For $u_k(n)$ less than 128, the amount of food added at trial n is $\Delta u_k(n)$, which is a function of $O_k(n)$, the average occupancy of node k over the previous 8 trials and a parameter c (with values 1, 2, or 4 determined by the experimenter). This function is: $\Delta u_k(n) = c$, if $O_k(n) = 0$ or 1; $\Delta u_k(n) = c(1 + O_k(n))$ if $O_k(n) > 1$.

Occupancy. The number of automata at node k at trial n is Occupancy$_k(n)$. The average value of Occupancy$_k(n)$ over the previous 8 trials is $O_k(n)$.

Motions of an automaton indexed i. Consider automaton i at node $k = \alpha, \beta$, at trial n. As in Fig. 22, it can remain where it is or move into any of the positions $\alpha - 1, \beta + 1; \alpha, \beta + 1; \alpha + 1, \beta + 1; \alpha - 1, \beta; \alpha + 1, \beta; \alpha - 1, \beta - 1; \alpha, \beta - 1; \alpha + 1, \beta - 1$; provided that it is not at the edge of the field. These motions are indexed with a number, ρ, between 0 and 8 (0 designating the act of remaining where it is).

Sensory capabilities. Automaton i senses the value of u at any node to which it can move. Hence it has 9 sensory inputs indexed by ρ.

Stomach content. The content of the stomach of automaton i at the start of trial n is $\theta_i(n)$, maximum value 128. If i is at node k, then $\theta_i(n)$ is incremented by $\Delta\theta_i(n) = \eta \cdot u_k(n) - \theta_i(n)$ if $u_k(n) > \theta_i(n)$. If $\theta_i(n) > u_k(n)$, the value of $\theta_i(n)$ is unchanged. $1 > \eta > 0$.

Age of i. The age of an automaton is the number of trials since it entered the simulation.

Aging and metabolism. For each trial the value of $\theta_i(n)$ is decremented by an amount F(age of i). This function has the following form (though several functions were used in the experiments):

F(age of i)	1	1	1	1	1	1	1	1	2	2	4	4	8	8	16	16	32	64
Age of i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	>16	>24

Decay. Automaton i is deleted from the simulation at trial n if $0 > \theta_i(n)$.

Experience weights. Let $\theta_i|p$ denote an increment of stomach food contingent upon one of the 9 motions indexed by p. Automaton i builds up a vector $G_i(n)$ of 9 weights $g_{ip}(n) = \sum \Delta\theta_i|p / \sum$ all $\Delta\theta_i$, the summation being taken over the entire history of the automaton.

Selecting motions. At trial n, positioned at node k, the automaton i has a sensory input of 9 u values, which are denoted $u_{ip}(n)$ relative to node k. It examines the weighted sensory input, having components $g_{ip}(n) \cdot u_{ip}(n)$, and selects the maximum if it exists. If no maximum exists, it selects an arbitary member of the maximal subset. The automaton moves in whichever direction is associated with the selection it has made. Thus if $g_{i1}(n) \cdot u_{i1}(n)$ is the selected maximum, then the automaton moves in direction 1 (i.e. from node $k \equiv \alpha, \beta$ to node $\alpha - 1, \beta + 1$).

Parameter P_i. The value of $P_i(n)$ is the average number of other automata that have occupied the same node as automaton i over the previous 8 trials ($P_i(n)$ is an external control variable).

Reproduction. If automaton i and automaton j occupy the same node k at trial n and if $\theta_i(n) > \mu$ and $\theta_j(n) > \mu$ then an automaton l is created at node k. An amount of food γ is subtracted from the stomach of i and the stomach of j and l is credited with an initial quantity of food, 2γ. Normally $\gamma = \frac{1}{4}\mu$ and $\mu = 64$.

Inheritance. The initial value of $G_l(n)$ is $\frac{1}{2}(G_i(n) + G_j(n))$. The initial value of $P_l(n)$ is $\frac{1}{2}(P_i(n) + P_j(n))$.

Processing. The automata are processed at trial n according to the magnitude of the $\theta_i(n)$. If several $\theta_i(n)$ are of the same value, the automata concerned are processed in an arbitrarily chosen order.

Density control mechanism. The reproductive parameter $\mu = P_i(n)/50$ for automaton i at trial n.

Parallel simulation. Since $u_k(n)$ is modified while an automaton at node k is processed, it is necessary to simulate parallel processing when determining the motion of an automaton. For this purpose a COPY ARRAY is made of the values of $u(n)$ for all 256 nodes at the beginning of trial n. From the COPY ARRAY the program (when processing node k) derives a SENSORY FIELD (k, n), which consists in the 8 values of u that belong to the 8 nodes accessible to an automaton located at node k. These values (rather than the prevailing values of $u(n)$) are used in determining the motion of an automaton.

Storage allocation. A node is constructed as a set of k-indexed words specifying $u_k(n)$, Occupancy$_k(n)$, and the list of k occupants at trial n. It has locations for the temporary storage of SENSORY FIELD and certain statistical data. Automaton i is a set of i-indexed words that specify $\theta_i(n)$, $G_i(n)$, $P_i(n)$, age of i, $\rho_i(n)$, and the name of the location automatons at trial n.

Specification 2 Model II Simulation

Nodes in the environment. Number and connectivity as in Specification 1.

Food store and incrementation.

$$u_k(n) = C \cdot \text{Occupancy}_k(n-4) + C \cdot \text{Occupancy}_k(n-3) + C \cdot \text{Occupancy}_k(n-2) + C \cdot \text{Occupancy}_k(n-1).$$

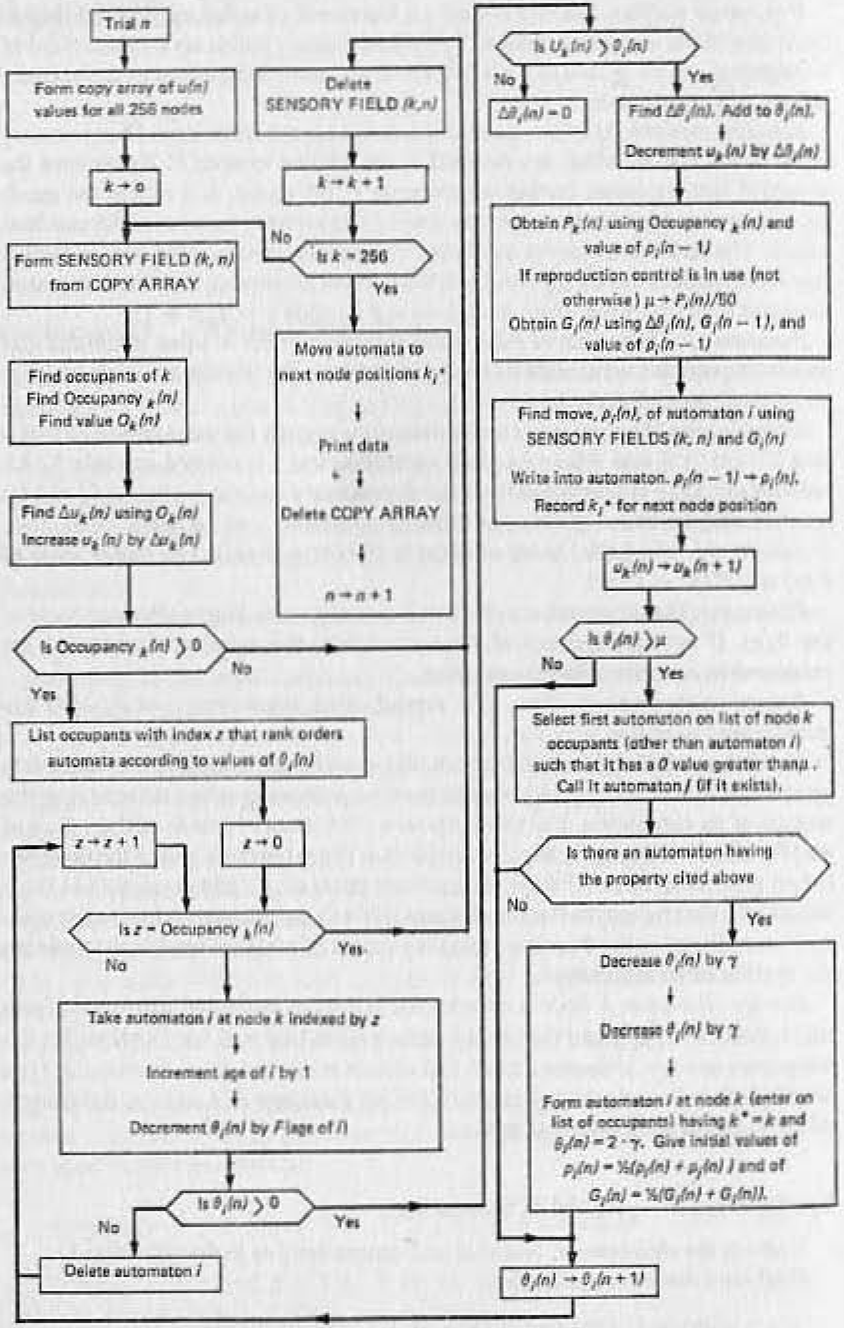

Figure 107 Flow-chart for Model I Simulation.

The automata do not 'eat' separately. The available food is equally divided between the occupants of a node and an automaton receives the YIELD of a node.

Occupancy. As in Specification 1; the average occupancy, however, is computed over only 4 trials. (For details, consult the flow chart of Figs. 107 and 108.)

Types of automata. There are 4 types of automata, denoted $-$, $|$, $+$, $\times$, and their possible motions are shown in Fig. 22.

Sensory capabilities. An automaton can sense the value of u at any node to which it can move. Since different 'types' of automata have different permitted motions, they also have different sensory fields.

Stomach content. Similar to Specification 1. As in Specification 1, however, the value of $\Delta\theta_i(n)$ is the same for all automata residing at node k at trial n. This value is the amount of food at node k divided by the number of automata residing at this node.

Age of i. As in Specification 1.

Aging and metabolism. As in Specification 1, except that the function F(age of i) can be differently defined for each type of automaton.

Decay. As in Specification 1.

Experience weights. These do not exist in Model II simulations.

Selecting motions. As in Specification 1, except that automaton i has no experience weights. Hence, at trial n, it selects a direction of motion associated with the maximum among its sensory inputs (if no maximum exists, it selects a motion randomly from among the maximal subsets).

Parameter P_i. This parameter does not exist in Model II simulations.

Reproductions. The conditions for reproduction are those described in Specification 1, but the results of reproduction depend on the types of automata that encounter one another and satisfy these conditions at the trial concerned. A pair of alternative reproduction rules, R_1 and R_2, are defined, and one of these is used in a particular experiment (unless stated otherwise, R_1 is employed). Using o to denote a combination of automata, p to denote the probability of a particular outcome, and $\Rightarrow$ to denote the progeny produced as an outcome, R_1 and R_2 are defined as follows.

$$
\begin{aligned}
R_1 = {} & -o- \Rightarrow - \\
& |o| \Rightarrow | \\
& -o| = |o- \Rightarrow p = \tfrac{1}{2}(+), p = \tfrac{1}{2}(\times) \\
& +o+ \Rightarrow + \\
& \times o \times \Rightarrow \times \\
& +o\times = \times o+ \Rightarrow p = \tfrac{1}{2}(+), p = \tfrac{1}{2}(\times) \\
& +o- = -o+ \Rightarrow \text{Infertile} \\
& \times o- = -o\times \Rightarrow \text{Infertile} \\
& +o| = |o+ \Rightarrow \text{Infertile} \\
& \times o| = |o\times \Rightarrow \text{Infertile}
\end{aligned}
$$

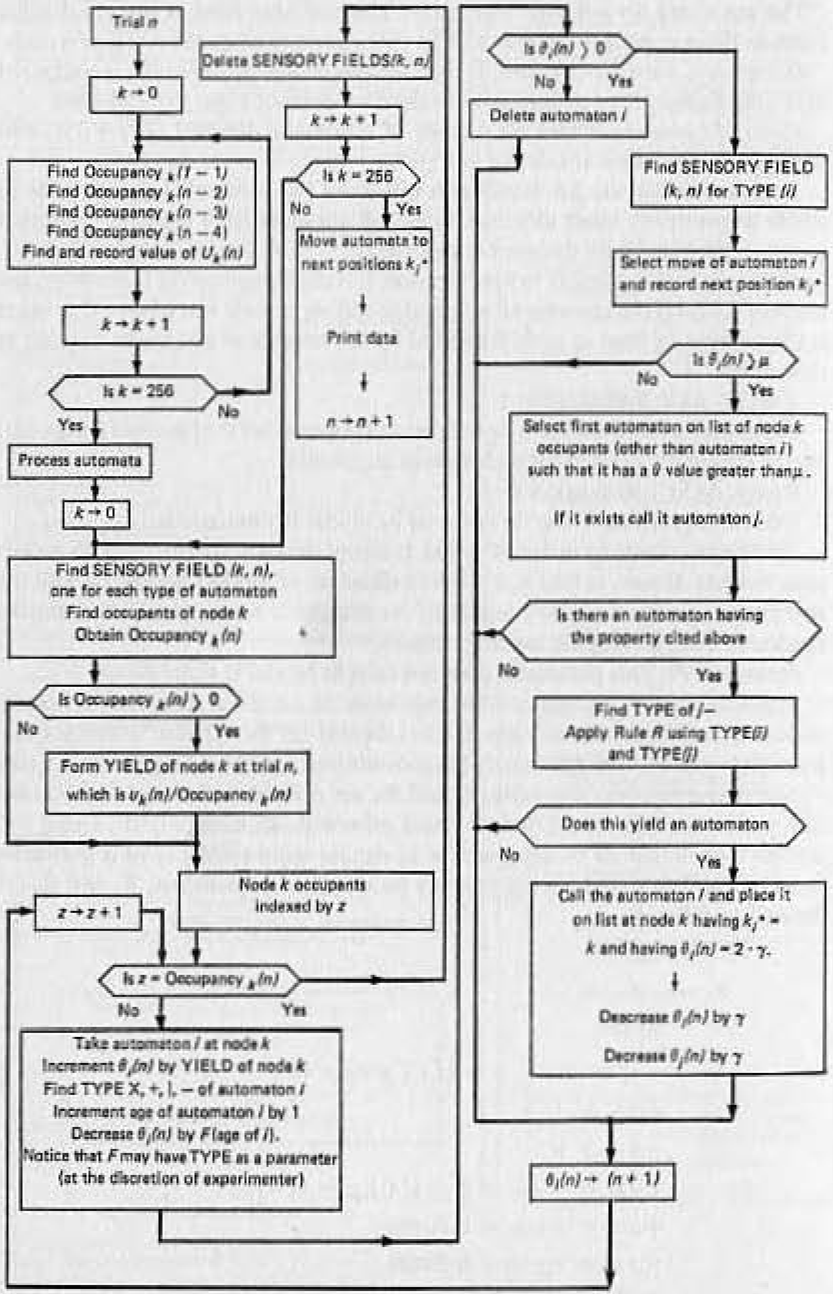

Figure 108 Flow-chart for Model II Simulation.

$$
\begin{aligned}
R_2 = -o- &\Rightarrow - \\
|o| &\Rightarrow | \\
-o| = |o- &\Rightarrow p(+) = \tfrac{1}{2}, p(\times) = \tfrac{1}{2} \\
+o+ &\Rightarrow p(+) = \tfrac{1}{2}, p(\text{Infertile}) = \tfrac{1}{2} \\
\times o \times &\Rightarrow p(\times) = \tfrac{1}{2}, p(\text{Infertile}) = \tfrac{1}{2} \\
+o\times = \times o+ &\Rightarrow p = \tfrac{1}{3}(+), \\
&\quad p(\times) = \tfrac{1}{3}, p(\text{Infertile}) = \tfrac{1}{3} \\
+o- = -o+ &\Rightarrow \text{Infertile} \\
\times o- = -o\times &\Rightarrow \text{Infertile} \\
+o| = |o+ &\Rightarrow \text{Infertile} \\
\times o| = |o\times &\Rightarrow \text{Infertile.}
\end{aligned}
$$

Inheritance. Apart from the type determination entailed by R_1 and R_2, there is no carryover from a given generation to the next generation. Indeed, there are no properties G or P to inherit.

Processing. The automata are processed in an arbitrarily selected order.

Parallel simulation. Since $u_k(n)$ is not modified in the course of processing node k, the COPY ARRAY is unnecessary and the current values of $u(n)$ are computed by the program at the beginning of a trial. SENSORY FIELD is derived from this collection of values, only there are 4 different SENSORY FIELDS, one for each type of automaton, namely $\times$, $+$, $|$, $-$.

Storage allocation. A node is a set of k-indexed words specifying $u_k(n)$ and the 4 previous occupancies used in computing the value of $u_k(n)$. An automaton is a set of l-indexed words that specify $\theta_l(n)$, age of l, and the location, at trial n, of the automaton.

Glossary

The items are arranged in an order conducive to rational development. The list is short enough to be scanned readily and the looked-for-item discovered without difficulty.

Set S A collection of unordered unitary elements. Membership in the collection is written '$\in$' so that $s \in S$ means 'element s is a member of set S'.

Subset X a 'portion of a set' (written, $X \subset S$) 'either a portion of or all of, a set' (written, $X \sqsubseteq S$). If the elements of subsets are specified explicitly the elements in question are enclosed in brackets (curly brackets, to distinguish them from *ordered* sets as below). For example, if $S = \{a,b,c,\}$ then some subsets are $X = \{a,b\}$ and $Y = \{a,c\}$.

Power Set The set of all subsets of a set (usually, including the *null* or *empty* set). For example, if S has elements $\{a,b,c\}$ then the Power set of S is $\{a\}$, $\{b\}$, $\{c\}$, $\{a,b\}$, $\{a,c\}$, $\{b,c\}$, $\{a,b,c\}$, *Null*.

Intersection, Union If X and Y are subsets of S then $X \cap Y$ is the *intersection* of X and Y (the subset of elements of S that belong to X *and* to Y) whilst $X \cup Y$ is the *union* of X with Y (the subset of elements of S belonging *either* to X *or* to Y *or both*).

Complement with respect to S The complement of X is $S - X$; the elements of S that do not belong to X.

Disjoint Subsets of S are subsets, X, Y ($X \subset S$, $Y \subset S$) such that $X \cap Y$ is void. Hence, $X \sqsubseteq S = Y$ and $Y \sqsubseteq S = X$.

Property That which is common to the members of a set (for example, if S is the set of things, X of red things, then X is the property 'redness').

Cartesian Product for two or more sets, S and T, their Cartesian product, $S \times T$, is the collection of all ordered pairs $\langle s,t \rangle$ with one element taken from S, and one from T (for example, all pairs of men and women). Ordered *pairs* are often called 2-tuples; this generalises, for an *ordered set of n elements*, to n-tuples $\langle s_1 \ldots s_n \rangle \in S_1 \times \ldots \times S_n$.

Relation R A relation, given as a listing, is a subset $R \subset S \times T$ (for example, if R is 'married to', then R contains all those men and women who are husband and wife). R may also be represented by a mapping or series of linkages between the elements of S and T. In a monogamous society R is a one to one mapping. Generally, R may consist in any one, one; one, many; many, one; or many, many mapping and it may involve all or only some of the elements (those involved are called R's *domain* and *codomain* or *range*). In general, also $R \subset S_1 \times \ldots \times S_n$, and n is called the adicity of the relation (thus $R \subset S \times T$ is a 2-adic or binary relation).

Function F A kind of relation; either a one to one or many to one mapping with specified domain. If the domain is less than an (understood) set the function is a 'partial' function. Since F is specified as a relation $F \subset S_1 \times \ldots \times S_n$, n is the adicity of the function (or, equivalently, F is a function of n variables).

Variable A named indexing device that assumes but one value at once. For example, a variable u that ranges over the elements s of S; a variable v that ranges over the elements t of T. The values of u denote elements of S, the values of v denote elements of T; in general, a vector $\langle u_1 \ldots u_n \rangle$ of n variables index the elements of $S_1 \times \ldots \times S_n$ and its values stand for n-tuples $\langle s_1 \ldots s_n \rangle \in S_1 \times \ldots \times S_n$.

Classes of Functions Functions are often classified by their domain/range for example, 'Boolean Functions' of a variable assuming values $\{0,1\}$ (or 'True/False') and having possible values $\{0,1\}$; an integer-arithmetic function (of a variable with integer values). Functions are also classified according to the number of variables needed to index their domain. Thus $F(u)$ is a 1 place function.

Assertoric Logic (In contrast to command logics and question logics considered in the text). A language together with rules for usage (syntax) and inference (see below). Statements in the language are given semantic interpretations; commonly as denoting relations and conditions that hold in one or more sets.

Proposition p (in a language) A statement like 's is red' or 's is married to t' and interpreted, if true, as '$s \in X$' or '$\langle s,t \rangle \in R \subset S \times T$'.

Predicate P An adjective. Equisignificantly, the name for a (usually partial) function from a set S to a set of truth values. Commonly, the truth value set is taken to be $\{1,0\}$ (indicating 'true' and 'false') but this is not always the case (see, for example, the Fuzzy sets in Appendix C). If u is a variable indicating members, s, of a set, S, then $P(u) = True$, is a *one place* predicate statement, meaning that the pointed at element has the *property* named by P. If P is used to name $X \subset S$ then X is (as before) a property and $P(u)$ is true for all elements S (pointed at by u) that are members of X. A 2-place predicate $Q(u,v)$ stands for a 2-adic relation. $Q(u,v)$ is true when u and v point at elements, $\langle s_0,t_0\rangle$, of S and T such that $\langle s_0,t_0\rangle$ is a married couple). In general, there are n-place predicates $Q(u_1 \ldots u_n)$.

Conjunction For propositions p,q written ,'$p \wedge q$' meaning that 'p *and* q' is true if p and q are both true.

Disjunction (unqualified and inclusive) written '$p \vee q$' meaning that '$p \vee q$' is true if either p or q or both are true.

Negation The proposition Not-p. If p asserts that $s \in X$ then Not-p asserts that $s \in S - X$, the complement of X in S.

Exclusive Disjunction For propositions, p, q, r, the exclusive disjunction written '$p \veebar q \veebar r$' means, if true, that one and only one of p, q, r is true. Consequently, exclusive disjunctions are often known as alternative sets, and are used for example, to represent alternative answers to a question.

Implication $p \rightarrow q$, means that *if* p is true, *then* q is true.

Equivalence $p \rightarrow q$ and $q \rightarrow p$; written $p \equiv q$. For example, $p \rightarrow q \equiv$ Not $p \vee q$.

Inference Typical inference rules are the primary rule (Modus Ponens of classical logic); if $p \rightarrow q$ and 'p' is true then it is legitimate to infer that q is true; and Modus Tollens (If '$p \rightarrow q$' is true and q is false then it is permissible to infer the falsity of p). Inference Rules figure as metastatements about the language employed to accommodate the logical system and are necessarily augmented by other (somewhat obvious) metatheoretical manipulative permissions; for example, substitution and replacement and logical truism or tautology (if q is derived by substitution and replacement from p then the statement $p \rightarrow q$ is a tautology).

If then else conditional imperative 'If A then B, else C' (where A, B, C are expressions in an action or programming language) means, when the label attached to this statement is encountered:
'If condition A is satisfied, bring about (do, secure or evaluate) B; if not, bring about C' (and, having done so, proceed to the next label).
The execution of a conditional imperative satisfies or brings about a function or a relation (it invariably *does* so, though other effects may be more obtrusive). So does the execution of an algorithm or program.

Assignment Statement (written $u \Rightarrow v$ or $u \Rightarrow 5$) meaning 'give u the value of v' or 'give u the value of 5'.

Program An (ordered) set of assignment statements, and conditional imperative statements, which is open to execution, as a process with a well specified beginning and ending.

Recursive Definition A non viciously circular definition, such as the following definition of 'oddments'.

Base X is an *oddment*; Y is an *oddment*.

Recursive Part An *oddment* placed *on* an *oddment*, is an *oddment*. For example, X on Y, or X *on* (X *on* Y) or Y *on* X *on* . . . (X *on* Y).

Exclusion or uniqueness Requirement Nothing other than these is an oddment.

It is assumed that whoever, or whatever, employs this definition is able to recognise X, Y and to perform the 'operation' *on* and to recognise the result of applying *on*.

Recursively or constructively specified functions

Base Recursive Functions are simple functions, namely, *Constant valued* functions (like $f(x)$ = *Constant*, for all values of x) *or simple order* functions (like the successor function, $f(x) = x + 1$ and functions that search an index or move an object in a constant manner; for example $f(x)$ = (*next above* x) or *Identity Functions* mapping a (possibly complex) element into a property or itself; thus $f(x) = x$, or $f(\langle a,b,c\rangle) = b$.

Recursive Part If composition, written 'o', means doing one thing after another then any composition of simple functions is a recursive function; for example, $f\ o\ g$, or $f\ o\ f\ o\ g$.

Any class of functions of which the member functions are orderable (by a *simple order* function) and are generated from some ancestor function by a common operation, is a recursive function.

Any functions of m variables generated by substitution, with constant value assignment, from a recursive function of $m + 1$ variables is a recursive function.

Exclusion Rule No other functions are recursive functions.

Compiler A program for translating programs, expressed conveniently in a 'high level' and humanly understandable (but, all the same 'formal') language such as ALGOL, or FORTRAN or PLANNER, into a machine code that is interpreted at a mechanical level (and is fussy enough to defeat human patience). No detailed knowledge of high level or machine languages is required; the interested reader may refer to B. Higman's *A Comparative Study of Programming Languages* (MacDonald, 1967).

Miscellaneous Mathematical terms A graph is a structure consisting of *nodes* (drawn as points) and connecting *arcs* that join some or all of the nodes. In a *directed graph* the arcs are directional and often called *edges*. Both nodes and edges are labelled (for example, nodes may represent members of a group; edges, paths of communication between the members). Formally a directed graph is a binary relation on a set of nodes. Though graphs are often invoked, familiarity with them is not essential. The interested reader should consult C. Berge's *The Theory of Graphs* (Methuen, 1962) and Busaker and Saaty's *Finite Graphs and Networks* (McGraw-Hill, 1965). Occasionally we refer to *Games* with the meaning given in *game theory* or *Decision Theory*. Again, little background knowledge is necessary, but readers intrigued by Game Theory will find Luce and Raiffa's *Games and Decisions* (Wiley, 1957) still the most readable and informative reference.

Miscellaneous Statistical Terms Used in discussing experimental results The following comments are superficial but easily augmented by reference to any elementary statistics text book. The *mean* of a set of observations is their average (arithmetic mean) value. The *variance* is an index of scatter about the mean (usually the *standard deviation*, abbreviated as SD is cited; this quantity is the square root of the variance). A significance level (cited as a probability or a percentage) is attached to any serious comment or finding; the usual values are 1% and 0·1% (or $p = 0{\cdot}01$ and $p = 0{\cdot}001$). Thus an assertion that any numerical quantity is significant at the 0·1% level (or $0{\cdot}001 \geqslant p$) means that the quantity might have assumed this value haphazardly or randomly with 1 chance in 1000 Statistical tests are employed in stating hypotheses about absolute values, *differences* between mean values, and *trends* (for example, 'the difference is significant at the 0·1% level'). *Correlation* coefficients measure the undirected (other than random) association between quantities; once again, correlations are quoted together with a properly calculated level of significance.

References

ACKOFF, R. (1973). Science in the systems age. *Operations Research*, May/June, 661–71.

ALEKSANDER, I. (1973). Random logic nets: stability and adaption. *Int. J. for Man-Machine Studies*, **5**(1): 115–32.

AMAREL, S. (1969). Representation of problems and goal directed procedures in computers. Bulletin ASC 1 No. 2, 8–35.

ANGEL, E. S. and G. A. BEKEY (1968). Adaptive finite models of manual control systems. *I.E.E.E. Trans. on Man-Machine Systems*, March 1968: 15–21.

APTER, M. (1966). *Cybernetics and Development*. Pergamon Press.

ARBIB, M. A. (1964). *Brains, Machines and Mathematics*. McGraw-Hill.

ASHBY, W. R. (1962). The self reproducing system. In: *Aspects of the Theory of Artificial Intelligence*. C. E. Muses (ed). Plenum Press.

ASHBY, W. R. (1964a). Constraint analysis. In: *General Systems Yearbook*, Vol. 9.

ASHBY, W. R. (1964b). The set theory of mechanism and homeostasis. In: *General Systems Yearbook*, Vol. 9. Reprinted in: *Automaton Theory and Learning Systems*, D. J. Stewert (ed), 1967: 23–51. Academic Press.

ASHBY, W. R. (1964c). *An Introduction to Cybernetics*. 2nd edn. Chapman and Hall.

ASHBY, W. R. (1968). Some consequences of Bremmerman's limit for information processing systems, pp. 69–76. In: *Cybernetic Problems in Bionics*. H. C. Oestricher and D. R. Moore (eds). Gordon and Breach.

ASHBY, W. R. (1969). Tables of identities governing information flows within large systems. *Commun. ASC*. **1**: 2.

ASHBY, W. R. (1970). Information flows within co-ordinated systems, pp. 57–64. In: *Progress of Cybernetics*. J. Rose (ed). Gordon and Breach.

ATKINSON, R. C. (1969). Models for memory. In: *Programmed Learning Research*. F. de Brisson and M. de Montmollin (eds). Dunod.

ATKINSON, R. C. and R. M. SHIFFRIN (1967). *Human Memory: A Proposed System and Its Control Processes*. Technical Report, 110. Stanford University.

BACH, E. (1964). *An Introduction to Transormational Grammars*. Holt, Rinehart and Winston.

BAKER, J. (1969). The uncertain student and the understanding computer. In: *Programmed Learning Research*. Proceedings of the NATO Symposium on Major Trends in Programmed Learning Research, Nice, 1968. Dunod.

BANERJI, R. B. (1970). *Theory of Problem Solving*. Elsevier, New York.

BANERJI, R. B. (1971). Some linguistic and statistical problems in pattern recognition, pp. 409–19. In: *Pattern Recognition*. Vol. 3. Pergamon Press.

BANNISTER, D. and J. MAIR (1968). *The Evaluation of Personal Constructs*. Academic Press.

BAR-HILLEL, Y. (1964). *Language and Information*. Addison Wesley.

BARICELLI, M. (1962). Numerical testing of evolution theories. *Acta Biotheoretica*, **16**: 69–126.

BARRALT-TORRIJOS, J. (1973). *A Programmatic Interpretation of Combinatory Logics*, (thesis), School of Information and Computer Science, Georgia Institute of Technology.

BARRALT-TORRIJOS, J. and L. CHIARAVIGLIO (1971). *On the Combinatory Definability of Hardware*. Report NSR-GN-655 (Feb.), School of Information Science, Georgia Institute of Technology.

BARRON, R. L. (1968). Self organising and learning control system. In: *Cybernetic Problems in Bionics*. H. L. Oestricher and D. R. Moore (eds). Gordon and Breach.

BATESON, C. (1972). *Our Own Metaphor*. Knopf.

BATESON, G. (1972). *Steps to an Ecology of Mind*. Intertext Books.

BATESON, G., D. JACKSON, J. HALEY and J. WEAKLAND (1956). Towards a theory of schizophrenia. *Behav. Sci.*, **1**: 251–69.

BEER, S. (1966). *Decision and Control*. Wiley.

BEER, S. (1973). *Fanfare for Effective Freedom*. 3rd Richard Goodman Memorial Lecture, Brighton Polytechnic, Feb 14.

BELBIN, R. M. (1969). *The Discovery Method in Training*. Training Information Paper H5, Dept. of Employment, HMSO.

BELL, G. I. (1973). Predator-prey equations simulating an immune response. *Math. Biosci.*, **16**(3–4): 291–315.

BELLMAN, R. (1957). *Dynamic Programming*. Princeton University Press.

BEURLE, R. L. (1954). *Activity in a Block of Cells Capable of Regenerating Impulses*. R.R.E. Memoranda, 1042 and 1043.

BEURLE, R. L. (1959). Storage and manipulation of information in the brain. *J. Inst. Elect. Engrs.*, **5** (new series).

BONNER, J. T. (1958). *The Evolution of Development*. Cambridge University Press.

BREMMERMAN, H. J. (1962). Optimization through evolution and recombination, pp. 93–106. In: *Self-Organising Systems*. M. Yovits, G. Jacobi and S. Goldstein (eds). Spartan Press.

BRICKER, P. D. (1955). The identification of redundant stimulus patterns. *J. Exp. Psych.*, **49**: 73–81.

BRILLOUIN, L. (1956). *Science and Information Theory*. Academic Press.

BRILLOUIN, L. (1964). *Scientific Uncertainty and Information*. Academic Press.

BROADBENT, D. E. (1971). *Decision and Stress*, Academic Press.

BRUNER, J., J. GOODNOW and G. AUSTIN (1965). *A Study of Thinking*. Wiley.

BURKS, A. W. (ed) (1970). *Essays on Cellular Automata*. University of Illinois Press.

CARNAP, R. (1950). *Logical Foundations of Probability*. University of Chicago Press.

CARNAP, R. and Y. BAR-HILLEL (1953). *An Outline of a Theory of Semantic Information*. MIT Research Laboratory, Electronics Rech. Rep. 247.

CARTLEDGE, J. W. and G. C. REJAC (1970). Simulation of the Tsembega ritual cycle. Term thesis. I.C. 5625, School of Information Science, Georgia Institute of Technology.

CHIARAVIGLIO, L. (1071). *Contribution to Summis Symposium*. Georgia Institute of Technology.

CHIARAVIGLIO, L., H. R. HAGEN and R. C. ROEHKASSE (1972). Abstract digital computers. In: *Annual Progress Report*. School of Information Science, Georgia Institute of Technology.

CHIARAVIGLIO, L., J. POORE and J. BARRALT-TORRIJOS (1971). *On the Combinatory Definability of Software*. Report NSR-GN-655 (March), School of Information Science, Georgia Institute of Technology.

CODD, E. F. (1968). *Cellular Automata*. Academic Press.

CONANT, R. and W. R. ASHBY (1970). Every good regulator of a system must be a model of that system. *Int. J. System Sci.*, **1**(2): 81–97.

GRAIK, K. (1966). In: *The Nature of Psychology*. S. Sherwood (ed). Cambridge University Press.

CROSSMAN, E. R. F. W. (1953). Entropy and choice time: the effect of frequency unbalance on choice response. *Quart. J. Exp. Psych.*, **5**: 41–51.

CROWDER, N. A. (1960). Automatic tutoring by intrinsic programming. In: *Teaching Machines and Programmed Learning*. I. A. A. Lumsdaine and R. Glaser (eds). Nat. Educ. Asson U.S.

DAVIS, I. K. and J. HARTLEY (1972). *Contributions to an Educational Technology*. Butterworth.

DAVIS, M. (1958). *Computability and Unsolvability*. McGraw-Hill.

DEXTER, M. (1972). Unpublished Ph.D. thesis. School of Information Sciences, Georgia Institute of Technology.

DIENES, D. (1972). *Petri Nets*. I.E.T. Memorandum, Open University.

ELMAGHRABY, S. E. (1970). Some network models in management science. In: *The Series Lecture Notes in Operations Research and Mathematical Systems*, 29. Springer-Verlag.

ELSHOUT, T. and M. ELSHOUT (1969). The programmed instruction of problem solving. In: *Programmed Learning Research*. F. de Brisson and M. Montmoullin (eds). Dunod.

ERNST, G. W. and A. NEWELL (1969). *G.P.S.: A Case Study in Generality and Problem Solving*. Academic Press.

EVANS, C. R., T. WILKINSON and P. LAMBERT (1967). *Auditory 'Stabilized Images'; Fragmentation and Distortion of Words with Repeated Presentation*. Report Auto 30, Autonomics Division, National Physical Laboratory, Teddington.

FEIGENBAUM, E. A. and G. FELDMAN (1963). *Computers and Thought*. McGraw-Hill.

FESTINGER, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press.
FINETTI, BRUNO DE (1962). Does it make sense to speak of good probability appraisers? pp. 357–64. In: *The Scientist Speculates*. I. J. Good (ed). Heinemann.
FITTS, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psych.*, **47**: 381–91.
FOGEL, L. J., A. J. OWENS and M. J. WALSH (1966). *Artificial Intelligence Through Simulated Evolution*. Wiley.
FU, K. S. (1968). *Sequential Methods in Pattern Recognition and Machine Learning*. Academic Press.
GABBAY, D. M. (1969). *A Model Theory for Tense Logic*. Report nos. 1 and 2. Applied Logic Branch. The Hebrew University of Jerusalem.
GAGNE, R. M. (1962). The acquisition of knowledge. *Psych. Rev.*, **69**(4): 355–65.
GAGNE, R. M. (1969). Learning categories and instructional strategies. In: *Programmed Learning Research*. Proceedings of the NATO Symposium on Major Trends in Programmed Learning Research, Nice, 1968. Dunod.
GAINES, B. R. (1968). *Automated Feedback Trainers for Perceptual Motor Skills*. Final Report, Ministry of Defence Contract on Servomechanisms in Operator Training Tasks.
GAINES, B. R. (1969a). Stochastic computers, pp. 66–76. In: *The Encyclopedia of Linguistics Information and Control*. A. R. Meatham and R. A. Hudson (eds). Pergamon Press.
GAINES, B. R. (1969b). Adaptively controlled instruction for a tracking skill. In: *Programmed Learning Research*. Proceedings of the NATO Symposium on Major Trends in Programmed Learning Research, Nice, 1968. Dunod.
GAINES, B. R. (1972). Axioms for adaptive behaviour. *Int. J. Man-Machine Studies*, **4**(2): 169–200.
GALLENTER, E. (ed.) (1959). *Automatic Teaching: The State of the Art*. Wiley.
GILBERT, T. F. (1962). Mathetics, Technology of Education. *J. of Mathetics*, **1**(1): 1–25.
GILSTRAP, L. O. (1971). *Keys to Developing Machines with High Level Artificial Intelligence*. ASMI, New York, April.
GLASER, R. (1962). Some research problems in automated instruction: instructional programming and subject matter structure. In: *Programmed Learning and Computer Based Instruction*. J. E. Coulson (ed). Wiley.
GLUSHKOV, V. M. (1966). *Introduction to Cybernetics*. Academic Press.
GOGUEN, J. A. The Logic of Inexact Concepts. *Synthese* **19**, 325–73.
GREY WALTER, W. (1953). *The Living Brain*. Penguin.
GREY WALTER, W. (1969). Electric signs of expectancy and decision in the human brain. In: *Cybernetic Problems in Bionics*. H. L. Oestreicher and D. R. Moore (eds). Gordon and Breach.
GUILFORD, J. P. (1956). The structure of intellect. *Psych. Bull.*, **53**: 267–93.
HARLOW, H. F. (1959). Learning set and error factor theory. In: *Psychology: a Study of a Science*. Vol. 2. Koch (ed). McGraw-Hill.
HARRAH, D. (1966). *Communication: A Logical Model*. MIT Press.

HARRAH, D. (1973). The logic of questions and its relevance to instructional science. *Instructional Sci.*, **1**(4): 447–68.
HELSON, H. (1964). *Adaptation Level Theory*, Harper and Row.
HESSE, M. (1963). *Models and Analogies in Science*. Sheed and Ward.
HEWITT, C. (1969). PLANNER: a language for proving theorems in robots, pp. 295–301. In: *Proceedings of the International Joint Conference on Artificial Intelligence*. Mitre Corp.
HEWITT, C. (1971) Unpublished Ph.D. thesis, MIT.
HICK, W. E. (1952). On the rate of gain of information. *Quart. J. Exp. Psychol.*, **4**: 11–26.
HOBBS, L. C., D. J. THEIS, J. TRIMBLE, H. TITUS and I. HIGHBERG (eds) (1970). *Parallel Processor Systems, Technologies and Applications*. Spartan Books.
HODGESON, A. and E. BENNET (1967). Structural communication. *Systematics*, **5**(3).
HODGESON, A. (1968). *A Communication Technique for the Future*. Ideas Curriculum Lab. Goldsmiths College, Univ. of London, No. 7.
HOLLAND, J. G. and B. F. SKINNER (1961). *The Analysis of Behaviour*. McGraw-Hill.
HOLT, A. W. (1968). *Final Report for the Information System Theory Project*. Rome Air Development Centre, Contract No. AF30 (602) 4211.
HOLT, A. W. and F. COMMONER (1972). *Events and Conditions*. Information System Theory Project, Applied Data Research Inc., New York, Contract No. DAH–CO4–68–C–0043.
HOMME, L. E. (1966). Contiguity theory: psychological management. *Psych. Rec.*, **16**.
HOWARD, N. (1966a). The theory of metagames. *General Systems Yearbook*, **11**: 167–86.
HOWARD, N. (1966b). The mathematics of metagames. *General Systems Yearbook*, **11**: 187–200.
HOWARD, N. (1973). *Paradoxes of Rationality*. MIT Press.
HUDSON, E. M. (1962). *An Adaptive Tracking Simulator*. I.R.E. Congress on Human Factors in Electronics. Long Beach, California.
HUNT, E. B., J. MARIN and P. J. STONE (1966). *Experiments in Induction*, Academic Press.
HYDEN, H. (1965). Activation of nuclear RNA of neurons and glia in learning, pp. 178–239. In: *The Anatomy of Memory*: 1. D. P. Kimble (ed). Science and Behaviour Books Inc.
IVANHENKO, A. G. and U. LAPPA (1967). *Cybernetic Forecasting Techniques*. Elsevier.
JONCKHEERE, A. R. (1954). A test of significance for the relation between m rankings and k ranked categories. *Brit. J. Math. Stat. Psych.*, **7**: 93–100.
KAGAN, J. (1966). Developmental studies in reflection and analysis, pp. 487–522. In: *Perceptual Development in Children*. A. H. Kidd and J. L. Ricoire (eds). University of London Press.
KAUFFMAN, S. A. (1971). Cellular homeostasis, epigeneis and replication in

randomly aggregated macromolecular systems. *J. Cyb.* (*Trans. Amer. Soc. for Cyb.*), 1 (Jan–March): 71–96.

KELLEY, C. R. and D. J. PROSIN (1968). *Adaptive Performance Measurement*. Final Report on Contract Nonr. 4986(00) Dunlop and Assoc. Inc. Santa Monica, California.

KELLEY, C. R. and E. J. KELLEY (1970). *A Manual for Adaptive Techniques*. Final Report on Contract Nonr. 4986(00) Dunlop and Assocs. Inc. Santa Monica, California.

KELLY, G. A. (1955). *The Psychology of Personal Constructs*. Vols. 1 and 2. Norton.

KERNER, E. H. (1957). Statistical mechanics of interacting biological species. *Bull. Math. Biophys.*, **19**: 121.

KERNER, E. H. (1959). Further considerations of statistical mechanics of biological associations. *Bull. Math. Biophys.*, **21**: 217.

KILMER, W. L., W. S. MCCULLOCH and J. BLUM (1969). A model of the vertebrate central command system. *Int. J. Man-Machine Studies*, **1**(3): 279–310.

KONORSKI, J. (1962). The role of central factors in differentiation. In: *Information Processing in the Central Nervous System*. I.U.S.P. Leiden.

LAING, R. D., S. PHILLIPSON and A. R. LEE (1966). *Interpersonal Perception*. Tavistock.

LANDA, L. V. (1971). The algorithmicization of instruction. Trans. by V. Bennet and edited by F. Kopstein. Educational Technology Publishers (Ms. copyright).

LEWIS, B. N. (1970). *Decision Logic Tables for Algorithms and Logical Trees*. HMSO. CAS Occasional Paper, 12.

LEWIS, B. N., I. S. HORABIN and C. P. CANE (1967). *Flow Charts, Logical Trees and Algorithms for Rules and Regulations*. HMSO. CAS Occasional Paper, 2.

LEWIS, B. N. and G. PASK (1964). The development of communication skills under adaptively controlled conditions. *Programmed Learning*, **1**: 69–88.

LEWIS, B. N. and G. PASK (1965). The theory and practice of adaptive systems, pp. 213–66. In: *Teaching Machines and Programmed Learning*. Vol. 2. *Data and Directions*. R. Glaser (ed). Nat. Educ. Assn. U.S.

LILLY, J. C. (1972). *Programming and Metaprogramming in the Human Biocomputer*. Julian Press.

LILLY, J. C. and D. SHIRLEY (1968). *Perception of Repeated Speech*. I.E.E. NPL Conference on Pattern Recognition. NPL.

LOEFGREN, L. (1968). An axiomatic explanation of complete self-reproduction. *Bull. Math. Biophys.*, **30**: 415–24.

LOEFGREN, L. (1972). Relative explanations of systems. In: *Trends in General Systems Theory*. G. Klir (ed). Wiley.

LORENZ, K. (1952). *King Solomon's Ring*. Methuen.

LOUGUET-HIGGINS, C. and D. MICHIE (1971). A holographic model for memory. *Nature*.

LUCE, R. D., R. T. BUSH and E. R. GALANTER (1963). *Handbook of Mathematical Psychology*. Vols. 1 and 2. Wiley.

LUMSDAINE, A. A. (1963). Assessing the effectiveness of instructional programmes. In: *Teaching Machines and Programmed Learning II*. R. Glaser (ed). Nat. Educ. Assn U.S.

MCCULLOCH, W. S. (1965). *Embodiments of Mind*. MIT Press.

MCCULLOCH, W. S. and W. PITTS (1943). A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.*, **5**: 115–33. Reprinted in *Embodiments of Mind*, W. S. McCulloch, MIT Press, 1965.

MCGILL, W. J. (1963). Multivariate information transmissions, pp. 84–102. In: *Readings in Mathematical Psychology*. R. D. Luce, R. T. Bush and E. R. Galanter. Wiley. Reprinted from *Psychometrika*, **19**: 97–116 (1954).

MANNA, Z. (1970). Correctness of non-deterministic programmes. *Artificial Intell.* **1**: 1.

MARKLE, S. M. (1960). Programmed instruction in English. In: *Teaching Machines and Programmed Learning I*. Nat. Educ. Assn U.S.

MARKOV, A. A. (1961). *Theory of Algorithms*. Works of the Math. Inst. I.M.V.A. Steklov. Vol. 42. Trans. Israel Program for Scientific Translations. Oldbourn Press.

MATHESON, J. E. (1964). *Optimum Teaching Procedures Derived from Mathematical Learning Models*. Report CS/2, Computer Co-ordinated Systems, Institute of Engineering Economic Systems, Stanford University.

MESAROVIC, M. D. (1962). Self-organisational systems. In: *Self-Organising Systems*. M. Yovitz, G. Jacobi and E. Goldstein (eds). Spartan Press.

MESAROVIC, M. D. (1963). *General Systems*. Proceedings of the 2nd IFAC Conference, Basle.

MILLER, G., E. GALANTER and K. H. PRIBRAM (1960). *Plans and the Structure of Behaviour*. Holt.

MINSKY, M. (1967). *Computation: Finite and Infinite Machines*. Prentice-Hall.

MINSKY, M. (ed.) (1968). *Semantic Information Processing*. MIT Press.

MINSKY, M. and S. PAPPERT (1969). *Perceptrons*. MIT Press.

MYHILL, J. (1970). The abstract theory of self-reproduction, pp. 206–18. In: *Essays on Cellular Automata*. A. W. Burks (ed). University of Illinois Press.

NAGEL, E. and J. R. NEUMANN (1959). *Godel's Proof*. Lund, Humphries and Co.

NEWELL, A., J. C. SHAW and H. A. SIMON (1962). The process of creative thinking. In: *Contemporary Approaches to Creative Thinking*. H. E. Gruber, G. Terrell and M. Wertheimer (eds). Atherton Press.

PASK, G. (1957). A teaching machine for radar training. *Automation Progress*, April: 214–17.

PASK, G. (1958). Electronic keyboard teaching machines. *J. Nat. Assn Educ. Commerce*. Reprinted (pp. 336–49). In: *Teaching Machines and Programmed Learning Vol. 1*. R. Glaser and A. Lumsdaine (eds). Nat. Educ. Assn U.S., 1960.

PASK, G. (1959). Physical analogues to the growth of a concept, pp. 877–922. In: *Mechanisation of Thought Processes*. A. Utley (ed). HMSO.

PASK, G. (1960). The growth process in the cybernetic machine, pp. 765–94. In: *Proceedings 2nd Congress International Association of Cybernetics*, Namur, 1958. Gauthier-Villers.

PASK, G. (1961a). *An Approach to Cybernetics*. Hutchinson.

PASK, G. (1961b). A proposed evolutionary model, pp. 229–54. In: *Principles of Self-Organisation*. H. Von Foerster and G. Zopf (eds). Pergamon Press.

PASK, G. (1967). The Control of Learning in Small Subsystems of an educational system. *IEEE Transactions on Human Factors in Electronics*, **1.** HFE. 8 pp. 78–93.

PASK, G. (1963). The use of analogy and parable in cybernetics with emphasis upon analogies for learning and creativity. *Dialectica*, **17**: 167–202.

PASK, G. (1964a). A proposed experimental method for the behavioural sciences, pp. 171–80. In: *Progress in Bio-Cybernetics*. Vol. 1. N. Weiner and J. Schade (eds). Elsevier.

PASK, G. (1964b). Self-organising systems involved in human learning and performance, pp. 247–335. In: *Proceedings 3rd Bionics Symposium, Dayton, Ohio, 1963*. USAF ASD-TDR-63-946, ASTIA.

PASK, G. (1965). Learning machines. (Survey paper at IFAC Conference, Basle, 1963), pp. 393–411. In: *Automatic and Remote Control*. Butterworths.

PASK, G. (1968). A cybernetic model for some kinds of learning and mentation, pp. 531–85. In: *Cybernetic Problems in Bionics*. M. C. Oestreicher and D. R. Moore (eds). Gordon and Breach.

PASK, G. (1969a). The computer simulated development of populations of automata. *Math. Biosci.*, **4**: 101–27.

PASK, G. (1969b). Strategy competence and conversation as determinants of learning. *Programmed Learning*, Oct., 250–67.

PASK, G. (1970a). The meaning of cybernetics in the behavioural sciences, pp. 15–45. In: *Progress of Cybernetics*. Vol. 1. J. Rose (ed). Gordon and Breach. Reprinted in *Cybernetica* (1970), No. 3: 140–59 and No. 4: 240–50 and in *Artoga Communications* (1971): 146–8.

PASK, G. (1970b). Cognitive systems, pp. 349–405. In: *Cognition: A Multiple View*. L. Garvin (ed). Spartan Press.

PASK, G. (1971a). A cybernetic experimental method and its underlying philosophy. *Int. J. Man-Machine Studies*, **3**: 279–337.

PASK, G. (1971b). Interaction between individuals: its stability and style. *Math. Biosci.*, **11**: 59–84.

PASK, G. (1972). A fresh look at cognition and the individual. *Int. J. Man-Machine Studies*, **4**: 211–16.

PASK, G. (1973a). Models for social systems and for their languages. *Instructional Sci.*, **1**(4): 395–445.

PASK, G. (1973b). *Conversation, Cognition and Learning: a Cybernetic Theory and Methodology*. Open University Monograph (abridged form). In press (Elsevier Press, 1974).

PASK, G. and B. N. LEWIS (1964). *Research on the Design of Adaptive Teaching Systems with a Capability for Selecting and Altering Criteria for Adaptation*. Annual Summary Report, No. 4, USAF Contract No. A F 61(052)–402. System Research Ltd.

PASK, G. and B. N. LEWIS (1967). The adaptively controlled instruction of a transformation skill. *Programmed Learning*, April: 74–86.

PASK, G. and B. N. LEWIS (1968). The use of a null-point method to study the acquisition of simple and complex transformation skills. *Brit. J. Math. Stat. Psych.*, **21**(1): 61–83.

PASK, G. and B. N. LEWIS (1969). The self-organisation of a three person task orientated group, pp. 291–311. In: *The Simulation of Human Behaviour*. Proceedings of the NATO Symposium, Paris, 1967. Dunod.

PASK, G., B. N. LEWIS and R. J. FELDMAN (1965). *A Cybernetic Investigation of Learning and Perception*. Final Report, USAF Contract No. AF61 (052)–640, ASTIA. System Research Ltd.

PASK, G. and G. MALLEN (1969). Some recent advances in adaptively controlled teaching systems, pp. 233–39. In: *Proceedings 5th International Congress on Cybernetics*. Namur: Gauthier-Villers.

PASK, G., G. MALLEN, M. ROBINSON, M. DOWSON and B. C. E. SCOTT (1967). *Uni-dimensional Tracking Skill*. Final Scientific Report, USAF Contract No. AF61 (052) 964. System Research Ltd.

PASK, G., G. MALLEN and B. C. E. SCOTT (1968). *Cybernetic Models for the Learning Process*. Final Scientific Report, USAF Contract No. F61 (052)–67C–0010. System Research Ltd.

PASK, G. and B. C. E. SCOTT (1970). *Learning Strategies and Individual Competence*. Final Report, SSRC Project No. HR983/1. System Research Ltd.

PASK, G. and B. C. E. SCOTT (1971). Learning and teaching strategies in a transformation skill. *Brit. J. Math. Stat. Psych.*, **24**: 205–29.

PASK, G. and B. C. E. SCOTT (1972a). Learning strategies and individual competence. *Int. J. Man-Machine Studies*, **4**: 217–53.

PASK, G. and B. C. E. SCOTT (1972b). *Uncertainty Regulation in Learning Applied to Procedures for Teaching Concepts of Probability*. Final Report, SSRC Project HR 1203/1. System Research Ltd.

PASK, G. and B. C. E. SCOTT (1973). CASTE: a system for exhibiting learning strategies and regulating uncertainties. *Int. J. Man-Machine Studies*, **5**: 17–52.

PASK, G., B. C. E. SCOTT and D. KALLIKOURDIS (1973). A theory of conversations and individuals. *Int. J. Man-Machine Studies*, **5**(4): 321–76.

PASK, G., B. C. E. SCOTT, G. MALLEN and M. DAWSON (1969a). *Comparative Study of Adaptively Controlled and Other Methods of Instruction*. Vols. 1 and 2. Final Report, MOD Contract No. A/70/GEN/9725. Systems Research Ltd.

PASK, G., B. C. E. SCOTT, W. LEWIS and M. ELSTOB (1969b). *Research into the Adaptively Controlled Instruction of Compensatory Tracking Skills*. Final Scientific Report, USAF Contract No. F61 052 68 C 0071. System Research Ltd.

PASK, G. and H. VON FOERSTER (1960). A predictive model for self-organising systems. Part 1. *Cybernetica*, **3**(4): 258–300.

PASK, G. and H. VON FOERSTER (1961). A predictive model for self-organising systems. Part 2. *Cybernetica*, **4**(1): 20–55.

PENROSE, L. S. (1962). On living matter and self-replication, pp. 258–72. In: *The Scientist Speculates*. I. J. Good (ed). Heinemann.

PEREIRA, L. (1973). Unpublished Ph.D. thesis. Institute of Cybernetics, Brunel University.

PETRI, C. A. (1965). *Communication with Automata*. Trans. by Clifford F. Greene Jr. A Supplement to Technical Documentary Report, 1, Prepared for Rome Air Development Center Contract AF30(602) = 3324.

PEW, R. C. (1966). Acquisition of hierarchical control over the temporal organisation of a skill. *J. Exp. Psych.*, **71**(5): 764–71.

PHILLIPS, L. (1970). Course notes for Bayesian statistics. Psychology Dept., Brunel University.

PIERCE, J. R. (1962). *Signal, Symbols and Noise*. Hutchinson.

POLYA, G. (1954). *Mathematics and Plausible Reasoning*: Vol. I, *Induction and Analogy in Mathematics*; Vol. II, *Patterns of Plausible Inference*. Princeton University Press.

PRESSEY, S. L. (1960). A machine for automatic teaching of drill material. In: *Teaching Machines and Programmed Learning*. A. A. Lumsdaine and R. Glaser (eds). Nat. Educ. Assn U.S., 1960.

QUILLIAN, M. R. (1969). The teachable language comprehender: a simulation programme and a theory of language. *Communs ACM*, **12**(8): 459.

RAPAPORT, A. (1967). Escape from paradox. *Sci. Amer.* July: 50–6.

RAPPAPORT, R. (1969). *Pigs for the Ancestors*. Yale University Press.

REITMAN, W. R. (1965). *Cognition and Thought*. Wiley.

RESCHER, N. (1966). *The Logic of Commands*. Routledge and Kegan Paul.

RESCHER, N. and A. URQUHART (1971). *Temporal Logic*. Springer-Verlag.

RESINES, S. (1966). *Mathematical Logic*. UNAM Press.

ROBINSON, M. (1973). Unpublished Ph.D. thesis. Institute of Cybernetics, Brunel University.

ROSENBLATT, F. (1961). *Principles of Neurodynamics, Perceptrons and the Theory of Brain Mechanisms*. Cornell Aeronautical Lab. Report No. VG 1196–G–8.

SCANDURA, J. M. (1970). The role of rules in behaviour; toward an operational definition of what (rule) is learned. *Psych. Rev.*, **77**: 516–33.

SCOTT, B. C. E. (1963). Unpublished Ph.D. thesis. Institute of Cybernetics, Brunel University.

SCOTT, B. C. E. and G. PASK (1973). The adaptively controlled instruction of keyboard skills. A series of three papers submitted to *Int. J. Man-Machine Studies*.

SEIGMAN, P. and W. STAPLETON (1971). *A Methodology Utilising Semantic Information Measures (SIM) for Conversational or Dialogue Experiments*. School of Information Science, Georgia Institute of Technology.

SHANNON, C. E. and W. E. WEAVER (1949). *Mathematical Theory of Communication*. University of Illinois Press.

SHIMURA, M. (1973). Fuzzy sets in rank-ordering objects. *J. Math. Anal. and Appl.* (in press).

SHROEDINGER, E. (1944). *What is Life?* Cambridge University Press.

SHUFORD, E. M. (1965). *Cybernetic Testing*. ESD–TR–65–467, Decision Sciences Lab., L. G. Hanscom Field, Beford, Mass.

SHUFORD, E. H., H. MASSENGILL and G. ALBERT (1966). Admissable probability measurement procedures. *Psychometrika*, **31**: 125.

SIMON, H. A. (1966). Thinking by computers and Scientific discovery and the psychology of problem solving. In: *Mind and Cosmos*. R. G. Colodny (ed). University of Pittsburgh Press.

SKINNER, B. F. (1968). *The Technology of Teaching*. Appleton-Century-Crofts.

SKINNER, B. F. (1969). *Contingencies of Reinforcement. A Theoretical Analysis*. Appleton-Century-Crofts.

SMALLWOOD, R. D. (1962). *A Decision Structure for Teaching Machines*. MIT Press.

SOMERHOFF, G. (1969). The Abstract Characterisation of Living Systems. In: *Systems Thinking*. L. Emery (ed). Penguin. (This paper summarises and updates work from 1950 onwards and is usefully compared with articles by other contributors to the book.)

SPENCER-BROWN, G. (1969). *The Laws of Form*. George Allen and Unwin.

STANISLAND, A. C. (1966). *Patterns of Redundancy*. Cambridge University Press.

STAVRINIDES, A. (1973). Unpublished Ph.D. thesis. Institute of Cybernetics, Brunel University.

STEINBUCH, K. VON (1961). Learning matrices. *Kybernetika*, **3**: 117–24.

STEWART, D. J. (ed) (1967). *Automaton Theory and Learning Systems*. Academic Press.

STOHLUROW, L. M. (1961). *Teaching by Machine*. OE 34010, CooperativeResearch Monograph 6. U.S. Department of Health, Education and Welfare.

TANNER, W. P. and J. A. SWETS (1954). A decision making theory of visual detection. *Psych. Rev.*, **61**: 401–09.

TARJAN, R. (1963). Problems of stability in control systems. In: *Optimising and Adaptive Control*. Bollinger, Truxall and Minar (eds). Instrument Soc. Amer.

THORPE, W. H. (1956). *Learning and Instinct in Animals*. Methuen.

TINBERGEN, N. (1953). *Social Behaviour in Animals*. Methuen.

TODA, M. (1962). Design for a fungus eater. *Behav. Sci.*, **7**: 75–87.

TURING, A. M. (1963). Computing machinery and intelligence. In: *Computers and Thought*. E. Feigenbaum and J. Feldmann (eds). McGraw-Hill.

TURING, A. M. (1969). Intelligent machinery. In: *Machine Intelligence*. Vol. 5. B. Meltzer and D. Michie (eds). Edinburgh University Press.

VAN DER VELDT, J. (1928). Cited in Woodworth, R. S. (1950).

VAN GICH, J. P. (1971). Changes in the mental content of work exemplified by lumber sorting operation. *Int. J. Man-Machine Studies*, **3**(1): 12–29.

VITANYI, P. (1973). Sexually reproducing automata. *Math. Biosci.* (in press).

VON FOERSTER, H. (1960). On self-organising systems and their environments. In: *Self-Organising Systems*. M. C. Yovitts and S. Cameron (eds). Pergamon Press.

VON FOERSTER, H. (1971). Molecular ethology. In: *Molecular Mechanisms of Memory and Learning*. C. Ungar (ed). Pergamon.

VON NEUMANN, J. (1969). *Theory of Self-Reproducing Automata*. A. Burks (ed). University of Illinois Press.

WADDINGTON, C. H. (1957). *The Strategy of The Genes*. George Allen and Unwin.

WALD, A. (1950). *Statistical Decision Functions*. Wiley.

WARREN, R. M. (1961). Illusory changes of distinct speech upon repetition—the verbal transformation effect. *Brit. J. Psych.*, **52**(3): 249–58.

WELFORD, A. T. (1968). *Fundamentals of Skill*, Methuen.

WIENER, N. (1948). *Cybernetics* (2nd edn., 1962). Wiley.

WILKINS, L. (1963). *Social Deviance*. Tavistock.

WINOGRAD, T. (1972). *Understanding Natural Language*. Edinburgh University Press.

WINSTON, P. (1970). *Learning Structural Descriptions by Example*. Project Mac TR–76, MIT Press.

WITTGENSTEIN, L. (1953). *Philosophical Investigations*. Trans. G. E. M. Anscombe. Blackwell.

WOODWORTH, R. S. (1950). *Experimental Psychology*. Methuen.

WYNNE EDWARDS, W. C. (1963). *Animal Dispersion*. Oliver and Boyd.

ZADEH, L. A. (1968). Fuzzy algorithms. *Inform. Contr.*, **12**: 94–102.

ZADEH, L. A. (1971). Toward a theory of fuzzy systems. In: *Aspects of Network and System Theory*. R. E. Kalman and N. DeClaris (eds). Holt, Rinehart and Winston.

ZADEH, L. A. (1973). Outline of a new approach to the analysis of complex systems and decision processes. In: *IEE Transactions on Systems, Man and Cybernetics*. Vol. SMC–3, 1, Jan 1973.

Subject Index

The index is selective and (deliberately) incomplete. Certain words or phrases appear often and are omitted; for example, 'Learning', 'Adaptation' (as psychological omissions); 'Feedback', 'System', 'Goal Directed', 'Alternative Set', 'Decision' (as Cybernetic or System Theoretic omissions); and 'Inference', 'Deduction/Induction' (in the field of logic and philosophy). The remaining, less common, words are usually indexed to points of discussion, rather than places where they are mentioned in the text.

Index of Names