

# Autonomous Learning in an Information Stream through Autopoiesis

Nikolaos Nanas<sup>1</sup>

<sup>1</sup>Center for Research and Technology Thessaly, Greece  
n.nanas@cereteth.gr

## Abstract

Nootropia is a complex, self-organizing system, inspired by the Theory of Autopoiesis and successfully applied so far to the challenging problem of profiling a user's information interests. In this paper for the first time, Nootropia is studied in the context of Artificial Life, as an autonomous system that can learn without human intervention. A series of experiments demonstrate that Nootropia can autonomously learn to identify documents belonging to a specific topic with minimal training. This is achieved through a deterministic process of self-organization, which, when coupled with a complex and dynamic information environment, gives rise to rich and unpredictable behavior. Nootropia is open to its environment and operates far from equilibrium, while it tries to maintain its identity within an information stream. Our exploration of the dynamics behind Nootropia's autonomous learning capabilities lead to interesting insights, which may extend beyond its successful application to the problem of profiling and towards a new research stream that uses Nootropia as a means for studying computational autopoiesis.

## Introduction

Humberto Maturana and Francisco J. Varelas' *Autopoietic Theory* describes a model of self-organisation (Varela et al., 1974; Maturana and Varela, 1980). In simple words, it states that a system's organisation is defined by its "structure" (its components (nodes) and their relations (links)) and the processes that this structure performs, which continuously regenerate the structure that produces them. Of particular interest to the current work is Varela's view of the immune system in the context of Autopoietic Theory. Varela treated the immune system as an *organisationally closed* network that reacts autonomously in order to define and preserve the organism's identity, in what is called *self-assertion* (Varela and Coutinho, 1991). Self-assertion is an on going process, since both the organism and the environment change over time.

Two types of change contribute to self-assertion. The network's *dynamics* refer to ongoing variations in the concentration of antibodies and play the role of reinforcement learning. The network's *metadynamics* are the result of the recruitment of new cells (produced by the bone marrow

and of the removal of existing cells. The network's metadynamics play the role of a distributed control mechanism that allows the network to maintain its viability by shifting its immune repertoire (Bersini and Varela, 1994). It is also important, that due to the interactions between antibodies, it is essentially the network itself that chooses which new recruited cells will survive in the network. According to Vaz and Varela, self-assertion is the natural consequence of this *endogenous selection* process (Vaz and Varela, 1978).

Stewart and Varela used a computational model to explore self-assertion (Stewart and Varela, 1991). Like the original computer simulation of a cell-like autopoietic structure in (Varela et al., 1974), Stewart and Varela's model involved a discrete two dimensional grid representation of shape-space, where antibodies are randomly introduced. The survival of antibodies on this grid depends on their affinity to other antibodies, with affinity being a function of the distance between two antibodies. The simulation gave rise to stable (but not static) patterns that were the result of the network's metadynamics. Similar self-assertion models have also been studied in (De Boer and Perelson, 1991) and (Bersini, 2002).

Discrete, two dimensional spaces have been the basis of many computational models of autopoiesis. A comprehensive review can be found in (McMullin, 2004). Although cellular automata on two-dimensional grids are known to be capable of universal computation, in the case of autopoiesis and self-assertion models in particular, the simulated environments are relatively simple. For instance, in the original computational model of autopoiesis, the environment where the cell-like structure is formed comprises particles that bond in the presence of a catalyst to form the cell's membrane. Similarly, in Stewart and Varela's model of the immune system the external environment consists of randomly generated antibodies in the shape space. In both cases, the computer simulations demonstrate visually, that despite the stochastic nature of the environment stable structures progressively emerge and manage to maintain their identity over time.

This paper suggests an alternative scientific methodology

for exploring autonomous behaviour through autopoiesis. It uses the Web as a source of real-world data for simulating a complex and dynamic information environment. In this environment, a profiling system, which has been inspired by the autopoietic view of the immune system, has to autonomously learn to identify specific information, in order to maintain its identity. A series of experiments demonstrates that this system is capable of autonomous learning through a process of self-organisation that dynamically controls the profile's structure. Although, the adopted information environment cannot be easily visualised, depicting how certain macroscopic variables vary over time, reveals a complex system that, although deterministic, is unpredictable. Small variations in the initial conditions can cause significant variations in the structural pathways the system follows as it interacts with its complex environment. The results also reveal an interesting relation between energy consumption and autonomous behaviour that requires further investigation.

### Profiling with Nootropia

According to (Mireille, 2008), profiling could be generally defined as:

“The process of ‘discovering’ correlations between data in databases that can be used to identify and represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category.”

In practice, when profiling an individual's (or a group's) information interests, a profile is built and continuously learns from the user's interaction with information and is used to evaluate the relevance of new, incoming information to these interests. Profiling in this case, is a challenging problem with analogies to the immune system's self-assertion process. To maintain its viability a profile has to be able to define and preserve the identify of the user's interests. It has to be able to learn a variety of interests and continuously adapt to changes in them.

These analogies inspired the design and development of Nootropia<sup>1</sup>, a profiling system that so far, has been successfully applied for adaptive filtering of textual information according to a user's (or a group's) interests. In its current form, Nootropia was first introduced in (Nanas et al., 2004) and since then, it has been extensively described and experimentally evaluated (see for instance (Nanas and De Roeck, 2009; Nanas et al., 2009, 2010b,a).

In Nootropia, the profile is a weighted network of features, e.g., a network of words extracted from the content of text documents. The links in this network capture correlations between features that appear regularly in the same

context, e.g., correlations between words that appear close to each other in text. A feature's weight measures its importance within the profile and a link's weight the strength of the correlation between two features. The profile is built and continuously adapts to interest changes through a process of self-organisation that adjusts the network's structure in response to user feedback (explicit or implicit). For instance, if a document is identified as relevant to the user's interests, then words in the profile that also appear in the document get reinforced at the expense of the words they are linked to. These local competitions cause a redistribution of weight between the profile's words (dynamics). Words in the document that do not already appear in the profile are recruited and those profile words that run out of weight are purged (metadynamics). The exact self-organisation process is described in detail in (Nanas and De Roeck, 2009).

To evaluate the relevance of an information item (e.g., document), the profile deploys a directional spreading activation process. Profile features (e.g., words) that also appear in the item get activated. In order of increasing weight, each activated feature disseminates part of its current activation towards the activated features with larger weights that it is linked to. The relevance score is then calculated as the weighted sum of the final activation of profile features. This non-linear evaluation process, which is described in detail in (Nanas et al., 2010a), implies a hierarchy of features, as activation is being channeled from the majority of features with small weights towards the “elite” of features with large weights. The structure of this implicit hierarchy, which continuously self-organises in response to the environment, defines the profile's collective reaction to incoming information.

The autopoietic properties of Nootropia are discussed in detail in (Nanas and De Roeck, 2009), where it is argued that Nootropia exhibits the basic characteristics of self-assertion models. It is a non-linear, self-organising system, that is open to its environment and operates far from equilibrium, constantly adjusting structurally, and hence behaviourally. It also involves both network dynamics and metadynamics with endogenous selection. Experiments performed in (Nanas and De Roeck, 2009) and (Nanas et al., 2010b) demonstrate Nootropia's ability to effectively adapt to a variety of interest changes through self-organisation. Further experiments and analysis indicate that it is the network's non-linearity which allows the profile to store additional information regarding a user's interests and thus remain specific even within high-dimensional spaces (Nanas et al., 2010b,a). In such spaces, comparative experiments between Nootropia and a vector-based profile containing the same weighted words, show that the additional information encoded by Nootropia's links contributes to an increase in accuracy of up to 50% (Nanas et al., 2010a). Nootropia's advantageous properties have already boosted the development of real world prototypes, such as the Personalised

<sup>1</sup>Greek word for: “an individual's or a group's particular way of thinking, someone's characteristics of intellect and perception”.

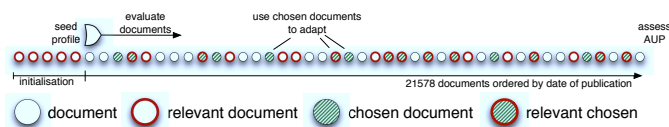


Figure 1: Experimental Process

News Aggregator described in (Nanas et al., 2010c).

In all of the above cases though, it is the user (or a group of users) that explicitly, or implicitly, provides the profile with relevant information to learn from. So what if we take the user out of the equation and ask the profile to autonomously identify and choose information to learn from? Will the profile be able to maintain its identity and what lessons can be learned from its autonomous learning behaviour? This paper deals with these questions experimentally in the context of Alife in general and of the Theory of Autopoiesis in particular.

## Experimental Evaluation

The performed experiments are a continuation of those referred above and use a variation of their methodology to test the ability of a profile to autonomously learn to identify documents belonging to a specific topic category. Once more the dataset used in the experiments is the Reuters-21578 document collection<sup>2</sup>. It includes 21578 news stories from Reuters newswire in 1987, ordered according to publication date and classified by human experts into 135 topic categories. The experiments focus on the 23 topics with at least 100 relevant documents in the dataset.

### Autonomous Profile

As it is exemplified in figure 1, for each of the 23 topics, a “seed” profile is initialised using the first five<sup>3</sup> documents in the collection belonging to the topic. The seed profile is then released in the information stream and traverses the 21578 documents in the collection in chronological order. The profile evaluates every individual document and assigns to it a score. If the assigned score is over a threshold the profile chooses the document for “consumption” and self-organises accordingly. For the current experiments, the threshold is calculated for each individual document as the average score assigned to the documents “consumed” so far. The process is repeated until all 21578 have been accounted for. The

<sup>2</sup>available at <http://www.daviddlewis.com/resources/testcollections/reuters21578/>

<sup>3</sup>Experiments were also performed for 1, 10 and 50 initialisation documents, but are not reported here due to space limitations. With just 1 initialisation document the seed profile is not developed enough to achieve the desired behaviour. As the number of initialisation documents increases from 5 to 10 and then to 50 the profile relies more on its initial condition rather than the subsequent learning process.

profile’s accuracy is then measured by calculating the Average Uninterpolated Precision (AUP) of the list comprising the documents in the collection ordered by decreasing score. A topics AUP is defined as the sum of the precision<sup>4</sup> at each point in the ordered list where a relevant document appears, divided by the total number of relevant documents. The essence of this accuracy metric is that documents relevant to the current topic should receive larger scores than irrelevant documents.

The above methodology establishes a challenging experimental task. Based only on its initial training with a small number of documents relevant to a topic, the profile has to autonomously learn to identify documents belonging to that topic. Ideally, the profile should choose all relevant documents and ignore the rest<sup>5</sup>. However, as it is depicted in figure 1, there are typically both false negatives and false positives. Not all relevant documents are chosen and not all chosen documents are relevant. Both the percentage of relevant documents chosen and the percentage of chosen documents that are relevant affect the profile’s accuracy. If the first percentage is small the profile ignores valuable input. If the second percentage is small then the profile may deviate away from the current topic of interest. It should also be noted that since the content of documents relevant to a topic may change over time, the profile has to be able to follow this drift. Overall, the choices the profile made so far define its current structure and consequently its future choices. So even small changes in the initial conditions can cause the profile to follow a very different trajectory. Out of an infinite number of possible network configuration the profile has to self-organise in such a way that it manages to maintain its (topical) identity within a complex and dynamic environment.

### Supervised and Random Profiles

In the experiments the accuracy and behaviour of the autonomous profile are juxtaposed with those of a *supervised profile* and of a *random profile*<sup>6</sup>. In both cases we start with an initially empty profile. Like before the profile is released in the information stream and evaluates the 21578 documents in chronological order. Unlike the autonomous profile, these two types of profile do not choose the documents to learn from autonomously. Whenever the supervised profile evaluates a relevant document it will always use it for learning, while it ignores all non-relevant documents. In other words, it is provided a priori with complete knowledge of which documents are relevant to the current topic of

<sup>4</sup>i.e., the ratio of documents relevant to that topic.

<sup>5</sup>It is assumed that the categorisation of documents by Reuter’s experts has been accurate.

<sup>6</sup>All three types of profile are built using Information Gain to extract the most important words in the training documents and a sliding window of size 20 for identifying correlations between the extracted words.

| topic code   | relevant docs | docs chosen | relevant chosen | rel. chosen/docs chosen | rel. chosen/total rel. | AUP autonomous | AUP supervised | AUP random | auto/supervised |
|--------------|---------------|-------------|-----------------|-------------------------|------------------------|----------------|----------------|------------|-----------------|
| earn         | 3987          | 715         | 694             | 0.97                    | 0.17                   | 0.694          | 0.732          | 0.349      | 0.949           |
| acq          | 2448          | 433         | 149             | 0.34                    | 0.06                   | 0.262          | 0.424          | 0.105      | 0.617           |
| money-fx     | 801           | 16          | 11              | 0.69                    | 0.01                   | 0.311          | 0.556          | 0.046      | 0.559           |
| crude        | 634           | 123         | 114             | 0.93                    | 0.18                   | 0.636          | 0.700          | 0.033      | 0.909           |
| grain        | 628           | 6           | 1               | 0.17                    | 0.00                   | 0.246          | 0.509          | 0.027      | 0.483           |
| trade        | 552           | 94          | 65              | 0.69                    | 0.12                   | 0.334          | 0.558          | 0.044      | 0.599           |
| interest     | 513           | 124         | 108             | 0.87                    | 0.21                   | 0.413          | 0.463          | 0.030      | 0.892           |
| wheat        | 306           | 77          | 53              | 0.69                    | 0.17                   | 0.430          | 0.490          | 0.012      | 0.878           |
| ship         | 305           | 54          | 6               | 0.11                    | 0.02                   | 0.029          | 0.436          | 0.011      | 0.066           |
| corn         | 254           | 8           | 3               | 0.38                    | 0.01                   | 0.322          | 0.275          | 0.010      | 1.171           |
| dlr          | 217           | 97          | 52              | 0.54                    | 0.24                   | 0.371          | 0.468          | 0.014      | 0.793           |
| oilseed      | 192           | 7           | 2               | 0.29                    | 0.01                   | 0.298          | 0.174          | 0.010      | 1.714           |
| money-supply | 190           | 719         | 64              | 0.09                    | 0.34                   | 0.051          | 0.184          | 0.012      | 0.279           |
| sugar        | 184           | 199         | 30              | 0.15                    | 0.16                   | 0.116          | 0.683          | 0.008      | 0.169           |
| gnp          | 163           | 8           | 3               | 0.38                    | 0.02                   | 0.384          | 0.424          | 0.013      | 0.907           |
| coffee       | 145           | 51          | 45              | 0.88                    | 0.31                   | 0.775          | 0.824          | 0.007      | 0.940           |
| veg-oil      | 137           | 182         | 21              | 0.12                    | 0.15                   | 0.082          | 0.459          | 0.012      | 0.180           |
| gold         | 135           | 10          | 5               | 0.50                    | 0.04                   | 0.763          | 0.768          | 0.005      | 0.993           |
| nat-gas      | 130           | 15          | 9               | 0.60                    | 0.07                   | 0.665          | 0.432          | 0.007      | 1.538           |
| soybean      | 120           | 7           | 2               | 0.29                    | 0.02                   | 0.421          | 0.285          | 0.005      | 1.479           |
| bop          | 116           | 230         | 45              | 0.20                    | 0.39                   | 0.175          | 0.310          | 0.005      | 0.566           |
| livestock    | 114           | 8           | 3               | 0.38                    | 0.03                   | 0.111          | 0.265          | 0.006      | 0.419           |
| cpi          | 112           | 119         | 26              | 0.22                    | 0.23                   | 0.065          | 0.285          | 0.005      | 0.229           |
| average      | 538.4         | 143.6       | 65.7            | 0.5                     | 0.1                    | 0.346          | 0.465          | 0.034      | 0.743           |

Table 1: Experimental Results. Columns from left to right: (1) topic code, (2) number of relevant documents in the collection, (3) number of documents chosen by the the autonomous profile, (4) number of chosen documents relevant to the current topic, (5) ratio of chosen documents that are relevant, (6) ratio of relevant documents chosen, (7) per topic AUP score for the autonomous profile, (8) per topic AUP score for the supervised profile, (9) per topic AUP for the random profile, (10) ratio of the supervised profile’s AUP achieved by the autonomous profile.

interest. The random profile, on the other hand, is provided with an equal number of randomly selected documents from the collection.

### Accuracy

Table 1 summarises for each topic, the choices made by the autonomous profile and the resulting AUP score and compares it to those of the supervised and autonomous profile. The results lead to the following observations:

- The accuracy of the random profile is the lowest (table 1 col. 9). The profile must learn from relevant documents to be accurate.
- As expected the supervised profile achieves the best overall performance (table 1 col. 8).
- The performance of the autonomous profile is satisfactory (table 1 col. 7). It achieves on average 74% of the supervised profile’s accuracy (table 1 col. 10).
- The autonomous profile achieves this level of accuracy although on average it only identifies 10% of the existing relevant documents per topic (table 1 col. 6). It appears that not all of the available relevant documents are required for increased accuracy. In fact, it is interesting that there are four topics (corn, oilseed, nat-gas, soybean) for which the autonomous profile clearly outperforms the supervised profile although, after its initialisation with five documents, it chooses a very small number of documents to learn from. It may be the case, that for certain topics with relatively small number of relevant documents in

the collection and distinct content, this is a better strategy. The seed profile overspecialises to the initialisation documents, but these are representative enough of the remaining relevant documents that ignoring them leads to better accuracy. In any case, this is not always the best strategy (e.g., topics grain, ship, and livestock).

- The satisfactory accuracy of the autonomous profile is mainly due to the fact that, on average, 50% of the documents chosen are indeed relevant. There is a clearer correlation between the profile’s accuracy and the percentage of chosen documents that are relevant. In general, if the percentage is small the accuracy of the autonomous profile is small and increases as the percentage increases. For percentages close to one the accuracy of the autonomous profile approximates that of the supervised profile. It is clear that the profile has to be selective when choosing the documents to learn from. Too many false positives can cause the profile to drift away from the current topic of interest.

### Behaviour

Nootropia is a complex system and it is not easy to visualise, or to analyse, its dynamic behaviour. In this paper, an attempt is made to understand how self-organisation contributes to the above autonomous learning capabilities, by observing certain macroscopic variables related to the profile’s nodes and their weights. The analysis of the network’s connectivity is part of ongoing work and will be included in future publications. Furthermore, due to space limitations,

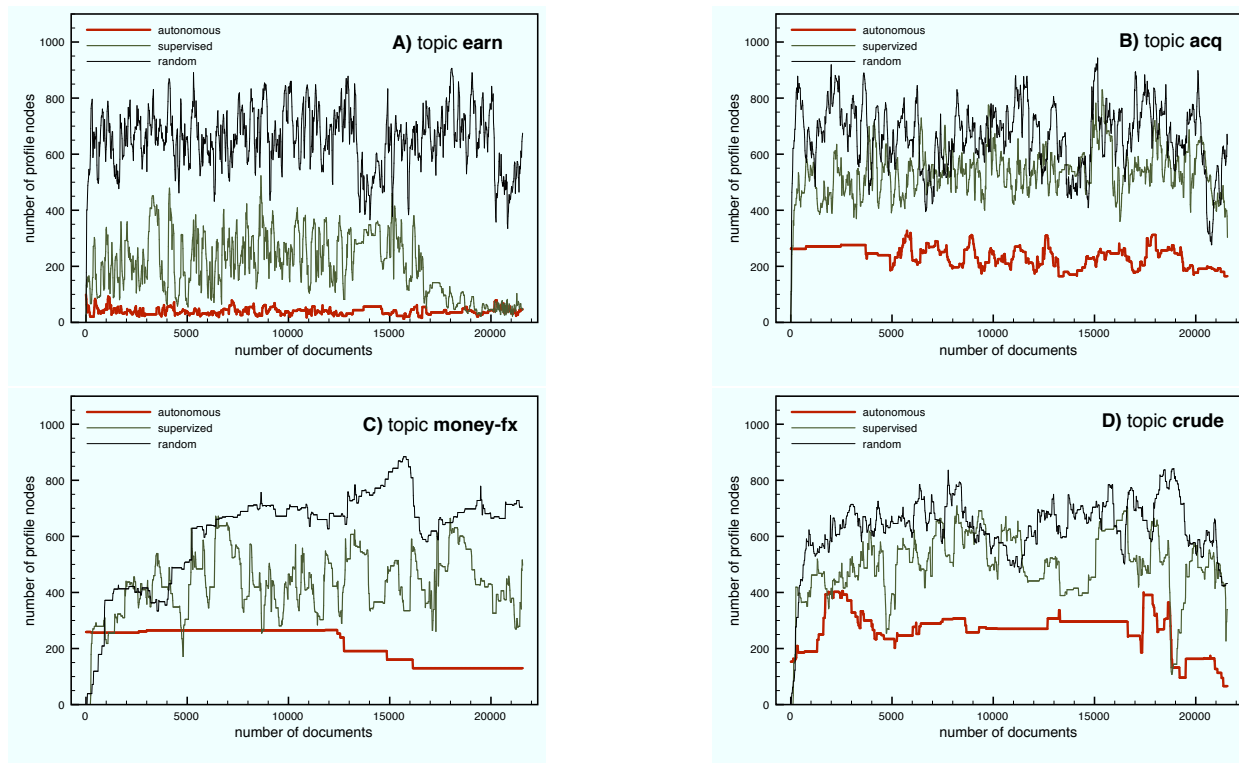


Figure 2: Fluctuations in the number of profile nodes along the information stream.

only the four topics with the largest number of relevant documents are indicatively chosen for this study.

The graphs in figure 2 depict for each of the topics earn, acq, money-fx and crude, how the number of nodes (words) in the profile (Y-axis) changes as it traverses the document collection (X-axis). These indicative graphs show that Nootropia is a dissipative, self-organising system that can dynamically control its size (and connectivity (Nanas et al., 2009)). Energy (word weight) flows through the profile with the addition of words and is dissipated when these words are purged. Although there are more than 20,000 unique words in Reuters-21578<sup>7</sup>, the number of nodes in the profile does not escalate above 1000. In all four cases, the three types of profile are easily distinguished based on the average number of words. The autonomous profile maintains the smallest number of words and the random profile the largest number of words, although it uses the same number of documents to learn from as the supervised profile. So these differences are not only due to differences in the number of training documents, but they also depend on the semantic diversity of these documents. The random profile is provided with randomly selected training documents from the collection, that may belong to any topic. These documents may include a greater variety of words and thus give rise to a profile

<sup>7</sup>After stop word removal and stemming.

with a larger number of nodes. The supervised profile uses the same number of documents relevant to a specific topic and so their vocabulary is more focused. For the same reasons, the autonomous profile appears to be the most focused profile type, with the least number of profile words and the mildest fluctuations. Apparently, the profile has the ability to choose documents that are semantically close to its initial composition and their vocabulary is already reflected in the profile. These documents do not have many new words to contribute to the profile and cause as a result smaller profile perturbations. It is also evident from these figures that the average number of words in each profile type varies from topic to topic and depends not only on the number of relevant documents, but also on the semantic characteristics of each topic. Finally, it is clear that in the case of topic money-fx (fig. 2 C ), the profile does not successfully identify appropriate documents to learn from, causing a decrease in the number of profile words and the poorest relative accuracy out of the four cases (see tbl. 1).

To further investigate the behaviour of Nootropia, figures 3 and 4 depict respectively, the average and aggregate weight of profile nodes through out the 21578 documents<sup>8</sup>. With the exception of the unsuccessful topic money-fx, the autonomous profile has the largest average weight, which

<sup>8</sup>Note that for visualisation reasons the Y-axis of the graphs in figure 3 has various scales.

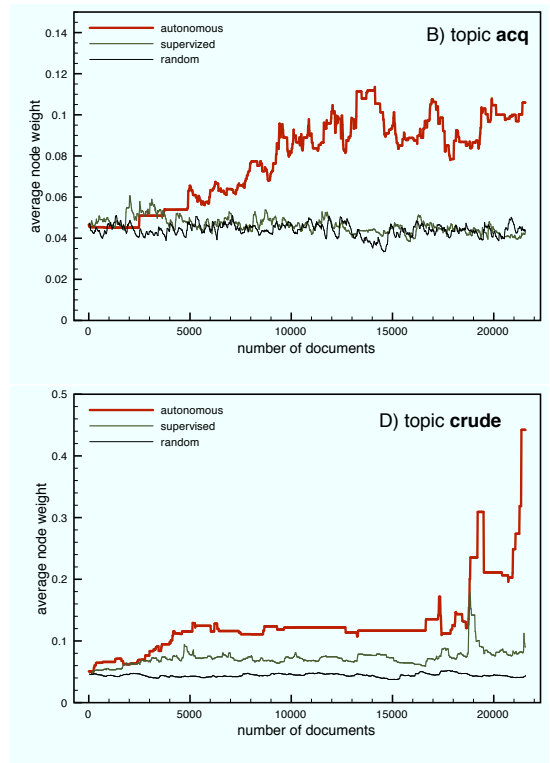
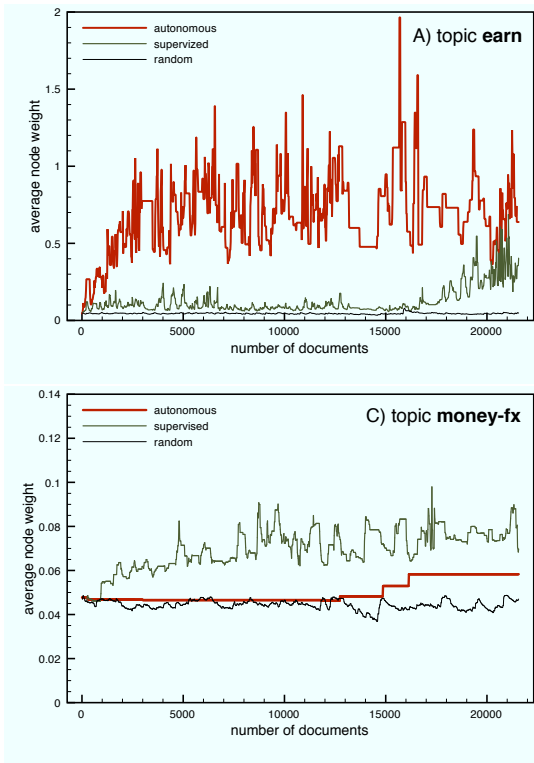


Figure 3: The average weight of profile nodes along the information stream.

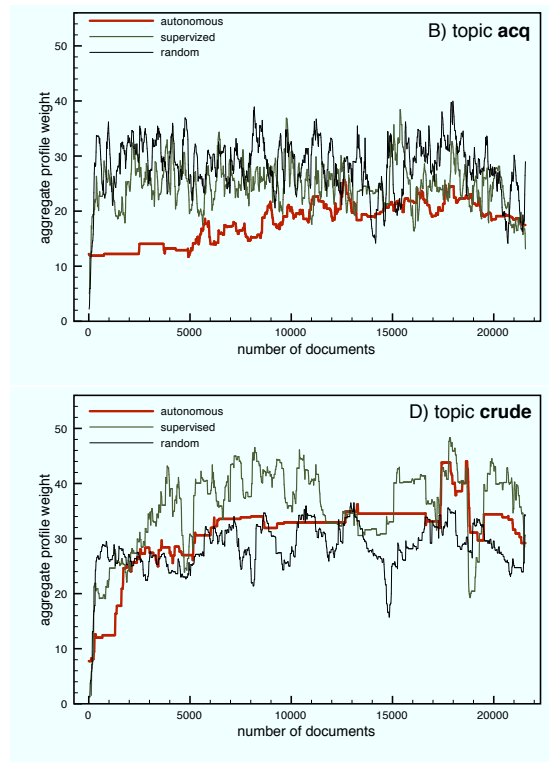
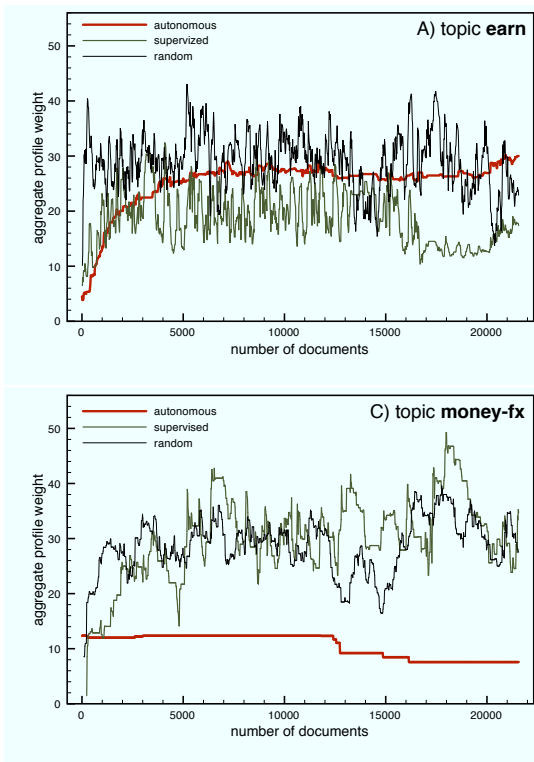


Figure 4: The aggregate profile weight along the information stream.

tends to increase along the process. There is also an apparent correlation between the scale of the average node weight and the accuracy of the profile (see table 1). Furthermore, the average weight does not depend only on the number of profile nodes. According to graphs B and D in figures 2 and 3, although the autonomous profile maintains in both cases approximately the same number of nodes, there is a significant difference in the average weight. This means that the average weight of profile nodes is possibly another macroscopic variable that characterises the behaviour and accuracy of the autonomous profile. It shows that the autonomous profile can effectively maintain and reinforce its identity. By choosing documents that are relevant to its initial semantic composition, the profile reinforces what has already been learned and remains specific to its area of interest, thus avoiding intense structural fluctuations.

The distinct behaviour of the autonomous profile is also reflected in the way it aggregates node weight. According to figure 4, with the exception again of topic money-fx, the autonomous profile progressively accumulates weight until it reaches a certain capacity, where it tends to stabilise. It is also interesting, that unlike the supervised and the random profile, the aggregation of weight by the autonomous profile is more progressive and with less fluctuations. Finally, it is notable that for topic earn, acq and crude the aggregate weight of the autonomous profile is comparable to that of the supervised and random profiles, despite the smaller number of documents used for learning (tbl. 1) and the smaller number of profile words (fig. 2).

## Discussion

The experimental results show that Nootropia is capable of autonomous learning within a complex information environment. The system's accuracy in itself is not the primary concern of this paper. It is already satisfactory enough, given the small amount of training data that are provided for initialisation and it can be further improved, e.g., through more elaborate thresholding mechanisms. What is important is that this unsupervised learning behaviour is the result of an autopoietic network's self-organisation in response to a diverse and changing information environment. The network "perceives" its environment through a non-linear spreading activation process that leads to increased specificity even within high-dimensional environments (Nanas et al., 2010a). As a result, the network can accurately identify and extract relevant information from the environment. The "cognitive", learning process is the result of the network's reaction to the extracted information and involves both the redistribution of node weights through local interactions (network dynamics) and the addition and removal of nodes (network metadynamics). The network becomes open to the environment: energy (weight) is absorbed from the environment, it is temporarily stored by the network and eventually, it is disseminated back to the environment. The distribution of

stored energy (weight) in the network imposes a hierarchy on the nodes that defines the network's response to the environment. When the network is forced to self-organise in response to random information then it becomes large (fig. 2), but the hierarchy of nodes remains shallow (fig. 3). The network is more volatile, because more nodes have small weights and can be more easily removed from the network, causing pronounced fluctuations in the number of nodes. On the contrary, relevant information reinforces what is already in the network with additional energy and the hierarchy of nodes grows higher. This increases the stability and specificity of the network and it becomes more likely that it will identify additional relevant information, leading to a positive feedback loop, which allows the profile to maintain its identity and to avoid strong perturbations.

Some interesting lessons can be learned from all the above:

- The World Wide Web can serve as a valuable source of real-world data, for simulating complex and dynamic environments to experiment in the domain of Artificial Life. These environments are multidimensional and cannot be visualised. They provide however a rich information world that lies somewhere in the middle of the range between the relatively simple 2D worlds of many computer simulations and the physical world. As in the case of Varela's 2D simulations (Varela et al., 1974; Stewart and Varela, 1991), the above experiments demonstrate that even in such a complex environment autopoiesis can still give rise to consistent, "meaningful" behaviour that can maintain a system's viability.
- If Nootropia is indeed an autopoietic system, or at least exhibits some autopoietic properties, then its experimental study highlights the importance of the environment during autopoiesis. Nootropia is organisationally closed, but it is the interaction with the environment that guides its structural and hence, behavioural development. It is structurally coupled to its environment and unlike existing 2D simulations, it is the richness of this environment that can give rise to a plethora of structural modifications and corresponding behaviours.
- It is even more interesting that the study of Nootropia's behaviour indicates a relation between energy and autopoiesis. The autonomous profile effectively accumulates energy per node, to reinforce its structure and hence its identity. As a consequence, even within a complex environment and despite the infinite number of possible structural pathways, the profile can be specific enough to choose the information that will lead to further energy aggregation and self-assertion. The role of energy during autopoiesis has been ignored by existing computational investigations and will become a major theme of the research endeavour that this paper initiates.



## Summary and Outlook

Computational Autopoiesis is already an established area of research in Alife. The most common approach involves simulating autopoietic (cell-like) structures on a discrete, two-dimensional space. The current work deviates from this practice. Nootropia is a profiling model that has been inspired by Varela's view of the immune system as an organisationally closed network of interacting antibodies, which reacts autonomously to define and preserve the host organism's identity. Nootropia has already been evaluated extensively and has produced significant results, both quantitatively and qualitatively. The past experimental work on Nootropia concentrated on supervised learning. In this paper for the first time, human intervention is kept minimal. A collection of news articles ordered according to publication date serves as an information stream and within this stream a profile, which has been initialised with a small number of articles relevant to a topic, has to autonomously identify and learn from more relevant articles. This autonomous profile is contrasted to a supervised profile that has complete knowledge of what is relevant and a random profile that chooses documents to learn from at random.

The accuracy of the autonomous profile is satisfactory. It clearly outperforms the random profile and achieves a level of accuracy that is on average 74% that of the supervised profile. What is important is that this level of accuracy is the result of self-organisation in response to the environment. The analysis of Nootropia's behaviour provides evidence that it can control its structure dynamically and in such a way that it effectively consumes and stores energy from the environment. The stored energy reinforces the network's structure and hence the profile's specificity. It becomes easier for the profile to identify more relevant information, leading naturally to self-assertion.

The experimental work in this paper demonstrates also that exploiting the web as a valuable source of real-world data for simulating complex and dynamic environments, can be a fruitful avenue of research in Alife. It is in such multidimensional information worlds that interesting complex structures and behaviours may arise as a natural consequence of autopoiesis. This paper is only a first step in this research avenue. Future steps involve a more extensive experimentation and analysis, including statistical analysis of the network's properties (e.g., degree distribution and clustering coefficient), but also, a more comprehensive exploration of the background theories and philosophies. This paper brings to our attention the relation between information, energy and life and creates one more connection between Alife and the domains of Thermodynamics and Energetics in general. Nootropia is a possible means for exploring this relation through computational autopoiesis and some interesting insights have been gained with the current work. Much more is of course required to be able to make bolder claims, or to draw more general conclusions.

## References

- Bersini, H. (2002). Self-assertion versus self-recognition: A tribute to Francisco Varela. In *ICARIS 2002: 1st International Conference on Artificial Immune Systems*, pages 107–112. University of Kent at Canterbury Printing Unit.
- Bersini, H. and Varela, F. (1994). The immune learning mechanisms: Reinforcement, recruitment and their applications. In *Computing with Biological Metaphors*, pages 166–192. Chapman Hall.
- De Boer, R. J. and Perelson, A. S. (1991). Size and connectivity as emergent properties of a developing immune network. *Journal of Theoretical Biology*, 149:381–424.
- Maturana, H. R. and Varela, F. J. (1980). *Autopoiesis and Cognition: the Realization of the Living*, volume 42 of *Boston Studies in the Philosophy of Science*. D. Reidel Publishing Company, Dordrecht, Holland.
- McMullin, B. (2004). 30 years of computational autopoiesis: A review. *Artificial Life*, 10(3):277–295.
- Mireille, H. (2008). *Defining Profiling: A New Type of Knowledge*, chapter 2, pages 303 – 326. Cross-Disciplinary Perspectives. Dordrecht Springer.
- Nanas, N. and De Roeck, A. (2009). Autopoiesis, the immune system and adaptive information filtering. *Natural Computing*, 8(2):387–427.
- Nanas, N., Uren, V., and De Roeck, A. (2004). Nootropia: a user profiling model based on a self-organising term network. In *Artificial Immune Systems, Third International Conference (ICARIS 2004)*, pages 146–160. Springer, Heidelberg, Germany.
- Nanas, N., Vavalis, M., and De Roeck, A. (2010a). A network-based model for high-dimensional information filtering. In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval, SIGIR '10*, pages 202–209, New York, NY, USA. ACM.
- Nanas, N., Vavalis, M., and De Roeck, A. (2010b). Words, antibodies and their interactions. *Swarm Intelligence*, 4(4):275–300.
- Nanas, N., Vavalis, M., and Houstis, E. (2010c). Personalised news and scientific literature aggregation. *Information Processing and Management*, 46:268–283.
- Nanas, N., Vavalis, M., and Kellis, L. (2009). Immune learning in a dynamic information environment. In *Artificial Immune Systems, 8th International Conference (ICARIS 2009)*, pages 192–205. Springer, Heidelberg, Germany.
- Stewart, J. and Varela, F. J. (1991). Morphogenesis in shape-space, elementary meta-dynamics in a model of the immune network. *Journal of Theoretical Biology*, 153:477–498.
- Varela, F. J. and Coutinho, A. (1991). Second generation immune network. *Immunology Today*, 12(5):159–166.
- Varela, F. J., Maturana, H., and Uribe, R. (1974). Autopoiesis: The organization of living systems, its characterization and a model. *Biosystems*, 5(4):187–196.
- Vaz, N. M. and Varela, F. (1978). Self and non-sense: An organism-centered approach to immunology. *Medical Hypotheses*, 4:231–267.