Azuaje F. "Discovering Genome Expression Patterns With Self-Organizing Neural Networks", in *Understanding and Using Microarray Analysis Techniques: A Practical Guide*, Berrar D, Dubitzky W and Granzow M, editors, London: Springer Verlag, 2002.

Chapter 15

DISCOVERING GENOMIC EXPRESSION PATTERNS WITH SELF-ORGANIZING NEURAL NETWORKS

Francisco Azuaje
Department of Computer Science, University of Dublin – Trinity College, Ireland, e-mail: Francisco.Azuaje@cs.tcd.ie

1. INTRODUCTION

Self-organizing neural networks represent a family of useful clustering-based classification methods in several application domains. One such technique is the *Kohonen Self-Organizing Feature Map* (SOM) (Kohonen, 2001), which has become one of the most successful approaches to analysing genomic expression data. This model is relatively easy to implement and evaluate, computationally inexpensive and scalable. In addition, it exhibits significant advantages in comparison to other options. For instance, unlike hierarchical clustering it facilitates an automatic detection and inspection of clusters. Unlike Bayesian-based clustering it does not require prior hypotheses or knowledge about the data under consideration. Compared to the *k-means* clustering algorithm, the SOM exemplifies a robust and structured classification process.

Self-organizing neural networks are based on the principle of transforming a set of *p-variate* observations into a spatial representation of smaller dimensionality, which may allow a more effective visualization of correlations in the original data. Murtagh and Hernández-Pajares (1995), among many others, have discussed the connections between SOMs and alternative data analysis techniques. Before its introduction to the area of functional genomics, SOMs had been extensively applied in different biomedical decision support tasks, including coronary heart risk assessment (Azuaje et al., 1998), electrocardiogram-based diagnostic studies (Papadimitriou et al., 2001) and tissue characterization in cancer studies (Schmitz et al., 1999).

Scientists may use SOMs to detect clusters of similar expression patterns. The SOM-based model was one of the first machine learning techniques implemented for the molecular classification of cancer. Golub and colleagues (1999) reported a model to discover the distinction between acute myeloid leukemia and acute lymphoblastic leukemia. The application of SOMs was part of a systematic expression monitoring method based on DNA microarrays. They were able to illustrate not only a classification process to distinguish known categories of leukemia samples, but also a class discovery process to identify unknown relevant subtypes. The authors suggested that it would be possible to achieve a sub-classification of higher resolution with a larger sample collection. Moreover, this classification technique may provide the basis for the prediction of clinical outcomes, such as drug response or survival. This research is a good example of how a SOM-based classifier together with other statistical tools may support a complex knowledge discovery function.

Another relevant study consisted of the application of SOMs to organize thousands of genes into biologically relevant clusters using hematopoietic differentiation data (Tamayo et al., 1999). This classification system indicated, for example, genes involved in differentiation therapy used in the treatment of leukemia. It discussed some of the key attributes that make the SOM an adequate clustering technique for expression data. It shows how SOMs can primarily be used to perform exploratory data analysis and facilitate visualisation-based interpretations. The authors developed *Genecluster*, which is a computer package to perform SOM-based classification of genomic expression data. It has assisted, for instance, the generation of interpretations relating to the yeast cell cycle, macrophage differentiation in HL-60 cells and hematopoietic differentiation across different cell lines (Tamayo et al., 1999).

Ideker and colleagues (2001) also used SOMs in an integrated approach to refining a cellular pathway model. Based on this method they identified a number of mRNAs responding to key perturbations of the yeast galactose-utilization pathway.

The remainder of this chapter addresses two important questions on self-organizing neural networks applications for expression data: a) How do these systems work?, and b) how can we use them to support genomic expression research?. It focuses on the application of SOMs in different expression data analysis problems. Advantages and limitations will be discussed. Moreover, an alternative solution based on the principle of adaptive self-organization will be introduced. This chapter will end with an overview of current challenges and opportunities.

Azuaje F. "Discovering Genome Expression Patterns With Self-Organizing Neural Networks", in *Understanding and Using Microarray Analysis Techniques: A Practical Guide*, Berrar D, Dubitzky W and Granzow M, editors, London: Springer Verlag, 2002.

2. SOMs AND MICROARRAY DATA ANALYSIS

The SOM is based on hypothetical neural structures called feature maps, which are configured and adapted by the effect of sensory signals or data observations (Kohonen, 2001). Their processing components, known as *neurones, prototypes* or *cells*, are spatially correlated after completing a learning or training process, such that those prototypes at nearby points on the resulting structure are more similar than those widely separated. Each prototype is associated with a weight vector $\mathbf{m}_i$. Thus, SOMs can be used to perform clustering functions (Murtagh and Hernández-Pajares, 1995). Figure 15.1 shows a typical SOM.

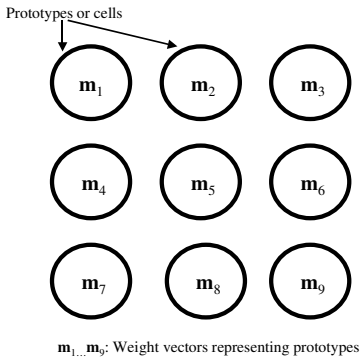

Figure 15.1. A typical SOM

2.1 The SOM Clustering Algorithm

The SOM learning algorithm transforms any p-dimensional space into an ordered two-dimensional coordinate system. Also one may say that the SOM algorithm implements a "nonlinear projection" of the probability density function, $p(\mathbf{x})$, of the input data vector $\mathbf{x}$ onto a two-dimensional space (Kohonen, 2001).

Given a number of samples, N, each one represented by a number of features, p, a Kohonen map (Kohonen, 2001) consists of a grid of k prototypes, $\mathbf{m}_j \in \Re^p$ (vector defined by p elements) (Figure 15.1). The main goal is then to define associations between each sample or observation and the prototypes represented on the map. The number of prototypes, k, and other learning parameters need to be defined by the user. Before starting the learning process the prototypes $\mathbf{m}_j$ are randomly initialized. Each of the k prototypes, $\mathbf{m}_j$, may also be encoded with respect to an integer coordinate pair $r_j \in Q_1 \otimes Q_2$. Where $Q_1 = \{1,\ldots, q_1\}$, $Q_2 = \{1,\ldots, q_2\}$ and $k = q_1 \times q_2$. Figure

15.2 illustrates a SOM consisting of 9 prototypes, which are used to categorise a number of samples. The SOM learning process is summarised as follows.

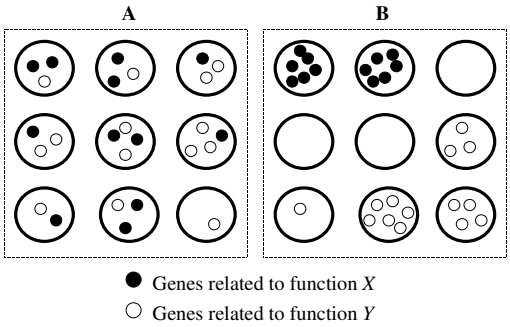

Figure 15.2. A SOM network before (panel A) and after (panel B) performing a learning process, based on a hypothetical data set of expression profiles linked to two classes of genes. The right panel indicates that the algorithm has successfully separated the classes under consideration

Each observation, $\mathbf{x}_i$, is processed one at a time. The first step in each *learning cycle* is to find the closest prototype $\mathbf{m}_j$ to $\mathbf{x}_i$ using, for example, the Euclidean distance in $\Re^p$. Then for all neighbours $\mathbf{m}_k$ of $\mathbf{m}_j$, the idea is to make $\mathbf{m}_k$ closer to $\mathbf{x}_i$, based on the following formula:

$$\mathbf{m}_{k\text{-new}} = \mathbf{m}_k + \alpha \times (\mathbf{x}_i - \mathbf{m}_k) \qquad \mathbf{m}_k \in N_j \qquad (1)$$

Where $\mathbf{m}_{k\text{-new}}$ represents the new value for $\mathbf{m}_k$, α is called the learning rate, and N_j represents the neighbourhood of $\mathbf{m}_j$, which always includes $\mathbf{m}_j$.

The main purpose of equation (1) is not only to move the SOM prototypes closer to the data, but also to develop a smooth spatial relationship between the prototypes. This process is summarized in Figure 15.3.

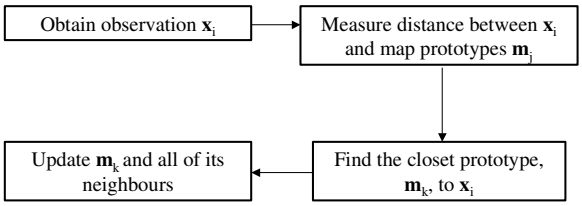

Figure 15.3. The SOM learning algorithm: a single learning cycle

The neighbours of $\mathbf{m}_j$ are defined to be all $\mathbf{m}_k$, such that the distance between r_j and r_k is *small*. Commonly this is calculated using the Euclidean distance, and *small* is defined by a threshold value, *Th*. The selection of the

size of N_j is crucial to achieve a proper clustering process. For example, if the neighbourhood is too small at the beginning of the learning process, the SOM will not be ordered globally. Thus, one can initiate it with a fairly wide N_j and let its size (threshold *Th*) decrease linearly during the learning process.

The performance of the SOM learning algorithm strongly depends on the selection of the learning rate, α. Typically α is linearly decreased from 1 to 0 over a few thousand learning cycles. For more information on the design principles of the SOM, the reader is referred to (Kohonen, 2001).

A SOM can also be seen as a constrained version of the k-means clustering algorithm. If we define *Th* small enough such that each neigbourhood contains only one prototype, then the spatial interrelation between the prototypes is not achieved. In that case it is possible to demonstrate that the SOM algorithm is a version of the k-means clustering method, which stabilizes at one of the local minima found by the k-means (Hasti et al., 2001).

Figure 15.2 illustrates a hypothetical situation, in which two types of genes, each one associated with a different biological function, are clustered based on their expression profiles. Panel A of Figure 15.2 shows a SOM at the very beginning of the learning process, while panel B portrays the clusters formed after completing a learning process. The prototypes are represented by circles, and the genes that are linked to each prototype are depicted randomly within the correspondent circle. One may, for example, run the algorithm during 2600 learning cycles through this data set of 26 genes (100 cycles for each gene), and let *Th* and α decrease linearly over the 2600 iterations. This example depicts a case in which a SOM network has successfully detected a class structure in a data set, which may allow one to differentiate its samples in terms of their patterns of functional similarity.

Once a SOM has been properly trained, one can use it to classify an unknown observation, which can also be referred to as a *testing* sample. In this situation the prediction process consists of identifying the closest SOM prototype to the sample under consideration, and use that prototype as its class or cluster predictor.

The following sub-section illustrates the application of the SOM to a genomic expression classification problem.

2.2 Illustrating Its Application

By way of example, this technique is first tested on expression data from a study on the molecular classification of leukemias. The data analysed consisted of 38 bone marrow samples: 27 acute lymphoblastic leukemia (ALL) and 11 acute myeloid leukemia (AML) samples. Each sample is described by the expression levels of 50 genes with suspected roles in this disease. These data were obtained from a study published by Golub and co-workers (1999). The original data descriptions and experimental protocols

can be found at the *MIT Whitehead Institute* Web site (http://www.genome.wi.mit.edu/MPR).

The data were normalised such that the mean and variance of the genes are set to 0 and 1 respectively, which is the traditional pre-processing method used in expression analysis. The SOM networks were trained with 3800 learning cycles. The initial value of the learning parameter α was equal to 0.1 in all of the clustering experiments. Both values for α and *Th* were linearly decreased during the learning processes.

Figure 15.4 displays the clustering results based on a SOM network, which is defined by two prototypes: A and B. All of the AML samples were grouped by prototype A. The samples belonging to the class ALL were assigned to prototype B, except two of them that were located in the first prototype. This configuration indicates that the cluster defined by the prototype A is representative of the class AML, and the cluster defined by the prototype B is associated with the class ALL. Therefore, one may argue that this learning process was able to distinguish between the classes ALL and AML based on the expression values of 50 genes (Golub et al. 1999).

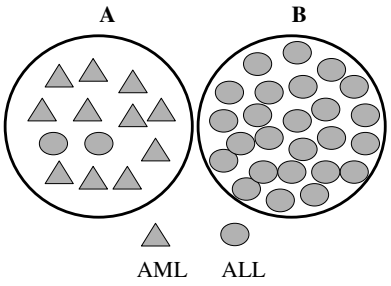

Figure 15.4. Expression data clustering using the SOM: two clusters of AML and ALL samples

This type of clustering technique may also be used to predict the existence of subclasses or discover unknown categories. Figure 15.5 displays the clustering results based on 4 prototypes, which were used to categorise the same leukemia data set.

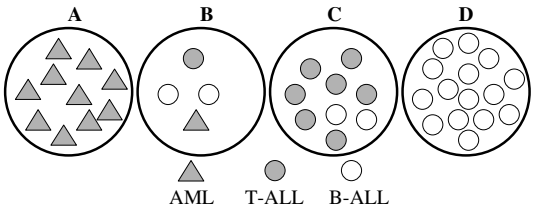

Figure 15.5. Expression data clustering using the SOM: Four clusters of AML and ALL samples. Clusters C and D are associated with two subtypes of ALL samples

These results again suggest that it is possible to distinguish AML from ALL samples. AML samples are encoded by prototype A, except one that was included in cluster B. Clusters B, C and D include the samples belonging to the ALL class. A previous systematic study of these data demonstrated that the ALL samples may indeed be classified into two subtypes: T-ALL and B-ALL (Golub et al., 1999). The SOM clustering results depicted in Figure 15.5 offers a useful insight into the existence of those subclasses. Based on the composition of the clusters obtained in Figure 15.5, one may point out, for example, that cluster C can be labelled as the T-ALL cluster, while cluster D identifies the samples belonging to B-ALL.

A second example deals with the molecular classification of diffuse large B-cell lymphoma (DLBCL) samples. The data consisted of 63 cases (45 DLBCL and 18 normal) described by the expression levels of 23 genes with suspected roles in processes relevant to DLBCL (Alizadeh et al., 2000). These data were obtained from a study published by Alizadeh and colleagues (2000), who identified subgroups of DLBCL based on the systematic analysis of the patterns generated by a specialized cDNA microarray technique. The full data and experimental methods are available on the Web site of their research group (http://llmpp.nih.gov/lymphoma).

In this case a SOM network was trained with 12,600 learning cycles, and the other learning parameters were defined as above. The data were normalised such that the mean and variance of the genes are set to 0 and 1 respectively. Figure 15.6 shows the clustering results based on two prototypes A and B.

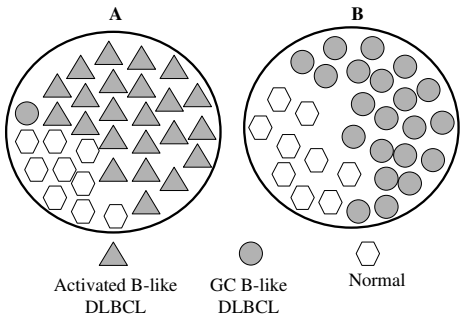

Figure 15.6. Expression data clustering using the SOM: Distinguishing subtypes of DLBCL

Because both clusters A and B include Normal samples, this clustering configuration does not clearly distinguish Normal from DLBCL samples. The reader is referred to (Alizadeh et al., 2000) for a discussion on the relationships between the Normal and DLBCL samples in terms of their expression patterns, which are indicative of different stages of B-cell differentiation. Nevertheless, these clustering results represent relevant information to recognise the two subtypes of DLBCL reported by Alizadeh et al. (2000): Activated B-like DLBCL and germinal centre B-like DLBCL (GC B-like DLBCL). In this case, cluster A can be labelled as the cluster representing Activated B-like DLBCL samples, and cluster B may be used to identify GC B-like DLBCL.

This section has dealt with the implementation and application of SOM networks for the analysis of expression data. The following section introduces some modifications to the original SOM, which may be useful to facilitate a knowledge discovery task based on this type of data.

3. SELF-ADAPTIVE AND INCREMENTAL LEARNING NEURAL NETWORKS FOR MICROARRAY DATA ANALYSIS

A number of research efforts have addressed some of the pattern processing and visualisation limitations exhibited by the original SOM. It has been shown how these limitations have negatively influenced several data mining, visualisation and cluster analysis applications (Alahakoon et al., 2000). A SOM system requires the user to predetermine the network structure and the number of prototypes. This trial-and-error task may represent a time-consuming and complex problem. Another important limitation is the lack of tools for the automatic detection of cluster boundaries. Different approaches

have been proposed to improve the original SOM algorithm. Investigations have suggested the application of *self-adaptive and incremental learning neural networks* (SANN), instead of static topology networks in order to improve several data classification applications (Nour and Madey, 1996), (Fritzke, 1994).

Some of these approaches aim to determine the prototype composition, shape and size of the self-organizing structure during the learning process. These learning techniques are well adapted to application domains, such as expression analysis, which are characterised by incomplete data and knowledge.

Recent advances include a neural network model known as *Double Self-Organizing Map* (Su and Chang, 2001), which has been suggested for data projection and reduction applications. The *Fast Self-Organizing Feature Map* algorithm (Su and Chang, 2000) aims to automatically reduce the number of learning cycles needed to achieve a proper clustering process. Other authors have proposed to combine the SOM approach and advanced supervised learning techniques. One example is the *Supervised Network Self- Organizing Map* (sNet-SOM) (Papadimitriou et al., 2001). In this case a variant of SOM provides a global approximation of a data partition, while a supervised learning algorithm is used to refine clustering results in areas categorised as ambiguous or more critical for discovery purposes. Other models designed to implement automatic map generation and cluster boundary detection include the *Growing Cell Structure Network* (GCS) (Fritzke, 1994), the *Incremental Grid Growing Neural Network* (IGG) (Blackmore, 1995) and the *Growing Self-Organizing Map* (GSOM) (Alahakoon et al., 2000). The following subsection illustrates the application of one of these techniques to the problem of recognising relevant genomic expression patterns.

3.1 A GCS-Based Approach To Clustering Expression Data

GCS is an adapted version of the SOM, which has been applied to improve a number of pattern recognition and decision support systems (Azuaje et al., 1999), (Azuaje et al., 2000). One type of GCS can be described as a two-dimensional space, where its prototypes are inter-connected and organised in the form of triangles. An initial topology for the GCS is organised as one two-dimensional triangle (Figure 15.7.a). The connections between cells reflect their separation distance on the prototype space. Like in the original SOM each cell is represented by a weight vector $\mathbf{m}_i$, which is of the same dimension as the input data. At the beginning of the learning process the weight vectors are assigned random values. The learning process comprises the processing of input vectors and the adaptation of weight vectors, $\mathbf{m}_i$. But unlike the SOM there is no need to define prototype neighbourhoods. Moreover, the learning rate, α, is substituted by two constant

values, ε_w and ε_n, which represent the learning rates for the closest prototype to a sample (winning cell) and its neighbours respectively. The value of these learning rates ranges between 0 and 1.

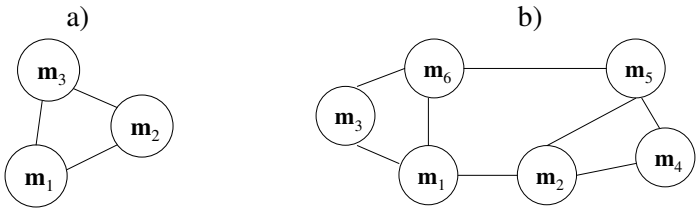

Figure 15.7. Growing Cell Structures. a) An initial topology of GCS. b) A GCS topology after a number of learning cycles

GCS also performs an adaptation of the overall structure by inserting new cells into those regions that represent large portions of the input data (Fritzke, 1994). Also, in some cases, when one is interested in more accuracy or when the probability density of the input space consists of several separate regions, a better modelling can be obtained by removing those cells that do not contribute to the input data classification. This adaptation process is performed after a number of learning cycles. Figure 15.7.b depicts a typical GCS after performing a number of learning cycles. The reader is referred to (Fritzke, 1994) for a complete description of this algorithm. Section 4 discusses some of the advantages and limitations of this type of models.

In order to exemplify some of the differences between the SOM and the GCS clustering models the hypothetical classification problem described in Section 2.1 is retaken. Panel A of Figure 15.8 depicts the results that one may have obtained using a standard SOM, whose shape and size were defined by the user. Panel B of the same figure portrays the type of results that one may expect from a GCS clustering model. In this situation the insertion and deletion of cells allowed the categorisation of the two types of genes into two separated regions of cells. Thus, one major advantage is the automatic detection of cluster boundaries. Moreover, the distance between cells may be used as a measure of similarity between groups of genes.

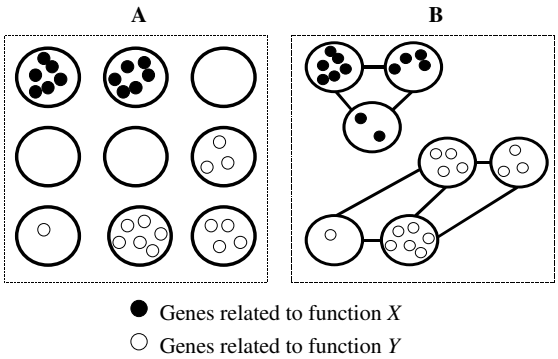

Figure 15.8. Comparing SOM-based (panel A) and GCS-based (panel B) clustering, using the hypothetical classification example introduced in Section 2.1

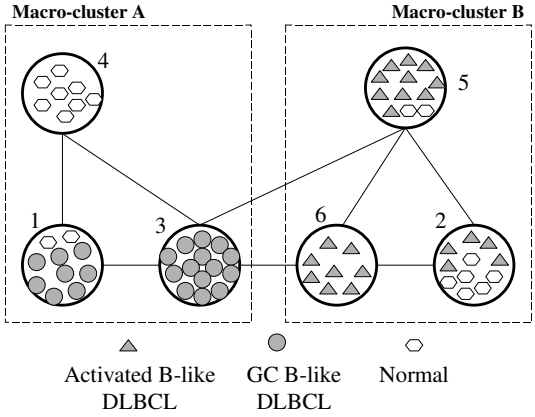

Figure 15.9. Expression data clustering using GCS: Distinguishing subtypes of DLBCL

Figure 15.9 shows the clusters obtained using a GCS and the DLBCL expression data presented in Section 2.2. The GCS network was trained with 2500 input presentation epochs (2500 x 63 learning cycles), inserting a new cell every 500 epochs and deleting irrelevant cells every 1000 epochs. The learning parameters, ε_w and ε_n, were equal to 0.095 and 0.010 respectively. For a complete description of this and other experiments the reader is referred to (Azuaje, 2001).

The resulting GCS consists of 6 cells or clusters containing the normal and DLBCL samples. The cell connections shown in Figure 15.9 do not reflect the weight vector distances. It shows that each cell corresponds to a representative cluster of the normal and DLBCL classes. For instance, cells 4 and 6 categorise only normal and DLBCL samples respectively. The majority of the samples recognised by Cells 1 and 5 belong to the class DLBCL. Cell 3 recognises samples belonging only to the category DLBCL. Cells 1 and 3

comprise all of the GC B-like DLBCL subjects. Cells 2, 5 and 6 represent the clusters encoding the Activated B-like DLBCL subjects. Thus, this GCS network consists of two regions or *macro-clusters*, A and B, which identify the GC B-like and the Activated B-like DLBCL subjects respectively. Unlike the results obtained from the SOM-based clustering, the GCS was also able to separate normal from DLCBL samples (Cell 4). Further descriptions and experimental procedures can be implemented to validate the statistical (Azuaje, 2001) and biomedical significance of these results.

4. DISCUSSION

This chapter has introduced the application of self-organizing neural networks for the analysis of genomic expression data. Several studies have suggested the SOM model as a basic approach to expression clustering (Section 2). Some of its advantages were illustrated and alternative solutions based on advanced principles of network self-organization were overviewed. It has been indicated that the application of SANN (Section 3) may support a deeper comprehension of a pattern discovery problem. This chapter has illustrated how a SANN model called GCS may be implemented to specify interesting molecular patterns or confirm known functional categories.

SANN systems, such as GCS, offer several advantages in relation to the SOM and other expression data classification techniques. In contrast to the SOM, SANN structures are determined automatically from the expression data. Most of these models do not require the definition of time-dependence of decay schedule parameters. SANN's ability to insert and delete cells allows a more accurate estimation of probability densities of the input data. Its capacity to interrupt a learning process or to continue a previously interrupted one, permits the implementation of incremental clustering systems. SANN have demonstrated its strength to process both small and high dimensionality data in several application domains (Alahakoon et al., 2000), (Azuaje et al., 2000), (Papadimitriou et al., 2001). Some SANN may be implemented in either unsupervised or supervised learning modes (Fritzke, 1994), (Papadimitriou et al., 2001). However, there are important limitations that need to be addressed. For example, in the GCS model there is not a standard way to define a priori the number of learning cycles and the exact number of cells required to properly develop a network. Some models, such the GSOM (Alahakoon et al., 2000), partially address this problem by introducing *spread factors* to measure and control the expansion of a network. In a number of applications it has been shown that techniques, like GCS and IGG, may be more susceptible to variations in the initial parameter settings than the SOM clustering model (Blackmore, 1995), (Köhle and Merkl, 1996).

There are additional problems that merit further research in order to contribute to the advance of clustering-based genomic expression studies. Among them: The implementation of hierarchical clustering using SANN,

faster clustering algorithms, specialised techniques for the processing of time-dependent or statistically-dependent data, and methods to automatically measure the contribution of a variable to the clustering results.

It is crucial to develop frameworks to assist scientists during the design and evaluation of clustering applications. Some of such guidelines and methods were examined in Chapter 13. Evaluation techniques may support not only the validation of clusters obtained from SOM, SANN or any other procedures, but also they may enable an effective and inexpensive mechanism for the automatic description of relevant clusters.

REFERENCES

Alahakoon D., Halgamuge S.K., Srinivasan B. Dynamic self-organizing maps with controlled growth for knowledge discovery. IEEE Transactions on Neural Networks 2000; 11: 601-614.

Alizadeh A.A., Eisen M.B., Davis R.E., Ma C., Lossos I.S., Rosenwald A., Boldrick J.C., Sabet H., Tran T., Yu X., Powell J.I., Yang L., Marti G.E., Moore T., Hudson J., Lu L., Lewis D.B., Tibshirani R., Sherlock G., Chan W.C., Greiner T.C., Weisenburger D.D., Armitage J.O., Warnke R., Levy R., Wilson W., Grever M.R., Bird J.C., Botstein D., Brown P.O., Staudt M. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000; 403:503-511.

Azuaje F., Dubitzky W., Lopes P., Black N., Adamson K., Wu X., White J. Discovery of incomplete knowledge in electrocardiographic data. Proceedings of the Third International Conference on Neural Networks and Expert Systems in Medicine and Healthcare; 1998 September 2-4; Pisa. World Scientific: Singapore, 1998.

Azuaje F., Dubitzky W., Black N., Adamson K. Improving clinical decision support through case-based fusion. IEEE Transactions on Biomedical Engineering 1999; 46: 1181-1185.

Azuaje F., Dubitzky W., Black N., Adamson K. Discovering relevance knowledge in data: a growing cell structure approach. IEEE Transactions on Systems, Man and Cybernetics, Part B: Cybernetics 2000; 30: 448-460.

Azuaje F. An unsupervised neural network approach to discovering gene expression patterns in B-cell lymphoma. Online Journal of Bioinformatics 2001; 1: 23-41.

Blackmore J. Visualizing high-dimensional structure with the incremental grid growing neural network. M.S. thesis, University of Texas at Austin, 1995.

Fritzke B. Growing cell structure--a self-organizing network for unsupervised and supervised learning. Neural Networks 1994; 7: 1441-1460.

Golub T.R., Slonim D.K., Tamayo P., Huard C., Gassenbeck M., Mesirov J.P., Coller H., Loh M.L., Downing J.R., Caligiuri M.A., Bloomfield C.D., Lander E.S. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999; 286:531-537.

Hasti, Trevor, Tibshirani, Robert, Friedman, Jerome, *The Elements of Statistical Learning*. NY: Sprineger, 2001.

Ideker T., Thorsson V., Ranish J.A., Christmas R., Buhler J., Eng J.K., Bumgarner R., Goodlett D.R., Aebersol R., Hood L. Integrated genomic and proteomic analyses of a systematically perturbated metabolic network. Science 2001; 292:929-933.

Köhle M., Merkl D. Visualizing similarities in high dimensional input spaces with a growing and splitting neural network. Proceedings of the International Conference of Artificial Neural Networks (ICANN'96); 1996; 581-586.

Kohonen, Teuvo, *Self-Organizing Maps*. Berlin: Sprineger, 2001.

Murtagh F., Hernández-Pajares M. The Kohonen self-organizing map method: an assessment. Journal of Classification 1995; 12:165-190.

Nour M.A., Madey G.R. Heuristic and optimization approaches to extending the Kohonen self organizing algorithm. European Journal of Operational research 1996; 93: 428-448.

Papadimitriou S., Mavroudi S., Vladutu L., Bezerianos A. Ischemia detection with a self-organizing map supplemented by supervised learning. IEEE Transactions on Neural Networks 2001; 12: 503-515.

Schmitz G., Ermert H., Senge T. Tissue-characterization of the prostate using radio frequency ultrasonic signals. IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control 1999; 46: 126-138.

Su M.C, Chang H.T. Fast self-organizing feature map algorithm. IEEE Transactions on Neural Networks 2000; 11: 721-733.

Su M.C, Chang H.T. A new model of self-organizing neural networks and its application in data projection. IEEE Transactions on Neural Networks 2001; 12:153-158.

Tamayo P., Slonim D., Mesirov J., Zhu Q., Kitareewan S., Dmitrovsky E., Lander E.S., Golub R. Intepretating patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. Proc. Natl. Acad. Sci. USA 1999; 96:2907-2912.