A Novel Large-Memory Neural Network as an Aid in Medical Diagnosis Applications

Hubert Kordylewski, Daniel Graupe, *Fellow, IEEE*, and Kai Liu

***Abstract*—This paper describes the application of a large memory storage and retrieval (LAMSTAR) neural network to medical diagnosis and medical information retrieval problems. The network is based on Minsky's knowledge-lines (k-lines) theory of memory storage and retrieval in the central nervous system. It employs arrays of self-organized map modules, such that the k-lines are implemented via link weights (address correlation) that are being updated by learning. The network also employs features of forgetting and of interpolation and extrapolation, thus being able to handle incomplete data sets. It can deal equally well with exact and fuzzy information, thus being specifically applicable to medical diagnosis where the diagnosis is based on exact data, fuzzy patient interview information, patient history, observed images, and test records. Furthermore, the network can be operated in closed loop with Internet search engines to intelligently use data from the Internet in a higher hierarchy of learning. All of the above features are shown to make the LAMSTAR network suitable for medical diagnosis problems that concern large data sets of many categories that are often incomplete and fuzzy. Applications of the network to three specific medical diagnosis problems are described: two from nephrology and one related to an emergency-room drug identification problem. It is shown that the LAMSTAR network is hundreds and thousands times faster in its training than back-propagation-based networks when used for the same problem and with exactly the same information.**

***Index Terms*—Adaptive systems, medical diagnosis, neural networks, pattern recognition, self-organizing feature maps.**

I. Introduction

THIS PAPER is concerned with the application of a large memory storage and retrieval (LAMSTAR) neural network to several medical diagnosis problems. The neural network discussed in this paper is a network specifically designed for large-scale memory storage and retrieval of information. The LAMSTAR network [1, Ch. 13]–[4] attempts to store and retrieve patterns in a computationally efficient manner, using tools of neural networks, especially self-organizing map (SOM)-based network modules, combined with statistical decision tools. The basic processing modules of the LAMSTAR network are modified Kohonen SOM modules [5]. In the LAMSTAR system, the information is stored via correlation links between individual neurons in separate SOM modules.

The input word is coded in terms of a real vector $\underline{X}$ given by

$$\underline{X} = \left[\underline{x}_1^T, \underline{x}_2^T, \ldots, \underline{x}_N^T\right]^T \tag{1}$$

where T denotes transposition.

Manuscript received July 9, 2001.

The authors are with the Department of Electrical Engineering and Computer Science, University of Illinois, Chicago, IL 60607-7053 USA.

Publisher Item Identifier S 1089-7771(01)07082-0.

In the training phase, a subset of subwords in the input word will represent the output of the network (diagnosis/decision). Each subword $\underline{x}_i$ is then channeled to a corresponding ith SOM module that stores data concerning the ith category of the input word. The network is thus organized to find a neuron in a set of neurons of a class (namely, in one SOM module) that best matches (correlates) with the input pattern.

By its structure, as described in Section II, the LAMSTAR network is uniquely suited to deal with medical diagnosis problems [2]–[4] where data are of many vastly different categories, where some categories may be missing for some patients (cases), where data are both exact and fuzzy, and where the vastness of data requires very fast algorithms. These features are rare to find, especially when coming together, in other neural networks. Still, the LAMSTAR improves its learning with increased case experience. It involves a degree of stochasticity to avoid rigidity in its decisions.

II. Outline of the LAMSTAR Network

A. Basic Structural Elements

The SOM structure employed in the LAMSTAR system adheres to fundamentals of the SOM structure, but it differs in details. Whereas in Kohonen's networks [5] all neurons of an SOM module are checked, in the LAMSTAR network, only a finite group of p neurons is checked at a time due to the huge number of neurons involved (the large memory involved). The final set of p neurons is determined by the weights (N_i), as shown in Figs. 1 and 2.

A winning neuron is determined for each input based on the similarity between the input (vector X in Fig. 2) and a weight vector W (stored information). For an input subword $\underline{X}_i$, the winning neuron is determined by minimization of a distance norm $||_*||$ given by

$$||\underline{X}_i - \underline{W}_{i,m}|| = \min ||\underline{X}_i - \underline{W}_{i,k}|| \quad \forall k \in \langle l, l+p\rangle; l \sim \{N_{i,j}\} \tag{2}$$

where

- m winning unit in ith SOM module;
- $(N_{i,j})$ weights to determine the neighborhood of top priority in module i;
- l first neuron to be scanned (determined by weights $N_{i,j}$).

B. Links Between SOM Modules (L Weights)

Information in the LAMSTAR system in encoded via correlation links $L_{i,j}$ (Figs. 1 and 2) between individual neurons in different SOM modules. The LAMSTAR system does not create

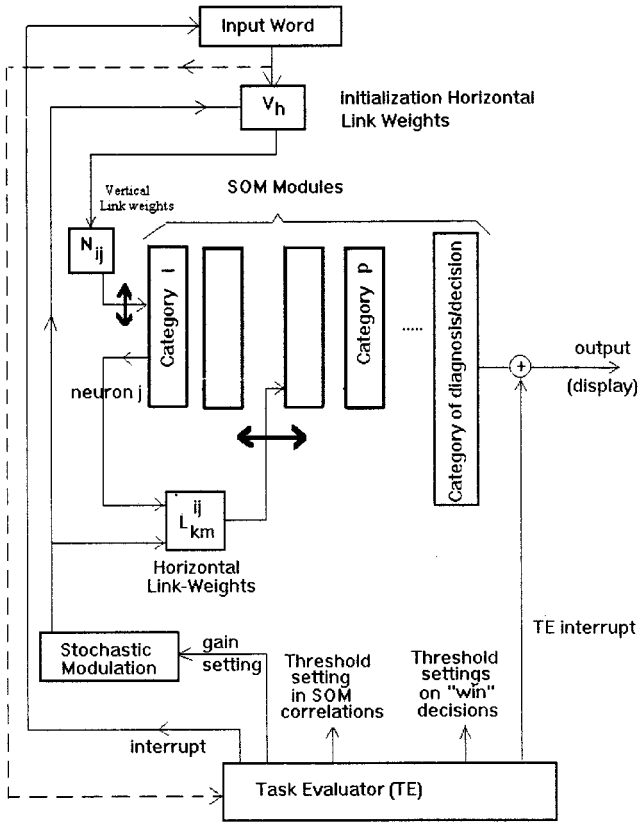

Fig. 1. General block diagram—LAMSTAR network. Task evaluation unit provides highest hierarchy of control—to modify tolerances and thresholds. Stochastic modulation unit introduces modulation noise to all settings of weights.

neurons for an entire input word. Instead, only individual subwords are stored in SOM modules (W weights), and correlations between subwords are stored in terms of creating/adjusting L links ($L_{i,j}$ in Figs. 1 and 2) that connect neurons in different SOM modules.

When the new input word is presented to the system during the training phase, the LAMSTAR network inspects all weight vectors (W_i) in SOM module i that corresponds to an input subword $\underline{x}_i$ to be store. If any stored pattern matches the input subword $\underline{x}_i$ within a preset tolerance, the system updates weights W according to the following procedure:

$$\underline{W}_{i,m}(t+1) = \underline{W}_{i,m}(t) + \alpha_i\Big(\underline{X}_i(t) - \underline{W}_{i,m}(t)\Big), \quad \text{for } m\text{: } \varepsilon_m < \varepsilon_i\,(\text{const}) \tag{3}$$

where

$\underline{W}_{i,m}(t+1)$	modified weights in module i for neuron m;
α_i	learning coefficient for module i;
ε_m	minimum error of all weight vectors W_i in module i [see (2)].

If no match was found, the system creates new pattern in the SOM module. It stores input subword $\underline{x}_i$ as a new pattern W_{in}, where index n is the first unused neuron in the ith SOM module. We repeat the above storage procedure for every input subword $\underline{x}_i$ to be stored.

Link weight values L are determined by evaluating distance minimization to determine winning neurons, where each win (successful fit) is counted by a count-up element associated with each neuron and its respective input-side links (Fig. 2) [1, Ch. 13], [6], [7]. The values of L links are modified according to

$$L^{k,m}_{i,j}(t+1) = L^{k,m}_{i,j}(t) - \alpha\left(L^{k,m}_{i,j}(t) - L_{\max}\right), \quad \text{for } L^{k,m}_{i,j}(t) > TH \tag{4a}$$

$$L^{k,m}_{i,j}(t+1) = 0, \quad \text{for } L^{k,m}_{i,j}(t) < TH \tag{4b}$$

where

TH	forgetting threshold (applies only to L weights);
$L^{k,m}_{i,j}$	links between neuron i in kth module and neuron j in mth module;
α	learning coefficient;
$L_{\max}$	maximal links value.

The count up (as the subsequent weight delay due to forgetting) set mean weight values to be stochastically modulated. Link weights $L_{i,j}$ decay over time, as a result of the learning formula of (4a) and (4b). Hence, if not chosen successfully, the appropriate $L_{i,j}$ will drop toward zero). This helps to avoid the need to consider a very large number of links, thus contributing to the network efficiency.

C. Retrieval of Information in the LAMSTAR Network

1) Input Word for Training and Information Retrieval: In applications such as medical diagnosis, the LAMSTAR system is trained for diagnosis by entering the symptoms/diagnosis pairs (or diagnosis/medication pairs). The training vectors are of the following form:

$$\underline{X} = \left[\underline{x}_1^T, \underline{x}_2^T, \ldots, \underline{x}_n^T, \underline{d}_1^T, \ldots, \underline{d}_k^T\right]^T \tag{5}$$

where $\underline{d}$ are subwords representing the output of the network (diagnosis/decision).

In the LAMSTAR's processing of data (storage and retrieval), the diagnosis subwords [$\underline{d}$ in (5)] are processed in the same manner as other subwords, namely, all punishment/reward feedbacks also apply to the diagnosis subwords. Therefore, one or more SOM module serve as output modules to output the LAMSTAR's decision/diagnosis.

The input word of (1) and (5) was shown to be a coded word (see Section II-A), comprised of subwords $(\underline{x}_i)$ that relate to various categories (input dimensions). Also, each SOM module of the LAMSTAR network corresponds to one of the categories of $\underline{x}_i$ such that the number of SOM modules equals the number of subvectors (subwords) $\underline{x}_n$ and $\underline{d}$ in $\underline{X}$ are defined by (5).

2) Channeling Weights for Fast Retrieval: In the LAMSTAR network, the input subword is channeled to only one SOM module at a time. To speed up the search process, a two-stage channeling process is employed. First, weight V_j (as in Figs. 1 and 2) determine which subword of any input word is to be the first examined, noting that from then on, the inter-module links $L_{i,j}$ will take over to consider other subwords. For application when input word categories are assumed *a priori* of equal importance, or no *a priori* information about input work categories is available, V_j weights

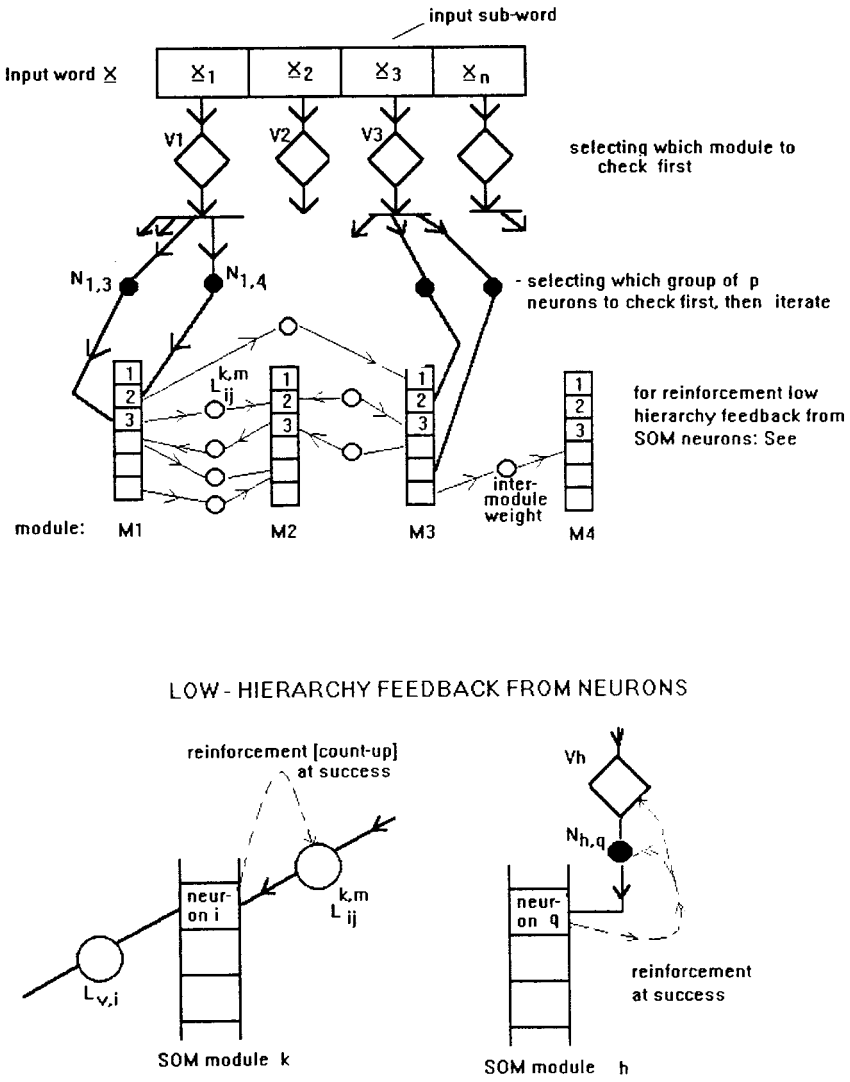

Fig. 2. Details of Fig. 1. (top) Links between SOM modules. (bottom) Low-hierarchy feedbacks from neurons that control weights N, V, and L used in the LAMSTAR.

should be modified by simple increment (reward) or decrement (punishment) functions as follows:

$$V_i(t+1) = \begin{cases} V_i(t) + \xi, & \text{for } V_i(t) < V_{\max} \\ V_i(t), & \text{for } V_i(t) \geq V_{\max} \end{cases} \quad (6)$$

where ξ denotes a small increment/decrement.

The SOM module is pseudorandom determined according to a probability distribution function (PDF) determined by weights V_j according to the following rule:

$$P_i = \frac{V_i}{\sum_{n=1}^{k} V_n} \quad (7)$$

where

P_i probability of choosing ith SOM module for initial search;

V_i weights associated with the ith SOM module;

K number of SOM modules with active inputs.

Furthermore, and is again most important to speed up this search, weights N_{ih}, as in Figs. 1 and 2 serve to assign priorities to certain neurons of the same SOM module. This is accomplished feedback based on counting (rewarding) past successes, as displayed in details in Fig. 2, to increase weights N_{ij} accordingly or to reduce it, if a "drought" has been observed in utilizing a certain memory, the latter being a forgetting feature.

3) SOM Correlations, Links, and Wandering Searches: The winning output of the SOM module [$\underline{W}_{i,m}$ in (2)] that has been reached by the above feedback process and the subsequent stochastic modulation will activate all the nodes associated with it in other SOM modules. A single concept subword automatically results in connecting to associated subwords with the highest $L_{i,j}^{k,m}$ link weights according to the following rule:

$$W_m^l = W_{m,n}, \qquad \text{for } L_{k,n}^{i,m} > L_{k,j}^{i,m} \quad \forall\, n \neq j \neq i \quad (8)$$

where

$L_{i,j}^{k,m}$ links between neuron i in kth module and neuron j in mth module;

$W_{n,k}$ weights of the neuron n (pattern) in mth SOM module.

Thus, after the activation of neurons in other SOM modules [via $L_{i,j}^{k,m}$ links of (8)] the retrieval pattern (word) would be

$$\underline{Y} = \left[\underline{W_1^l}, \ldots, \underline{W_{i-1}^l}, \underline{W_{i-1}}, \underline{W_{i+1}^l}, \ldots, \underline{W_n^l}\right] \quad (9)$$

where

$\underline{W_i^l}$ weights that results from neurons activation via $L_{i,j}^{k,m}$ links;

$\underline{W_i}$ weights of a neuron from SOM module selected by procedure described above.

D. Determination of Winners in Output Modules

The resulting retrieval [$\underline{Y}$ of (9)] is subsequently checked at the SOM level concerning correlations between stored subwords and input subwords. The links are then reinforced in cases of a successful retrieval to "strengthen" the weights of such links according to (5).

Therefore, if for a given SOM module, a winning neuron has been determined (for its corresponding input subword), and then the highest L-valued links to subwords relating to another dimension (other subword) of the input word (and which are or may be stored in another SOM module) are being examined. The final full word retrieval [$\underline{Y}$ of (9)] is than accepted/rejected by the Task Evaluation Unit of Fig. 2.

The acceptance formula is

$$\text{Test} = \begin{cases} 1, & \text{for } \sum_{i=0}^{n} z_i < \varepsilon \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

where

ε error preset by the user in 0.1–0.25-m range (where m is the number of subwords);

z_i logical fit or misfit for each subword calculated as follows:

$$z_i = \begin{cases} 1, & \text{for } ||\underline{X}_i - \underline{W}_{i,m}|| < \varepsilon_i \\ 0, & \text{otherwise} \end{cases} \quad (11)$$

where

$\underline{W}_i$ weights of the retrieved information [see (8)];

ε_1 vigilance parameter (usually in 0.1–0.25 range);

$\underline{X}_i$ weights of ith category in the input word.

Once a retrieval has been completed the feedback signals begin the adjustment of all weights (V, L, N) that were involved in the retrieval according to (4a) and (4b).

Through the L_{ij} links, the LAMSTAR network facilitate extrapolation and interpolation to previous related subwords that were missed or mistaken given an input word. As was mentioned above, in the retrieval phase, the network uses only one subword of the input word (selected via weights V). Other subwords are used only for comparison with the input pattern, and evaluation of difference by the Task Evaluation Unit [see (11)] to determine a fit between the interpolation/extrapolation of previously stored and new information. While the LAMSTAR system extrapolates entire subwords (as was shown above) via L links [i.e., via (8)], the interpolation takes place inside specific subwords, namely, inside individual SOM modules [via (2)]. The interpolation feature of the LAMSTAR system is a result of interpolation capabilities of the SOM structure employed.

III. Medical Diagnosis Applications of the LAMSTAR Network

A. General Discussion

Nearly all clinical decisions are based on many categories of data. Some categories are often fuzzy, while some are exact, and often categories pieces are missing (incomplete data sets). As mentioned in Section II-B, the LAMSTAR network can be trained with incomplete data or category sets. Therefore, due to its features, the LAMSTAR neural network is a very effective tool in such situations. The knowledge base of the system contains a mathematical extract of a series of cases with known outcome inputted to the system in the training phase. As an input, the system accepts data defined by the user, such as: 1) patient's age; 2) height; 3) weight; or 4) very specific data, as is shown in the diagnostic cases presented below. The system then builds the patient model (based on data from past experience and training) and searches the stored knowledge to find the best approximation/description to the clinical features/parameters given as input data. Thus, the LAMSTAR's function is to help the physicians to tackle a specific clinical problem by providing information in the form of, say: 1) possible diagnosis; 2) facts about a disease; 3) suggested medication; 4) medication dosage; 5) potential adverse drug reaction; 6) recommended therapy; and 7) prediction of the patient's conditions.

In medical diagnosis situations, the LAMSTAR system can be used as a: 1) teaching aid; 2) diagnosis aid; 3) tool for data analysis; 4) classification tool; and 5) prediction tool.

The LAMSTAR network can provide multidimensional analysis of input variables that can, for example: 1) assign different weights (importance) to the items of data; 2) find correlation among input variables; or 3) perform identification, recognition, and clustering of patterns. Being a neural-network algorithm, the LAMSTAR system can do all this without reprogramming per each diagnostic problem. The following subsections discuss the application of the LAMSTAR network to various medical problems and compare performance with other neural networks applied to the same problems using the same data. The examples considered below are: 1) patient diagnosis after removal of kidney stones; 2) renal cancer diagnosis; and 3) diagnosis of drug abuse in an emergency-room situation (unconscious patient). The examples presented below illustrate the scope of applications of the LAMSTAR network.

B. Case Study of ESWL Medical Diagnosis Problem

1) Problem Statement: In this example, the LAMSTAR system is applied to aid in a typical urological diagnosis problem. It evaluates a patient's condition and provides long-term forecasting after removal of renal stones via extracorporeal shock-wave lithotripsy (ESWL). The ESWL procedure breaks very large renal stones into small pieces that are then naturally removed from the kidney with the urine. Unfortunately, the large kidney stones appear again in 10%–50% of patients (1–4 years post surgery). It is difficult to predict (due to the large number of analyzed variables) with

TABLE I
INPUT DATA USED IN THE MEDICAL DIAGNOSIS-AID EXAMPLE (ESWL)

Category	Description	Parameter
1	**Age**	1 - 100
2	**Gender**	Female, Male
3	**Race**	Black, White, Hispanic, Oriental
4	**Stone Chemistry**	Calcium, Cystine
5	**Stone Location**	Parenchyma, Pelvis, Ureter, Cylyx
6	**Stone Configuration**	Staghorn, Abnormal Anatomy, Cylyceal
7	**Stone Location in Kidney**	Bilateral
8	**Acid Type and Levels**	Metabolic, Hypercalciuria/uricosuria
9	**Culture Last**	Catheter
10	**Time of the Surgery**	Time (in years)
11	**Stone partition**	Fragments, Volume
12	**Retreatment Procedure**	Medical Terms for Retreatment
13	**Medical Therapy**	Allopurinol, Thazides, Both
14	**Volume**	1 - 20 (ml)
15	**Previous Stones**	1 - 20 (# of stones)
16	**History of Prev. Stones**	Type, Other Stones
Diagnosis	**Long Term Forecast**	Success / Failure

reasonable accuracy (over 50%) if the surgery was a success or a failure. A system that correctly predicts which patients are in danger for stone recurrence after ESWL can dramatically cut down costs of treatment by reducing the need for subsequent EWSL. When the recurrence of stones is detected early enough, the very costly ESWL treatment can usually be replaced by: 1) use of medications and 2) more aggressive surveillance. The LAMSTAR system predicts success or failure of the surgery from correlation among the variables and not from the variables alone.

2) Structure and Format of the Analyzed Data: In this particular example, the input data (denoted as a "word" per each analyzed case, namely, per each patient) are divided into 16 subwords (categories). The length in bytes per each subword in this example varies from 1 to 6 bytes. The subwords describe patient's physical and physiological characteristics such as: 1) patient demographics; 2) stone's chemical composition; 3) stone location; 4) laboratory assays; 5) follow-up; 6) re-treatments; and 7) medical therapy. The system attempts to predict the result (failure/success) by analyzing the correlations among the subwords (categories) variables provided by the user. It then automatically adjusts the weights and the mapping of the correlation links accordingly (Section II-C). The system's categories for this example are defined by a category number (in this case, 1–16, as shown in Table I) with the associated meanings and parameters.

3) Test Results for Medical Diagnosis Problem: The LAMSTAR network was trained with 66 input words (patient cases), each containing actual patient data. The data were obtained from the Urology Department, University of Illinois Medical Center, Chicago, IL [8]. The system attempts to predict the treatment's outcome by analyzing the correlations among the subwords (categories) variables provided by the user. It then automatically adjusts the weighting and mapping correlation links [see (4a), (4b), and (6)].

Table II compares results of the LAMSTAR network and the back-propagation (BP) neural network [8], as applied to exactly the same training and test data sets. While both networks model the problems with high accuracy, the results show that the LAMSTAR network is over 1000 times faster in this case. The difference in training time is due to the incorporation of an unsupervised learning scheme in the LAMSTAR network, while the BP network training is based on error minimization in a 37-dimensional space (when counting elements of subword vectors), which requires over 1000 iterations.

Both networks were used to approximate calculation for the Wilks' Lambda [9], [10], which is a reflectance of importance for each input variable. The Wilks' test was used to determine which input variables are meaningful with regard to system performance. In clinical settings, the test is used to determine the importance of specific parameters in order to limit the number of patient's examination procedures. In the BP network, the Wilks' test was modeled by sequential removal of the input nodes, as well as removal of all combinations of the input nodes. The BP network was initially trained and the testing error ε_1 was recorded. Afterward, one or more input nodes were disabled

TABLE II
PERFORMANCE COMPARISON OF THE LAMSTAR NETWORK AND THE BP NETWORK FOR THE RENAL CANCER AND ESWL DIAGNOSIS

	Renal Cancer Diagnosis		ESWL Diagnosis	
	LAMSTAR Network	**BP Network**	**LAMSTAR Network**	**BP Network**
Training Time	0.08sec	65 sec	0.15 sec	177 sec
Test Accuracy	83.15 %	89.23%	85.6%	78.79%
Negative Specificity	0.818	0.909	0.53	0.68
Positive Predictive Value	0.95	0.85	1	0.65
Negative Predictive Value	0.714	0.81	0.82	0.86
Positive Specificity	0.95	0.85	1	0.83
Wilks' Test Computation time	< 15 mins	weeks	< 15 mins	Weeks

Comments:
Positive/Negative Predictive Values – ratio of the positive/negative cases that are correctly diagnosed to the positive/negative cases diagnosed as negative/positive.
Positive/Negative Specificity – the ratio of the positive/negative cases that are correctly diagnosed to the negative/positive cases that are incorrectly diagnosed as positive/negative.

and the network was retrained and retested with the initial data sets, and the resulting error was recorded as ε_2. A similar procedure was also applied to the LAMSTAR network. The two resulting errors (ε_1 and ε_2) were used to determine statistical significance of the disabled variables, namely, determine statistical significance of the increase in the errors values. Applying the Wilks' test [9], [10] to both networks to evaluate the significance of each input variable, the BP network took several weeks to run, while the LAMSTAR network required only a few minutes. This is due to the fact that, in most combinations of disabled (input) variables, the retraining of the LAMSTAR network was not necessary (since each variable is stored in a separate SOM module). For performance measurements of both networks, we also calculated: 1) test accuracy and 2) positive/negative specificity (the ratio of the positive/negative cases that are correctly diagnosed to the negative/positive cases that are incorrectly diagnosed as positive/negative, positive/negative predictive values (fraction of the positive/negative cases that are correctly diagnosed to the positive/negative cases diagnosed as negative/positive).

C. Renal Cancer Diagnosis Problem

1) Problem Statement and Structure of the Analyzed Data: The history of patients with renal cell tumors is intriguing. Some patients live for many years, while others succumb soon after surgery due to a metastatic disease. In this case study, we attempted to predict if patients will develop a metastatic disease after surgery for removal of renal-cell tumors. The input variables were grouped into sub-words describing patient's demographics, bone metastases, histologic subtype, tumor characteristics, and tumor stage.

The system's categories for this example are defined by a category number (in this case, 1–13, as shown in Table III), with the associated description and parameters.

2) Test Results for the Renal Cancer Diagnosis Problem: In this case study, we used 232 data sets (patient record), i.e., 100 sets for training and 132 for testing. The performance comparison of the LAMSTAR network versus the BP network are also summarized in Table II. As we observe, the LAMSTAR network is not only much faster to train (over 1000 times), but clearly gives better prediction accuracy (85%, as compared to 78% for BP) with less sensitivity. Applying the Wilks' test [9], [10] to both networks produced similar results, indicating that patient's age, gender, and histological subtype are the most significant variables in the prediction of the development of a metastatic disease (all mentioned variables have a p-value of the generalized likelihood ratio test <0.05). The test also show that variables that are not critical for the system's performance are: 1) tumor size and 2) bone metastases.

D. Diagnosis of Drug Abuse for Emergency Cases

1) Problem Statement: In this case study, the LAMSTAR network is used as a decision support system to identify the type of drug used by an unconscious patient who is brought to an emergency room (data obtained from Maha Noujeime, University of Illinois at Chicago). A correct and very rapid identification of the drug type will provide the emergency-room physician with the immediate treatment required under critical condi-

TABLE III
INPUT DATA USED IN THE MEDICAL DIAGNOSIS-AID EXAMPLES (RENAL CANCER)

Category	Description	Parameter
1	Race	Black, White, Hispanic, Oriental
2	Diagnosis Date	1 – 100 (Months)
3	Gender	Female, Male
4	T – Stage	Stage describing tumor size, and the spreading outside the kidney (adrenal, renal vein, vena cava, tissues outside).
5	N – Stage	Stage describing tumor size, and the spreading outside the kidney.
6	M – Stage	Stage describing tumor size, and the spreading outside the kidney.
7	Nephrectomy	Subword describing if nephrectomy was performed and the date (if available)
8	Date of the Surgery	Time (in years)
9	Lung Metastases	Present/Not present
10	Bone Metastases	Present/Not present
11	Histologic Subtype	10 subtypes of cancer histologic subtypes
12	Tumor Size	Size in cm
Diagnosis	Long Term Forecast	Treatment choice

TABLE IV
SYMPTOMS DIVIDED INTO FOUR CATEGORIES FOR DRUG-ABUSE DIAGNOSIS PROBLEM

CATEGORY 1	CATEGORY 2	CATEGORY 3	CATEGORY 4
Respiration	Pulse	Euphoria	Physical Dependence (mostly unavailable)
Temperature	Appetite	Conscious Level	Psychological Dependence (mostly unavailable)
Cardiac Arrhythmia	Vision	Activity Status	Duration of action (mostly unavailable)
Reflexes	Hearing	Violent Behavior	Method of Administration (needle/oral)
Saliva Secretion	Constipation	Convulsions	Urine Drug Screen (mostly unavailable)

Comments:
1. Each category is a vector and thus constitutes a subword (the input word having four subwords).
2. Unavailable data are empty (missing data).
3. Category 4, being mostly empty(unavailable), can be estimated by the interpolative feature of the LAMSTAR network, as it can be trained off-line.

tions, whereas wrong or delayed identification may prove fatal and when no time can be lost, while the patient is unconscious and cannot help in identifying the drug. The LAMSTAR system can diagnose to distinguish between five groups of drugs: alcohol, cannabis (marijuana), opiates (heroin, morphine, etc.), hallucinogens (LSD), and central nervous system (CNS) stimulants (cocaine) [11]. In the drug-abuse identification problem, diagnosis cannot be based on one or two symptoms since, in most cases, the symptoms overlap. The drug-abuse identification is a very complex problem since most of the drugs can cause opposite symptoms depending on additional factors like: 1) regular/periodic use; 2) high/low dose; and 3) time of intake [11]. The diagnosis is based on a complex relation between 21 input variables arranged in four categories (subword vectors) representing drug-abuse symptoms. Most of these variables are easily detectable in an emergency-room setting by simple evaluation (Table IV). The large number of variables often makes it difficult for a doctor to properly interrelate them under emer-

gency-room conditions for a correct diagnosis. An incorrect diagnosis and subsequent incorrect treatments may be lethal to a patient. For example, while cannabis and cocaine require different treatments, when analyzing only the mental state of the patient, both cannabis and large doses of cocaine can result in the same mental state classified as mild panic and paranoia.

Furthermore, often not all variables can be evaluated for a given patient. In an emergency-room setting, it is impossible to determine all 21 symptoms, and there is no time for either a urine test or other drug tests.

2) Results: The LAMSTAR network was trained with 300 sets of simulated input data of the kind considered in actual emergency-room situations. The testing of the network was performed with 300 data sets (patient cases), some of which have incomplete data (in an emergency-room setting, there is no time for urine or other drug tests). Due to the specific requirements of the drug-abuse identification problem (abuse of cannabis should never be mistakenly identified as any other drug), the training of the system consisted of two phases. In the first phase, 200 training sets were used for unsupervised training, followed by the second phase, where 100 training sets were used in online supervised training with punishment coefficients, shown in (4a), (4b), and (6), increased when cannabis was incorrectly identified.

The LAMSTAR network successfully recognized 100% of cannabis cases, 97% of CNS stimulants, and hallucinogens (in all incorrect identification cases, both drugs were mistaken with alcohol), 98% of alcohol abuse (2% incorrectly recognized as opiates), and 96% of opiates (4% incorrectly recognized as alcohol).

IV. Conclusions

In this paper, we have shown that the LAMSTAR neural network provides not only very fast memory retrieval and efficient storage, but also interpolation and extrapolation of input data based on stored information. Furthermore, this is accomplished in the face of possibly incomplete data, which is equally well processed if it is exact or fuzzy. The fast retrieval capability of the network results partly from the need to analyze only one sub-word of the input word at a time. The rest of the sub-words are retrieved by analyzing not the input word, but correlation links (link weights) associated with the sub-word. The implementation of pseudorandom modulation of link weights prevents the network from getting stuck. In the retrieval process, the network not only extrapolates the input word, but also interpolates the input in cases of minor variations in data. As we have shown in the studied cases, the performance (success rate) of the LAMSTAR neural network is as good (or better) than the performance of the widely used BP neural network, while the training time is reduced by a factor of over 1000 and, in more complicated cases (more sub-words or categories per case), by a factor of a few thousand or more. Thus, the LAMSTAR network can be very effective in problems where the training domain is not well defined, and where it is difficult to create reliable training sets, which is exactly the situation one faces in cases of medical diagnosis. Furthermore, conventional networks such as BP require complete data sets.

References

[1] D. Graupe, *Principles of Artificial Neural Networks*, Singapore: World Scientific, 1997.
[2] D. Graupe and H. Kordylewski, "A large memory storage and retrieval neural network for adaptive retrieval and diagnosis," *Int. J. Software Eng. Knowledge Eng.*, vol. 8, no. 1, pp. 115–138, 1998.
[3] H. Kordylewski and D. Graupe, "Applications of the LAMSTAR neural network to medical and engineering diagnosis/fault detection," in *Proc. 7th Artificial Neural Networks in Eng. Conf.*, St. Louis, MO, 1997.
[4] H. Kordylewski, D. Graupe, and K. Liu, "Medical diagnosis applications of the LAMSTAR neural network," in *Proc. Biol. Signal Interpretation Conf.*, Chicago, IL, 1999.
[5] T. Kohonen, *Self-Organizing and Associative Memory*, 2nd ed. Berlin, Germany: Springer-Verlag, 1988.
[6] D. Graupe and W. J. Lynn, "Some aspects regarding mechanistic modeling of recognition and memory," *Cybernetica*, vol. 3, pp. 119–141, 1970.
[7] M. L. Minsky, "K-lines: A theory of memory," *Cogn. Sci.*, vol. 4, pp. 117–133, 1980.
[8] C. S. Niederberger *et al.*, "A neural computational model of stone recurrence after ESWL," in *Int. Eng. Appl. Neural Networks Conf.*, 1996, pp. 423–426.
[9] D. F. Morrison, *Multivariate Statistical Methods*. New York: McGraw-Hill, 1996, p. 222.
[10] S. Wilks, "The large sample distribution of the likelihood ration for testing composite hypothesis," *Annu. Math. Stat.*, vol. 9, pp. 2–60, 1938.
[11] L. J. Bierut *et al.*, "Familiar transmission of substance dependence: Alcohol, marijuana, cocaine, and habitual smoking," *Arch. Gen. Psychiatry*, vol. 55, no. 11, pp. 982–988, 1998.

Hubert Kordylewski received the Engineer degree in electrical engineering from the Warsaw University of Technology, Warsaw, Poland, in 1992, and the M.S. and Ph.D. degrees in electrical engineering from the University of Illinois at Chicago, in 1995 and 1999, respectively.

He is currently a Professor and Research Director at the Graduate School, Knowledge System Institute, University of Illinois at Chicago. His research interests involve applications of artificial neural networks to signal processing and data modeling.

Daniel Graupe (M'71–SM'83–F'85) received the B.S.M.E., B.S.E.E., and Dipl. Eng. in electrical engineering degrees from the Technion, Israel Institute of Technology, and the Ph.D. degree in electrical engineering from the University of Liverpool, Liverpool. U.K.

He is currently a Professor of electrical engineering and computer science, Professor of bioengineering, and Adjunct Professor of physical medical and rehabilitation at the University of Illinois at Chicago. He was Bodine Chair Professor and Distinguished Professor of electrical engineering at the Illinois Institute of Technology. He has authored four textbooks, one of which was translated into Russian and Servo-Kroatian, and another into Chinese. He has also authored chapters in several other books, over 85 journal papers, and over 100 conference papers.

Dr. Graupe served as Associate Editor of the IEEE Transactions on Circuits and Systems (1989–1991). He was chairman of the IEEE Circuits and Systems (CAS) Technology Committee on Medical Image and Signal Processing (1987–1991).

Kai Liu, photograph and biography not available at time of publication.