Sertraline in Psychiatric Practice : a Topographical Study

Ernestina Politi, Carlo Balduzzi, Enirico Smeraldi, Laura Bellodi
Department of Neuropsychiatric Sciences
Università degli Studi di Milano
Ospedale S.Raffaele Ville Turro
via Prinetti, 29 - Milano
Telephone # 39-2-2643 3315
politi.ernestina@hsr.it
e_politi@yahoo.it

Abstract

We propose an operational research contribution to clinical psychopharmacology ; the key problems of drug selection and outcome prediction are tackled in a retrospective study about Mood and Anxiety Disorders focusing on sertraline, a selective inhibitor of serotonin reuptake. A three-step approach to data coding, clinical modeling and rule extraction is proposed, based on topographical techniques (Kohonen's Self-Organizing Maps) and information theory (Shannon entropy and mutual information). Clinical data are bitwise sampled, allowing an unbiased definition of system metrics. Uncertainty measures are introduced for a real-world sized approach to clinical practice ; top-down induction decision trees (TDIDT) for drug administration are proposed, and a default logic of prescription is analyzed in the light of direct clinical experience and available literature data.

Keywords : Sertraline ; Cybernetics ; Operational research ; Information theory ; Self-organization ; Neural networks.

1. Introduction

The present paper deals with the problem of predicting the outcome of a psychotropic drug treatment, according to a cybernetic standpoint. Its main concern is therefore to suggest some procedures for evaluating the degree of predictability inside complex bodies of phenomenological knowledge, rooting in clinical practice ; its path runs aside from the more traditional controlled clinical trials, which are carried out within selected boundaries and protected conditions, according to the golden rules of positivistic inference.

Whichever the standpoint chosen, inferential or cybernetic, the *act* of inquiring and the *object* of such inquiry belong to distinct epistemological domains (Black, 1937 ; Aiello, 1986). Clinicians choose a given drug for treatment according to a complex conceptual frame, which evolves out of two distinct topological fields : a *default logic* yielding from a body of psychiatrical knowledge, which is inherited as a data structure (Reiter,

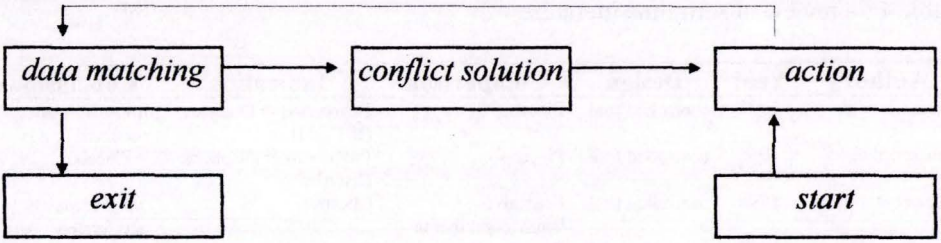
1980 ; Cohen, 1985) ; and a procedure of casewise prototyping (*scripting*), based on the sequence of events which appears from time to time the most feasible, thus exhibiting an adaptive and context-sensitive behaviour (Abelson, 1981).

The inferential approach to prediction in clinical psychopharmacology is, in turn, mainly concerned with the definition of a default operational logic - that is, a general class of experimental objects, and a comparatively small set of standardized procedures for handling such objects (Bhatnagar, 1986 ; Anderson, 1987 ; Pearl, 1988). On the other hand, the contribution of a cybernetic approach can be expressed in terms of *prototyping* and *optimization.*

The suitable tools for this latter approach belong to the field of parallel distributed processing (Fu, 1983 ; Kaufmann, 1993 ; Kohonen, 1995/1997) and system modeling paradigms (Barnett, 1981 ; Pearl, 1986 ; MacLane, 1991); these tools provide the observer with the skill to capture and identify the real-life system by building a *black box* model, usually containing one or more complex, nonlinear functions (Cheng, 1984 ; Arthur, 1994). The first step in cybernetic system identification shall satisfy the requirements of both efficiency and economy (Minsky, 1975 ; Collins, 1988) ; the observer will be able to prototype a numeric model, featuring the highest achievable degree in learning or self-organizing capabilities (Plotkin, 1993 ; Ito, 1994 ; Kaufmann, 1994), and exhibiting a reliably adaptive behaviour when faced with new data. Put in other words, we could think about a real-world complex system as an analogue signal with a high informational content (although affected by a potentially high *Signal/Noise ratio*), while the cybernetic prototype is a digitally sampled model of such a system, which enables a faithful reproduction of its main features and a precise structural definition of the original sampling space (Chaitin, 1987 ; Dubois, 1989).

The second step that we propose deals with optimization (Wolpert, 1996). One can think about it as a procedure of digital signal processing, where some kind of transformation or filter is applied, in order to identify the most relevant wavelengths (that is, the maximized information) with respect to the previsional target, and a new generation of operational rules (scripts) is generated, according to the identified topological properties. Scripts can be therefore revised in a cyclical fashion, according to the many different objects, roles, scenes and starting conditions merged into the system (Gunji, 1996). What is finally achieved is an algorithm for feeding information back to the system agents, resulting in updated, more efficient and more specialized scripts. In this phase of computation, a comparatively new set of coding, problem designing and data analysis techniques can be useful in the task of working out the desired new operational rules, namely self-organizing neural networks and top-down induction decision tree programming (Quinlan, 1993).

The cyclical procedure proposed here can be reduced to a basic cycle of recognition-and-action, namely a '*TOTE ring*' (Test-Operation-Test-Exit) ; the process (Buchanan, 1984 ; Carrier, 1989) can be described by means of a simple flowchart, as follows :



Cybernetic prototyping therefore requires a series of three conceptual and operational steps :

- *coding phase* - where the whole available body of information is considered, according to a relational conception of data ;
- *modeling phase* - where observed events and conditions are assigned to classes and class relations are defined, in the most efficient way with respect to the target of reducing the structural functions in the model ;
- *programming phase* - where an ordered organization of data is identified, so as to build a comparatively small set of operational programs.

The structural relationships described by an efficient script supply a numerical model of the desired condition of mutual adaptation between stimuli (action of drugs) and responses (clinical evaluation of outcome) ; both actors, clinician and patient, gain some knowledge about the new *frame* portraying the *object-drug*, rather than about the object itself (Szolovits, 1978). The phenomenology of interaction between drugs and biological/clinical actors identifies a *structural coupling* conceivable as an operational script in two senses : a) development of a clinical action, b) development of resources which are able to modify the frame of the clinical object (Sandell, 1985).

Our study is based on retrospective data, gathered from the clinical records of psychiatric outpatients suffering from Anxiety and/or Affective Disorders, who received a pharmacological treatment with sertraline (Murdoch, 1992). Individual reports filed in our records often appeared remarkably complex, under a both clinical and therapeutical standpoint, with anamnestic evidence of frequent trial-and-error psychopharmacological search procedures, featuring several different drugs and a waxing-and-waning clinical course over time. Moreover, data available in literature seem themselves to point to a quite complex *default logic* about the clinical use of sertraline ; while showing no particular advantage over other SSRIs in the treatment of Anxiety and Affective Disorders altogether (Thase, 1996 ; Franchini, 1997 ; Kocsis, 1997 ; Tourigny-Rivard, 1997), sertraline is considered by several Authors as the first choice drug in a series of highly selective pathological conditions. Such clinical experiences have produced, over the last years, a wealth of inferential studies about this drug ; Table 1 provides a

Table 1 : a review of sertraline literature

Author	Year	Design	Comparison	Indication	Conclusions
Goodnick et al.	1997	controlled trial	Placebo	Depression in Diabetes (NIDDM)	improvement observed
Yonkers et al.	1996	controlled trial	Placebo	Premenstrual Disphoric Disorder	> Placebo
Ricca et al.	1996	controlled trial	Cognitive-behavioural therapy	Obesity	weight loss accelerated
Zanardi et al.	1996	controlled trial	Paroxetine	Delusional Major Depression	> Paroxetine
Greist et al.	1995	controlled trial	Placebo	Obsessive Compulsive Disorder	> Placebo
Mendels et al.	1995	controlled trial	Placebo	Premature Ejaculation	> Placebo
Swartz et al.	1997	uncontrolled trial	SSRIs	Frontotemporal Dementia	improvement observed
Constantino et al.	1997	uncontrolled trial	SSRIs	Aggressive behaviour in adolescents	no major improvement
McConville et al.	1996	uncontrolled follow up study	None	Major Depression in Adolescents	improvement observed
Steingard et al.	1997	open label trial	None	Transition Anxiety in autistic children	improvement observed
Thakore et al.	1996	open label trial	Various Neuroleptics	Negative symptoms in Chronic Schizofrenia	improvement observed
Hellings et al.	1996	open label trial	None	Mental Retardation in Autistic Disorder	improvement observed
Brady et al.	1995	open label trial	None	PTSD with Alcohol Dependence	improvement observed
Scott et al.	1995	open label trial	None	Depression in Multiple Sclerosis	improvement observed
Dow et al.	1997	prospective study	Fluoxetine, Nortrypt., Desipr.	PTSD with Major Depression	SSRIs performing better
Kafka	1994	prospective study	Sequential use of Fluoxetine	Paraphilias	improvement observed
Popli et al.	1997	case series	None	Psychotic symptoms	exacerbation observed
Mukand et al.	1996	case series	None	Pathological Crying/Laughing	improvement observed
Volicer et al.	1994	case series	None	Alzheimer Disease with Depression	improvement observed
Ranen et al.	1996	case report	None	Aggressiveness in Huntington Disease	improvement observed
Dodt et al.	1997	letter	-	Risperidone-induced OCD	-
Sprenger	1997	letter	-	Nocturnal Enuresis	-
Young	1997	letter	-	Hallucinogen Perception Disorder	-
Rubenstein et al.	1996	letter	-	Transvestic Fetishism	-
Kalivas et al.	1996	letter	-	Neurotic Excoriations	-
El-Khatib et al.	1995	letter	-	Body Dysmorphic Disorder	-
Czepowicz et al.	1995	letter	-	Social Phobia	-
Papp et al.	1995	letter	-	Pulmonary Disease with Anx./Mood Disorders	-
Rahman et al.	1995	letter	-	Trichotillomania with HIV Infection	-
Frankenburg et al.	1994	letter	-	ADHD and Tourette	-

summary of recently published data on sertraline, including clinical trials and case reports focusing on specific clinical conditions and complex strategies of treatment, featuring frequent drug associations. However, no exhaustive analysis is available to date about the *quality* of this psychopharmacological knowledge of sertraline, nor about the *inner organization* of such knowledge. We may think that the optimal use of sertraline is narrowed to a few, somewhat atypical clinical conditions, but we lack an overall conceptual frame about the drug's ideal conditions of use.

In this scenario, a cybernetic inquiry may well candidate for a process of identification and optimization of a *practice* which lies perhaps beyond the boundaries of inferential description, however hiding some privileged information which may deserve a more formal approach. The real-world clinical 'signal' can thereby be sampled and processed with the aid of advanced numerical tools, in a way that should enable the observer to easily test the goodness of fit of his/her hypotheses ; this represents the final step in a process of transformation of models (reframing) into rules of production (rescripting).

2.Method

a. coding phase : gathering information

This study was developed after a base of knowledge deriving from a pool of clinical information ; data were gathered through the case descriptions filed by six senior psychiatrists into the clinical records of an outpatient population, being observed over a 12 month period. The sample of population included 69 outpatients, who were treated with sertraline for a wide range of clinical conditions.

We developed a project of data encoding which would allow us to identify a consistent base of knowledge, keeping in mind that numbers would have to encompass as much as possible of the real-world clinical experience. A 'molecular' level of coding was defined for all available data - i.e. for 65 distinct epidemiological, diagnostic, clinical and pharmacological bits of information recorded by our clinicians ; data were thus re-engineered into a *discrete*, *granular* and *highly connected* pool of information (Wyatt, 1991). The chosen bitwise binary model is fairly well known in the fields of pattern recognizing and digital signal processing ; in summary, it features an unbiased *a priori* definition of stronger and weaker cells of information, according to the assumption of equiprobability and equipotence in the starting conditions of the universe under observation (Dubois, 1980 ; Grosoff, 1986). Bitwise coding was also applied to continuous variables ; for each of them, the observed range was splitted into a convenient number of discrete classes, each gathering approximately the same number of real observations.

Conventions and details of the adopted coding method are reported in the left side of Table 2.

Table 2 : variables and codes.

Entropy H (x, 1-x) is computed, according to Shannon, as $(_i\ p_i \log (1 / p_i)$, where p_i is the probability of event *i* occurring over *n* possible events ($n = 2$ in the case of a binary code).
For a given output, x, we assume : $H (X|y_i) = (_i\ p (x_i|y_i) \log (1 / p (x_i|y_i)$ as the entropy (uncertainty measure) for output x, given input y_i ; in the whole set of observations,

$$H (X|Y) = (_i\ p (y_i)\ H (X|y_i) = (_i\ (_j\ p (x_i, y_j) \log (p (y_i) / p (x_i, y_j))$$

is the uncertainty measure for output X, and

$$I (X ; Y) = (_i\ p (x_i) \log (1 / p (x_i)) - (_i\ (_j\ p (x_i, y_j) \log (p (y_i) / p (x_i, y_j)) = (_i\ (_j\ p (x_i, y_j) \log (p (x_i, y_j) / p (x_i)\ p (y_j))$$

is the mutual information between each output x and each input y, where $i = 2$ for each input variable and $j = 4$ for distinct outcome classes ; logarithm base is 4, i.e. a complete mutual information between inputs and ouputs is described when $I (X ; Y) = 2$.

Bit	variable	class	code	frequency	H(x,1-x)	H(x|y)	I(x;y)
1	*Sex*	-	0,1	0.725	0.849	0.993	**0.015**
2	*Age*	*≥64*	1111111	0.101	0.474	0.977	**0.045**
3	*(if ≤ 28 code* **0000000***)*	*55-63*	0111111	0.203	0.728	0.968	**0.065**
4		*50-54*	0011111	0.319	0.903	0.952	**0.096**
5		*46-49*	0001111	0.464	0.996	0.970	**0.060**
6		*40-45*	0000111	0.609	0.966	0.959	**0.081**
7		*33-39*	0000011	0.725	0.849	0.962	**0.076**
8		*29-32*	0000001	0.855	0.597	0.988	**0.025**
9	*Major Affective Disorder*	*bipolar*	0,1	0.087	0.597	0.984	**0.032**
10	*(M.A.D.)*	*recurrent*	0,1	0.478	0.999	0.975	**0.049**
11		*psychotic*	0,1	0.029	0.189	0.987	**0.026**
12	*Anxiety Disorder*	*Panic Dis.*	0,1	0.377	0.956	0.998	**0.003**
13	*(A.D.)*	*Generalized A.D.*	0,1	0.073	0.375	0.989	**0.022**
14		*Social Phobia*	0,1	0.014	0.109	0.965	**0.071**
15		*Somatoform*	0,1	0.029	0.189	0.980	**0.039**
16		*Agoraphobia*	0,1	0.145	0.597	0.986	**0.028**
17		*Obsess. Compuls. Dis.*	0,1	0.217	0.755	0.983	**0.034**
18	*Other Disorders*	*Tic Disorder*	0,1	0.015	0.109	0.994	**0.013**
19		*Abuses*	0,1	0.116	0.518	0.983	**0.034**
20		*Impulse Control Dis.*	0,1	0.029	0.189	0.987	**0.026**
21		*Eating Disorder*	0,1	0.058	0.319	0.988	**0.023**
22	*Lifetime drugs*	*BDZ*	0,1	0.898	0.473	0.977	**0.047**
23		*TCA*	0,1	0.536	0.996	0.972	**0.056**
24	*Other simultaneous drugs*	*(any)*	0,1	0.739	0.828	0.962	**0.076**
25	*Treatment (months)*	*≥22*	1111111	0.058	0.319	0.977	**0.045**
26	*(if ≤ 1 code* **0000000***)*	*15-21*	0111111	0.174	0.667	0.967	**0.065**
27		*10-14*	0011111	0.333	0.918	0.949	**0.103**
28		*8-9*	0001111	0.493	0.999	0.919	**0.161**
29		*5-7*	0000111	0.609	0.966	0.868	**0.265**
30		*3-4*	0000011	0.768	0.781	0.891	**0.219**
31		*2*	0000001	0.870	0.130	0.885	**0.231**
32	*Axis II Disorder*	*(any)*	0,1	0.304	0.887	0.988	**0.024**
33	*Other clinical features*	*Depressive*	0,1	0.725	0.849	0.962	**0.077**
34	*and complains*	*Anxious*	0,1	0.696	0.887	0.986	**0.029**
35		*Somatic*	0,1	0.217	0.755	0.975	**0.050**
36	*Age at onset*	*≥46*	1111111	0.145	0.597	0.981	**0.039**
37	*(if ≤ 16 code* **0000000***)*	*38-45*	0111111	0.275	0.849	0.981	**0.039**
38		*32-37*	0011111	0.406	0.974	0.967	**0.067**
39		*29-31*	0001111	0.507	0.999	0.974	**0.052**
40		*24-28*	0000111	0.638	0.945	0.995	**0.010**
41		*21-23*	0000011	0.768	0.781	0.980	**0.040**
42		*17-20*	0000001	0.870	0.559	0.998	**0.004**

Table 2 (continued)

Bit	variable	class	code	frequency	H(x,1-x)	H(x|y)	I(x;y)
43	***Recent life events***	-	0,1	0.333	0.918	0.974	**0.052**
44	***Illness in Ist degrees***	*M.A.D.*	0,1	0.522	0.999	0.990	**0.020**
45		*A.D.*	0,1	0.304	0.887	0.982	**0.036**
46		*Sociopathy*	0,1	0.189	0.698	0.945	**0.110**
47	***N. visits***	*≥ 25*	1111111	0.159	0.633	0.994	**0.013**
48	***(if ≤ 2 code 0000000)***	*15-24*	0111111	0.275	0.849	0.980	**0.039**
49		*11-14*	0011111	0.377	0.956	0.976	**0.049**
50		*9-10*	0001111	0.507	0.999	0.979	**0.042**
51		*7-8*	0000111	0.638	0.945	0.980	**0.039**
52		*5-6*	0000011	0.754	0.806	0.987	**0.026**
53		*3-4*	0000001	0.942	0.319	0.950	**0.101**
54	***Sertraline side effects***	-	0,1	0.348	0.932	0.974	**0.053**
55	***Other lifetime SSRIs***	*Fluoxetine*	0,1	0.232	0.781	0.956	**0.083**
56	***and side effects***	*s.e.*	0,1	0.101	0.474	0.951	**0.097**
57		*Fluvoxamine*	0,1	0.304	0.887	0.987	**0.025**
58		*s.e.*	0,1	0.145	0.597	0.978	**0.046**
59		*Paroxetine*	0,1	0.174	0.667	0.947	**0.107**
60		*s.e.*	0,1	0.073	0.375	0.980	**0.039**
61		*Citalopram*	0,1	0.116	0.518	0.976	**0.049**
62		*s.e.*	0,1	0.087	0.426	0.961	**0.078**
63	***Psychiatric admissions***	*(lifetime)*	0,1	0.333	0.918	0.981	**0.039**
64	***Clinical improvement***	-	0,1	0.725	0.849	-	-
65	***Long-term outcome***	-	0,1	0.812	0.698	-	-

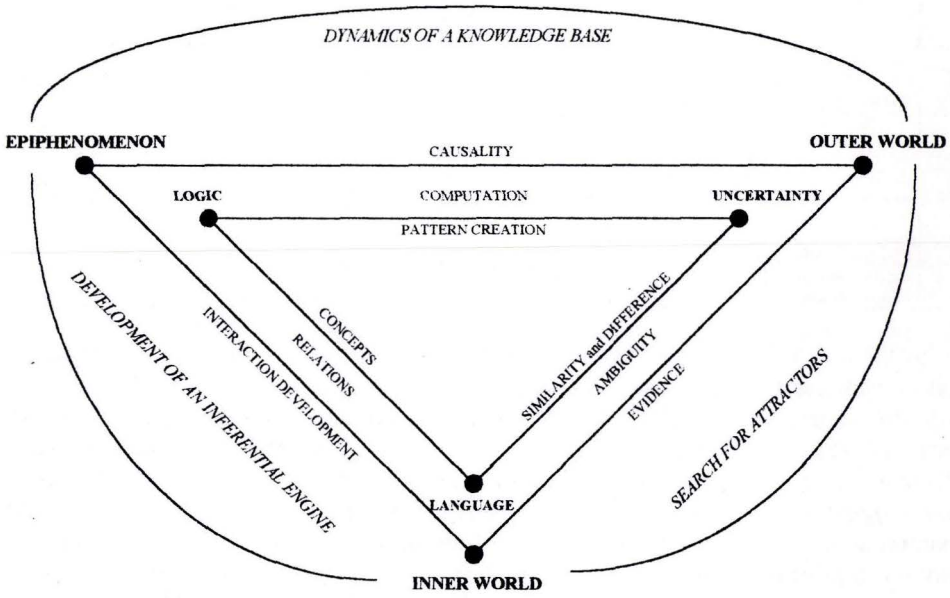
As to the outcome of treatment, we defined a 'global clinical improvement' criterion (bit 64 : efficacy) along with a 'reliability over time' criterion (bit 65 : duration), in order to properly classify four distinct classes of outcome : *responders* to a full-length sertraline treatment (code 11) ; *partial responders*, being described in the earlier phases of clinical improvement, thereby not allowing for a long-term evaluation (code 10) ; *non responders*, where no major improvement was achieved over a convenient time span (code 01) ; finally *dropout* patients, whose assessment was untimely interrupted, and no significant benefit had been achieved by the time of the latest available observation (code 00). Again, respective frequencies and codes are given in Table 2.

We assumed that the pool of information obtained so far could mirror the summed effects of the interconnected epidemiological, diagnostical, pharmacological and clinical dominia defining the dynamic system under study. Shannon entropy (Shannon, 1949) was computed after each variable, to yield a measure of uncertainty according to an informational perspective. Moreover, we computed a measure of conditional entropy, H (X|Y), between each bit of input information and the pair of bits coding for treatment outcome (i.e. a measure of uncertainty of any given symbol transmitted from a binary source in a finite realization), and the related measure of mutual information, I (X;Y), defining the degree of certainty associated to the transmission of symbol x (outcome codes) from the native source, given to receive the symbol y (input codes) over the binary channel (Swingler, 1995). Values for informational measures are reported in the right side of Table 2, along with the adopted formulas and conventions.

b. modeling phase : measuring and organizing information

In synthesis, the dominion under study can be modeled by a circular space (Ehresmann, 1997) with an inscribed polygon, bearing three key elements on its vertices : causal logic, uncertainty about the outer world, natural language as inner symbolic expression (see Figure 1).

Figure 1 : evidence, causality and interaction.

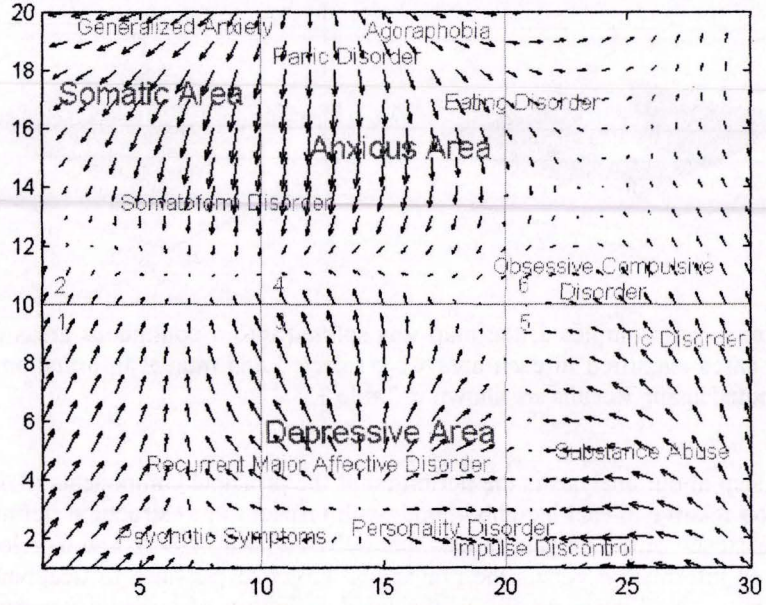


The modeling phase of analysis aims at defining a topographical organization of the sampled universe, within linear boundaries ; it is reasonable to presume that different areas in data correspond to individual partitions of the overall pool of information, where a specific portion of data gains relevance in view of its particular topological status.

A topographical projection with such properties is the main feature of unsupervised neural network architectures such as Kohonen's Self Organizing Map (Kohonen, 1988), which was chosen for this study (see Figure 2 for architectural details and training parameters). Kohonen's projection creates a reduced-dimensionality virtual space - in our case, two dimensions in a 30*20 grid of neurons - where separate regions are defined for each case in the training set over the whole training cycle, displaying in the end some discrete patterns, or 'bubbles', of activation (Utsugi, 1996). This projective plain is sampled at discrete intervals (neurons) and can be splitted into contiguous portions (i.e. self-similar subspaces, corresponding in our cases to six evenly spaced,

squared areas of 10*10 units) called *neighborhoods*, where self-similarity (Yin, 1995) is guaranteed by the function of lateral inhibition adopted by the map (a variant of 'mexican hat' function : see Figure 2, lower section for details).

Figure 2 : Kohonen self-organizing map.



A view of Kohonen's topographic projection. Arrows point to areas of successful treatment (steady global clinical improvement). Our map consisted of a 30*20 neuron grid, with random initial synaptic weights ((0.05). Training consisted in 100 presentations of the whole sample, in a random order ; after each case, the winner neuron was selected according to the least euclidean distance criterion from the index vector. Synaptic weights were then updated according to the Hebbian rule :

$$(w = (* (* |x - w|$$

where ((learning rate) is linearly decreasing over training time (starting value = 0.1) and (is defined as :

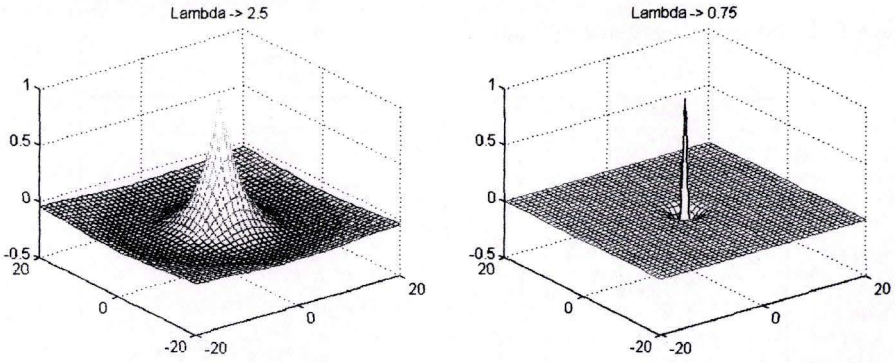
$$(= (1 - d^2 / 2 * (^2) * \exp (- d^2 / 2 * (^2)$$

where d is the euclidean distance between the winner and all other neurons in the grid, and (is also linearly decreasing over time (starting value = 2.5, end value = 0.75) so as to gradually reduce the bubble of activation around the winner.

Function (defines a neighborhood including :

a) the winner neuron
b) a ring ('bubble') of partial activation
c) an outer ring of inhibition.

The plots below illustrate the reducing diameters of the neighborhood for values of (corresponding to the starting and ending values used in this experiment.



After training was completed, the map was splitted into 6 contiguous areas of 10*10 neurons ; cases classified in each area were selected and mutual information I (x ; y) was computed again. Results are shown in Table 3.

The next step in our analysis is the definition of the principal components in the binary information relative to each neighborhood, with respect to : a) structural definition and inner consistency of the neighborhoods themselves (Erwin, 1992), and b) relevance of the selected information versus the previsional target of response to treatment. These concepts correspond, respectively, to the *local* definition of synaptic weights (Ritter, 1989) and to the measure of mutual information which was previously introduced over the whole sample. Table 3 displays the results of these piecemeal computation, after the map was divided into its six squares ; it can be noted that this topological partition causes an increase in the previsional power of at least certain input variables, acting as a digital filter applied to the original signal.

In short, the procedure employed so far can be summarized in two points : a) apply a splitting algorithm on complex data being digitally sampled with a high resolution power, b) apply a digital filter in order to select the locally relevant pieces of information.

Table 3 : the 6 areas in the map

Area (N of subjects)	Clinical features	f req.	outcome frequencies 00	01	10	11	Outcome predictors	I (x ; y)
1 (11)	*Sex F*	1.00	.091	.182	.091	.636	Treatment (10 months	**1.302**
	Other drugs	.72					Lifetime fluoxetine	**1.026**
	M.A.D. in 1st degrees	.72					Age (55 years	**0.517**
	Recurrent M.A.D.	.63						
2 (12)	*Anxious area*	.75	.167	.083	.000	.750	Treatment (2 months	**1.333**
	P.D.	.66					Onset (37 years	**1.037**
	Depressive area	.66					Antisocial	**1.037**
							1st degrees	
	A.D. in 1st degrees	.50						
3 (15)	*Recurrent M.A.D.*	1.00	.067	.133	.067	.733	Agoraphobia	**0.470**
	Depressive area	1.00					OCD	**0.144**
	Sex F	.93						
	Onset ≤ 28 years	.73						
4 (10)	*Sex F*	.80	.100	.300	.000	.600	Treatment (2 months	**1.257**
	Anxious area	.70					Age (32 years	**0.579**
	P.D.	.60						
	Depressive area	.50						
5 (11)	*Depressive area*	.81	.182	.182	.182	.455	Treatment (2 months	**0.317**
	Anxious area	.81					Lifetime admissione	**0.317**
	OCD	.72					Depressive area	**0.105**
	Axis II	.45					Anxious area	**0.039**
6 (10)	*Anxious area*	.50	.000	.200	.200	.600	Sertraline S.E.	**0.892**
	OCD	.50					Treatment (2 months	**0.892**
	PD	.40						
	Somatic area	.20						

c. programming phase : graph construction

Once a local hierarchy has been defined, we can think of a global tree of information pointing to the whole sampled universe. Each area in this universe has a measurable informational definition that we could use as a membership function for projection or representation of any given case ; as the final step, we were able to build a decision tree that took into account the local measures of mutual information between a few selected wavelenghts in the sample and their respective previsional targets.

The adopted model was a variant of Top-Down Induction Decision Tree (TDIDT), where a minimal set of conditions was used to represent the whole set of available information (Quinlan, 1993) ; by iteratively including input variables (starting with the one with the highest local previsional power, and proceeding in a decreasing order, until all cases in the sample were correctly classified), we defined a set of six local sub-trees (*direct acyclical graphs*) performing a near-optimal search in the sample, based on the described topographical euristics. It should be noted that the search space in the unreduced binary tree - i.e., in the absence of an efficient digital filtering mechanism - would be quite too wide for any exhaustive search algorithm ; a system with 63 binary input variables has 2^{63} possible solutions - a number that raises the dreaded ghost of intractability and NP-completeness.

Besides solving the combinatorial dilemma, the TDIDT approach deals in this case with areas of self-similarity whose topological and metric properties should make sense to the clinician. Input variables are treated as points of accumulation in a structural and previsional perspective (Ginberg, 1985), with the self-organization paradigm serving as a first-order filter to enhance the originally high S/N ratio. Local graphs are shown in Figure 3.

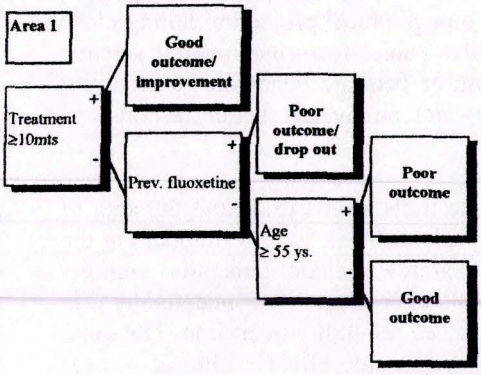
3.Results

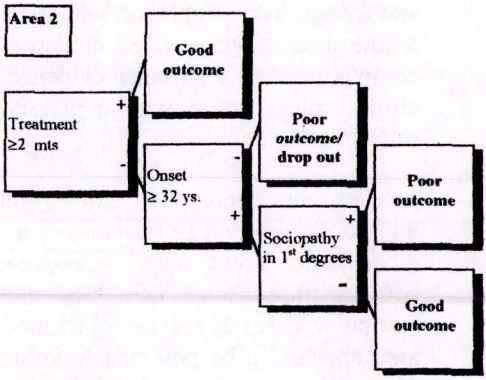
Kohonen's map builds a discrete topological landscape where a continuous kaleidoscope of conditions is depicted, portraying a pool of clinical experience which undergoes a process of partition, as if through the action of a potential varying across an electrophoretic field (Cortes, 1987). This 'energy field' captures a continuous and harmonic proportion of clinical conditions and diagnostic informations, with a first sight evidence of a clear-cut separation between 'pure' diagnostic prototypes - falling along the edges of the map, due to the learning algorithm which proceeds through decreasingly wide bubbles of activation - and more complex assemblies of clinical conditions, falling towards the center of the map (Figure 2).

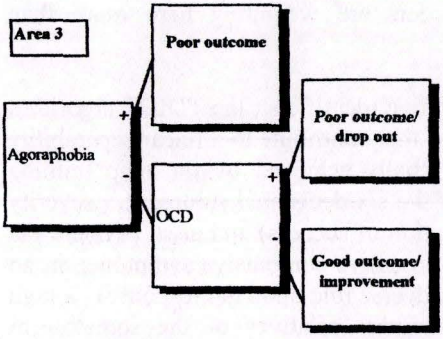
The background in Figure 2 is a gradient plot of the synaptic weights corresponding to bit 64 (clinical improvement) ; arrows are drawn away from areas of unsuccessful treatment, moving towards subspaces with a better expectations of response. At least two well defined basins of attraction are recognizable in the map, corresponding to a major 'wave front' between panic/phobic and major affective areas (middle), and to a smaller borderline domain between obsessive-compulsive spectrum disorders and major depression areas (lower, right side) ; in this latter region the synaptic weights for bit 64 are rather high, while bit 65 (not shown for the sake of readability) turns out to be quite lower, thus representing partial or short-lasting responders to sertraline.

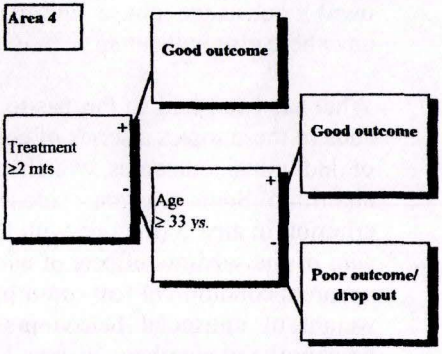
Cases of recurrent major depression and depressive syndrome with marked bodily complains exhibit here a better response to sertraline than bipolar disorders do ; an all-or-nothing pattern of action is suggested by the lack of partial responders in area 2

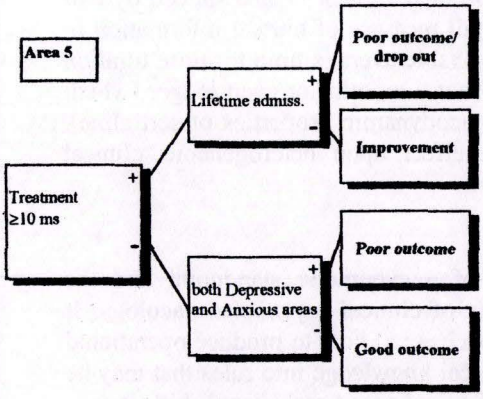
Figure 3 : local decision trees

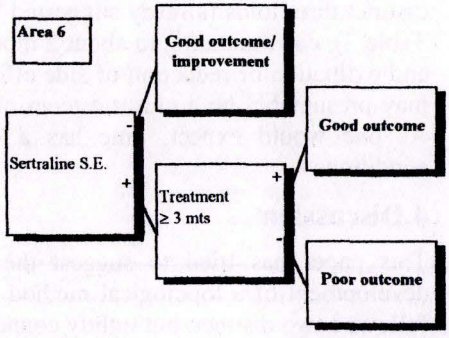












(somatoform neighborhood) and area 4 (panic disorder neighborhood). A similar statement applies to the domain of anxiety disorders ; we can describe a prototype of panic disorder featuring a marked processual nature, a strong agoraphobic component, and a cognitive completion which accounts for a topographical projection falling close to the area of generalized anxiety ; and a 'bodily' panic, featuring marked somatic complains, with a stronger evidence of a dynamic or periodic behaviour, a recurrent clinical presentation with depressive symptoms, and an overall better response to sertraline treatment.

Anxiety with depressive features and no anancastic traits seems to denote the area of highest expectation of response to a pharmacological treatment with sertraline. On the other hand, areas with obsessive-compulsive features include generally younger patients, more often men, and exhibit a fairly higher degree of uncertainty ; in particular, cases of successful treatment show a reduced reliability over time. The same area appears to be particularly vulnerable to the drug's side effects ; clinical records typically mention sertraline after a wealth of other drugs, the most common being fluvoxamine (which thereby may turn out to have an interesting default logic of its own) ; substance abuse and personality disorders are weighting here more than anywhere else in the map.

What has been said so far, has to do with descriptive identifiers ; the TDIDT algorithm adds to these topics a series of cellular conditions that contribute to a linear separability of individual outcomes, which can be only partially achieved by the map training algorithm. Some relevant evidences stem out of the six decisional subtrees : a severity criterion in area 5 (the one with the least expectation of success) and area 3 (where the sum of the separate effects of agoraphobia and obsessive-compulsive symptoms, in an apparent condition of real comorbidity, plays an adverse role upon the response) ; a high weight of antisocial behaviours among first degrees relatives of the somatoform neighborhood members, in area 2 ; the presence of time thresholds in the age at onset of illness, or in present age (areas 1 and 4). Duration of treatment is also spliced by two distinct thresholds (already suggested by the overall measure of mutual information in Table 2), corresponding to about 2 months (a reasonable average time for dose titration and extinction or reduction of side effects) and about 6 months, or even longer (which may presumably be a closer esteem of the pharmacodynamic properties of sertraline). As one would expect, time has a differential effect upon heterogeneous clinical conditions.

4.Discussion

This paper has tried to suggest the adoption of a cybernetic standpoint and the development of a topological method in the study of clinical psychopharmacology. It followed two distinct, but tightly connected approaches : a) how to produce operational scripts that can synthesize a growing body of clinical knowledge into rules that may be immediately applied and revised, in the interest of an enhanced previsional skill (Peters, 1993) ; b) how to identify relevant topological areas that correspond to discrete

phenotypes of drug action, which may be usefully developed in a pure research perspective, eventually linking to clinical, molecular and genetic findings (Andreassen, 1991).

This methodological aims mark a first step towards an open, dynamical research model that may satisfactorily address a key question in clinical psychopharmacology : what is really relevant to know, while meeting a patient for the first time and along the later course of treatment, and where to focus a clinician's attention in order to capture the ongoing process. Each drug acts in the realm of the complex events of illness as a *structure modifier* - a putative catalyser of events and elements belonging to the system under study (Arbit, 1982) ; at the same time, all the relevant items in this catalytic process need to be identified in the open structure of a knowledge base, which shall be suitable for validation and reproduction, and able to encompass both stronger and weaker cells of information into an open frame allowing for a thorough communication.

We used a topological tool in order to define a somewhat intermediate level of dimensionality - one that causes information to collapse into a new frame with many degrees of freedom (Sel, 1997), but small enough for a procedure of validation and optimization ; the subsequent use of entropy measures aims at the computation of a previsional esteem, implying the least possible dependence upon noisy conditions (Clowes, 1971). Logic requires closed worlds, and order requires closed rules of organization and a non-uniform distribution of information, while every feasible symbolic convention must be realistically complex, in order to convey a real-world-sized contribution to clinical knowledge (Dählback, 1989).

Above all, this study is concerned with defining a measure of 'drug administrability' in some specific conditions, belonging to a possibly selective, but not too narrow clinical dominion. We investigated a somewhat obscure default logic for a drug with paradoxical evidences of a first-choice status in some closed partitions and scenarios of the clinical experience, and a last-in-the-chain status in the realm of complex or multifaceted clinical conditions. At least in the latter case, the unspoken script for sertraline administrability seemed to sound in the beginning like «Use sertraline when you can't do without it», a statement compounding a strategy of choice that may well be considered optimal in reducing uncertainty under two respects : 1) «Use sertraline when anything else has failed» ; and 2) «Use sertraline when nothing can be expected to work better». The same frame was thereby investigated from within, in order to maximize the contribution coming from clinical experience ; we only assumed that a clinical record would enclose a great deal of information about what had been said, observed and done by clinicians, laying the ground to the further work of rule extraction.

5. References

1. Abelson, R.P. (1981) The psychological status of the script concept. American Psychologist 36 : 715-729.

2. Aiello, L., Cecchi, C. and Sartini, D. (1986) Representation and use of metaknowledge. Proc. of the IEEE 74 (10) : 1304-1321.
3. Anderson, J.R. (1987) Methodologies for studying human knowledge. Behavioral and Brain Sciences 10 : 467-505.
4. Andreassen, S., Jensen, F.V. and Olesen, K.G. (1991) Medical expert systems based on causal probabilistic networks. Int J Biomed Comput 28 : 1.
5. Arbit, M.A. (1982) Cooperative computation and the cybernetic society. In: Progress in Cybernetics and Systems Research, ed. R. Trappl, Vol. 9 : 3-12. Washington, Hemisphere.
6. Arthur, W.B. (1994) On evolution of complexity. In: Cowan, G.A., Pines, D. and Meltzer, D., eds. Complexity Metaphors, Models and Reality, 65-81, 1st ed. Reading, MA, Addison-Wesley
7. Barnett, J.A. (1981) Computational methods for a mathematical theory of evidence. Proc. 7th IJCAI-81, 868-875, Vancouver, BC.
8. Bhatnagar, R.K. and Kanal, L.N. (1986) Handling uncertain information: a review of numeric and non-numeric methods. In: Uncertainy in Artificial Intelligence, ed Kanal, L.M. and Lemmer, J.F., 3-25, Amsterdam, Elsevier.
9. Black, M. (1937) Vagueness, an exercise in logical analysis. Philosophy of Science 4 : 427-455.
10. Brady, K.T., Sonne, S.C. and Roberts, J.M. (1995) Sertraline treatment of comorbid posttraumatic stress disorder and alcohol dependence. Journal of Clinical Psychiatry 56 (11): 502-5.
11. Brown, W.A. and Harrison, W. (1995) Are patients who are intolerant to one serotonin selective reuptake inhibitor intolerant to another ? J Clin Psychiatry 56: 30-34.
12. Buchanan, B. and Shortliffe, E.H. (1984) Rule-Based Expert Systems. The MYCIN Experiments of the Stanford Heuristic Programming Project. Reading, MA, Addison-Wesley.
13. Carnap, R. (1950) Logical Foundations of Probability. Chicago, IL, University of Chicago Press.
14. Carrier, D.H. and Wallace, A.W. (1989) An epistemological view of decision-aid technology with emphasis on expert systems. IEEE Trans. on SMC 19 (5) : 1021-1028.
15. Chaitin, G.J. (1987) Information, Randomness and Incompleteness. Singapore, World Scientific.
16. Cheng, Y. and Fu, K.S. (1984) Conceptual clustering in knowledge organization. Proc. 1st Conference on Artificial Intelligence Applications, 274-279, Denver, CO.
17. Clowes, M.B. (1971) On seeing things. Artificial Intelligence 2 : 79-166.
18. Cohen, P.R. (1985) Heuristic Reasoning about Uncertainty. Los Alto, CA, Morgan Kaufmann.
19. Collins, A. and Michalski, R.S. (1988) The logic of plausible reasoning: a core theory. George Mason University, Fairfax, VA, Machine Learning and Inference Lab, Report, MLI-88-14.
20. Constantino, J.N., Liberman, M. and Kincaid, M. (1997) Effects of serotonin

reuptake inhibitors on aggressive behavior in psychiatrically hospitalized adolescents: results of an open trial. Journal of Child and Adolescent Psychoparmacology 7 (1): 31-44.
21. Cortes, C., Krogh, A. and Hertz, J.A. (1987) Hierarchical associative networks. Journal of Physics Ann 20 : 4449-4455.
22. Czepowicz, V.D., Johnson, M.R., Lydiard, R.B., Emmanuel, N.P., Ware, M.R., Mintzer, O.B., Walsh, M.D. and Ballenger, J.C. (1995) Sertraline in social phobia. Journal of Clinical Psychopharmacology 15 (5): 372-3.
23. Dählback, N. (1989) A symbol id not a symbol. Proc. IJCAI-89, 8-14, Detroit, MI.
24. Dodt, J.E., Byerly, M.J., Cuadros, C. and Christensen, R.C. (1997) Treatment of risperidone-induced obsessive-compulsive symptoms with sertraline. American Journal of Psychiatry 154 (4): 582.
25. Dow, B. and Kline, N. (1997) Antidepressant treatment of posttraumatic stress disorder and major depression in veterans. Annals of Clinical Psychiatry 9 (1): 1-5.
26. Dubois, D. and Prade, H. (1980) Conditioning in possibility and evidence theories - a logical viewpoint. Proc. Intl. Conf .on Information Processing and Management of Uncertainty in Knowledge-Based-Systems, 401-407, Urbino.
27. Dubois, D. and Prade, H. (1989) Measure-free conditioning, probability and non-monotonic reasoning. Proc. IJCAI-89, 1110-1114.
28. Ehresmann, C.A. and Vanbremeersch, J.P. (1997) Information processing and simmetry-breaking in memory evolutive systems. BioSystems 43 : 25-40.
29. El-Khatib, H.E. and Dickey, T.O. (1995) Sertraline for body dysmorphic disorder. Journal of the American Academy of Child & Adolescent Psychiatry 34 (11) : 1404-5.
30. Erwin, E., Obermayer, K. and Schulten, K. (1992) Self-organizing maps : ordering, convergence properties and energy functions. Biological Cybernetics 67 : 47-55.
31. Franchini, L., Gasperini, M., Perez, J., Smeraldi, E. and Zanardi, R. (1997) A double-blind study of long-term treatment with sertraline or fluvoxamine for prevention of highly recurrent unipolar depression. Journal of Clinical Psychiatry, 58 (3) : 104-7.
32. Frankenburg, F.R. and Kando, J.C. (1994) Sertraline treatment of attention deficit hyperactivity disorder and Tourette's syndrome. Journal of Clinical Psychopharmacology 14 (5): 359-60.
33. Fu, K.S. (1983) A step towards unification of syntactic and statistical pattern recognition. IEEE Trans. on PAMI 5 (2) : 200-205.
34. Ginberg, M.L. (1985) Does probability have a place in nonmonotonic reasoning? Proc. 9th IJCAI-85, 107-110, Los Angeles, CA.
35. Goodnick, P.J., Kumar, A., Henry, J.H., Buki, V.M. and Goldberg, R.B. (1997) Sertraline in coexisting major depression and diabetes mellitus. Psychopharmacology Bulletin 33 (2): 261-4.
36. Greist, J., Chouinard, G., DuBoff, E., Halaris, A., Kim, S.W., Koran, L., Liebowitz, M., Lydiard, R.B., Rasmussen, S., White, K., et al. (1995) Double-blind parallel comparison of three dosages of sertraline and placebo in outpatients with obsessive-compulsive disorder. Archives of General Psychiatry 52 (4): 289-95.

37.Grosof, B.N. (1986) An inequality paradigm for probabilistic knowledge. In: Uncertainty in Artificial Intelligence, ed Kanal, L.N. and Lemmer, J.F., 259-275, Amsterdam, Elsevier.
38.Gunji, Y.P., Sadaoka, H. and Ito, K. (1996) Bootstrapping system defined by inconsistent relation between Boolean and non-Boolean algebra. Appl. Math. Comput. 79 : 43-98.
39.Hellings, J.A., Kelley, L.A., Gabrielli, W.F., Kilgore, E. and Shah, P. (1996) Sertraline response in adults with mental retardation and autistic disorder. Journal of Clinical Psychiatry 57 (8): 333-6.
40.Ito, K. and Gunji, Y.P. (1994) Self-organization of living systems toward criticaly at the edge of chaos. BioSystems 33 : 17-24.
41.Kafka, M.P. (1994) Sertraline pharmacotherapy for paraphilias and paraphilia-related disorders: an open trial. Annals of Clinical Psychiatry 6 (3): 189-95.
42.Kalivas, J., Kalivas, L., Gilman, D. and Hayden, C.T. (1996) Sertraline in the treatment of neurotic excoriations and related disorders. Archives of Dermatology 132 (5): 589-90.
43.Kauffman, S.A. (1993) The Origin of Order: Self-organization and Selection and Evolution . Oxford University Press, Oxford.
44.Kauffman, S.A. (1994) Whispers from Carnot: the origins of order and principles of adaptation in complex nonequilibrium systems. In: Cowan, G.A., Pines, D. and Meltzer, D., eds. Complexity Metaphors, Models and Reality, 83-160, 1st ed. Reading, MA, Addison-Wesley
45.Kocsis, J.H., Zisook, S., Davidson, J., Shelton, R., Yonkers, K., Hellerstein, D.J., Rosenbaum, J. and Halbreich, U. (1997) Double-blind comparison of sertraline, imipramine, and placebo in the treatment of dysthymia: psychosocial outcomes. American Journal of Psychiatry, 154 (3): 390-5.
46.Kohonen, T. (1988) Content-Addressable Memories, 2nd ed. Berlin, Springer Verlag.
47.Kohonen, T. (1995/1997) Self-Organizing Maps. Berlin, Springer-Verlag. 1st ed. 1995, 2nd ed. 1997.
48.MacLane, S. (1991) Categories for the working mathematician, Berlin, Springer.
49.Maturana, H. R. and Varela, F.J. (1988) The tree of Knowledge. Boston, MA, New Science Library.
50.McConville, B.J., Minnery, K.L., Sorter, M.T., West, S.A., Friedman, L.M. and Christian, K. (1996) An open study of the effects of sertraline on adolescent major depression. Journal of Child and Adolescent Psychoparmacology 6 (1): 41-51.
51.Mendels, J., Camera, A. and Sikes, C. (1995) Sertraline treatment for premature ejaculation. Journal of Clinical Psychopharmacology 15 (5): 341-6.
52.Minsky, M.A. (1975) A framework for representing Knowledge. In: Winston, P. ed. The psycology of computer vision, 211-277, McGraw-Hill, New York.
53.Mukand, J., Kaplan, M., Senno, R.G. and Bishop, D.S. (1996) Pathological crying and laughing: treatment with sertraline. Archives of Physical Medicine & Rehabilitation 77 (12): 1309-11.

54.Murdoch, D. and McTavish, D. (1992) Sertraline. A review of its pharmacodynamic and pharmacokynetic properties, and therapeutic potential in depression and obsessive-compulsive disorder. Drugs 44 (4): 605-622
55.Papp, L.A., Weiss, J.R., Greenberg, H.E., Rifkin, A., Scharf, S.M., Gorman, J.M. and Klein, D.F. (1995) Sertraline for chronic obstructive pulmonary disease and comorbid anxiety and mood disorders. American Journal of Psychiatry 152 (10): 1531.
56.Pearl, J. (1986) Fusion, propagation and structuring in belief networks. Artificial Intelligence 29 : 241-288.
57.Pearl, J. (1988) Probabilistic Reasoning in Intelligent Systems : Network of Plausible Inference. San Mateo, CA, Morgan Kaufmann.
58.Peters, M. and Broughton, P.G.M. (1993) The role of expert systems in improving the test requesting patterns of clinicians. Ann Clin Biochem 30-52.
59.Plotkin, H. (1993) Darwin Machines and the Nature of Knowledge. Cambridge, MA, Harvard University Press.
60.Popli, A.P., Fuller, M.A. and Jaskiw, G.E. (1997) Sertraline and psychotic symptoms: a case series. Annals of Clinical Psychiatry 9 (1): 15-7.
61.Quinlan, J.R. (1993) C4.5 : Programs for machine learning. San Mateo, CA, Morgan Kaufmann.
62.Rahman, M.A. and Gregory, R. (1995) Trichotillomania associated with HIV infection and response to sertraline. Psychosomatics 36 (4): 417-8.
63.Ranen, N.G., Lipsey, J.R., Treisman, G. and Ross, C.A. (1996) Sertraline in the treatment of severe aggressiveness in Huntington's disease. Journal of Neuropsychiatry & Clinical Neurosciences 8 (3): 338-40.
64.Reiter, R. (1980) A logic for default reasoning. Artificial Intelligence 13 : 1-2, 81-132.
65.Ricca, V., Mannucci, E., Di Bernardo, M., Rizzello, S.M., Cabras, P.L. and Rotella, C.M. (1996) Sertraline enhances the effects of cognitive-behavioral treatment on weight reduction of obese patients. Journal of Endocrinological Investigation 19 (11): 727-33.
66.Ritter, H. and Kohonen, T. (1989) Self-organization semantic maps. Biological Cybernetics 61, 241-254.
67.Rubenstein, E.B. and Engel, N.L. (1996) Successful treatment of transvestic fetishism with sertraline and lithium. Journal of Clinical Psychiatry 57 (2): 92.
68.Sandell, H.S.H. and Bourne, J.R. (1985) Expert systems in medicine: a biomedical engineering prospective. Crit Rev Biomed Eng 12-95.
69.Scott, T.F., Nussbaum, P., McConnell, H. and Brill, P. (1995) Measurement of treatment response to sertraline in depressed multiple sclerosis patients using the Carroll scale. Neurological Research. 17 (6): 421-2.
70.Sel, R. (1997) Dissociation as complex adaptation. Medical Hypotheses 48 : 205-208.
71.Shannon, C.E. (1949) The Mathematical Theory of Communication. Urbana, IL, University of Illinois Press.
72.Sprenger, D. (1997) Sertraline for nocturnal enuresis. Journal of the American

Academy of Child & Adolescent Psychiatry 36 (3): 304-5.
73. Steingard, R.J., Zimnitzky, B., DeMaso, D.R., Bauman, M.L. and Bucci, J.P. (1997) Sertraline treatment of transition-associated anxiety and agitation in children with autistic disorder. Journal of Child and Adolescent Psychopharmacology 7 (1): 9-15.
74. Swartz, J.R., Miller, B.L., Lesser, I.M. and Darby, A.L. (1997) Frontotemporal dementia: treatment response to serotonin selective reuptake inhibitors. Journal of Clinical Psychiatry 58 (5): 212-6.
75. Swingler, K. (1995) Practical methods for applying neural networks to industrial and commercial systems. Master's thesis, Stirling University, Stirling.
76. Szolovits, P. and Pauker, S.G. (1978) Categorical and probabilistic reasoning in medical diagnosis. Artif Intell 11-115.
77. Thakore, J.H., Berti, C. and Dinan, T.G. (1996) An open trial of adjunctive sertraline in the treatment of chronic schizophrenia. Acta Psychiatrica Scandinavica 94 (3): 194-7.
78. Thase, M.E., Fava, M., Halbreich, U., Kocsis, J.H., Koran, L., Davidson, J., Rosenbaum, J. and Harrison, W. (1996) A placebo-controlled, randomized clinical trial comparing sertraline and imipramine for the treatment of dysthymia. Archives of General Psychiatry 53 (9): 777-84.
79. Tourigny-Rivard, M.F. (1997) Pharmacotherapy of affective disorders in old age. Can J Psychiatry 42 (suppl. 1): 10S-18S.
80. Utsugi, A. (1996) Topology selection for self-organizing maps. Network : Computation in Neural Systems 7 : 727-240, available at http ://www.aist.go.jp/NIBH/~b0616/Lab/index-e.html.
81. Volicer, L., Rheaume, Y. and Cyr, D. (1994) Treatment of depression in advanced Alzheimer's disease using sertraline. Journal of Geriatric Psychiatry & Neurology 7 (4): 227-9.
82. Wolpert, D.H. (1996) The lack of a priori distinctions between learning algorithms. Neural Computation 8 : 1341-1390.
83. Wyatt, J. (1991) Computer-based knowledge systems. Lancet, 338, 1431.
84. Yin, H. and Allison, N.M. (1995) On the distribution and convergence of feature space in self-organizing maps. Neural Computing 7 : 1178-1187.
85. Yonkers, K.A., Halbreich, U., Freeman, E., Brown, C. and Pearlstein, T. (1996) Sertraline in the treatment of premenstrual dysphoric disorder. Psychopharmacology Bulletin 32 (1): 41-6.
86. Young, C.R. (1997) Sertraline treatment of hallucinogen persisting perception disorder. Journal of Clinical Psychiatry 58 (2): 85.
87. Zanardi, R., Franchini, L., Gasperini, M., Perez, J. and Smeraldi, E. (1996) Double-blind controlled trial of sertraline versus paroxetine in the treatment of delusional depression. American Journal of Psychiatry 153 (12): 1631-3.