KNOWLEDGE MEDIA KMi INSTITUTE

Nootropia:

A Self-Organising Agent for Adaptive Document Filtering

Tech Report kmi-04-2

Nikolaos Nanas, Victoria Uren, John Domingue and Anne de Roeck

The Open University

Nootropia: a Self-Organising Agent for Adaptive Document Filtering

Nikolaos Nanas (Phd Student) , Victoria Uren, John Domingue
Knowledge Media Institute, The Open University

Anne de Roeck
Department of Computing and Mathematics, The Open University

Abstract

This paper presents Nootropia, a self-organising information agent, capable of evaluating documents according to a user's multiple and changing interests. In Nootropia, a hierarchical term network that takes into account term dependencies is used to represent a user's multiple topics of interest. Non-linear document evaluation is established on that network based on a directed spreading activation model. We then introduce a process for adjusting the network in response to changes in user feedback. We argue that Nootropia exhibits self-organising characteristics, which, as demonstrated experimentally, allow Nootropia to adapt to a variety of simulated interest changes.

1. Introduction

In recent years, advances in digital media, network and computing technologies have caused an exponential growth of the digital information space that is accessible to individuals. We are now faced with the cumbersome task of selecting out of this glut of accessible information, information items that satisfy our interests, i.e. "relevant information". This is the problem that is usually referred to as "Information Overload" [6].

The purpose of Personal Information Agents or Assistants is to alleviate a user's information overload. This typically involves a tailored representation of a user's interests, a *user profile*. The user profile is used to evaluate information items (e.g. documents) according to the user's interests and should be able to adapt to changes in them. Such changes are caused by changes in the user's environment and knowledge. They are dynamic and can range from modest, but potentially fast, short-term changes, to occasionally radical, but more progressive, long-term changes.

In this paper we present *Nootropia*[1], a self-organising information agent that can evaluate documents according to a user's changing interests. In Nootropia, the user profile is a hierarchical term network, on which we establish non-linear document evaluation (section 3). Adaptation is then achieved through a self-organising process that allows the profile to respond structurally to variations in feedback (section 4). Initial experiments indicate Nootropia's ability to adapt to a variety of simulated interest changes (section 5).

1 Greek word for: "an individual's or a group's particular way of thinking, someone's characteristics of intellect and perception"

2. Related Work

Personal information agents have been applied for the filtering of netnews [22, 3, 7, 21, 20] and e-mails [24, 10], and for autonomously searching the WWW [8, 12]. In all of the above cases an agent has to be able to evaluate textual information according to a users changing interests. For that purpose, models and techniques from the domains of Information Retrieval (IR) and Adaptive Information Filtering (AIF) are adopted.

The Vector Space Model (VSM) [17] is commonly used for profile representation. The VSM ignores term dependencies. It is for that reason linear and can only represent a single topic of interest [14]. More than one single-topic profile are required to represent a user's multiple interests. Based on user feedback, each profile is then adapted separately, typically using linear learning algorithms like Rocchio's [15]. These assume a steady change of interests, reflected by a constant learning coefficient [5, 18]. The learning coefficient is appropriately adjusted over time when reinforcement learning is used [19]. To account for the difference in the pace of short- and long-term changes, dual profiles with separate learning coefficients have also been suggested [22, 2].

Inspired by biology, Genetic Algorithms (GAs) constitute a different solution to profile adaptation. Typically a population of profiles (or agents) is maintained that collectively represent the user interests [20, 24, 8, 12, 1, 9]. The population evolves according to user feedback. Individual profiles that better represent the user interests dominate over less successful profiles in the population. Furthermore, in many cases, evolution of the population as a whole is combined with learning at the individual profile level. Linear learning algorithms [20], Hebbian [1] and reinforcement learning [9, 8] have been employed for that purpose. With the hybridisation of GAs with learning, these systems tackle the trade-off between fast, short-term changes and occasionally radical, but progressive, long-term changes.

Finally, in contrast to the above systems, which employ profile representations that ignore term dependencies, there are a couple of connectionist information agents that take into account the lexical correlations between terms in the same phrase. They use single-topic profiles, which are adapted separately through linear learning of term and link weights [21], or a combination of unconstrained and constrained Hebbian learning [7].

In summary, to evaluate documents according to a user's multiple and changing interests, existing information agents are forced to use more than one profile. Based on user feedback each profile is then adapted separately using some learning algorithm, or the profiles are adapted collectively using GAs. Since multiple profiles have to be maintained and adapted, these approaches, and especially GAs, are computationally expensive [23]. Furthermore, they assume that topics are unrelated and ignore any topic-subtopic relation between them. Finally, they imply a large number of parameters, like learning coefficients and relative importance weights, that require optimization which may have to be performed separately for each individual user. To our knowledge, no information agent exists that tackles the above problem with a single profile.

3. Profile Representation and Document Evaluation

In Nootropia, a user's multiple interests are represented with a hierarchical term network. Given a set of documents about various topics that the user has specified as interesting, the network is synthesised in three steps. Briefly (for more details see [14]):

1. Informative terms are extracted from the interesting documents using a term weighting method that we call *Relative Document Frequency* (RelDF) [13]. Extracted terms populate the profile.
2. Correlations between profile terms in the interesting documents are identified within a sliding window of 10 contiguous terms. Two profile terms are linked if they appear at least once within the window, i.e. within the same context. A weight $w_{ij} \in (0, 1]$ is then assigned to the link between two extracted terms t_i and t_j using equation 1. fr_{ij} is the number of times t_i and t_j appear within the window, fr_i and fr_j are the number of occurrences of t_i and t_j in the interesting documents and d is the average distance between the two linked terms. Two extracted terms that appear next to each other have a distance of 1, while if l words intervene between them the distance is $l + 1$. The result of the first two steps is a flat network of terms (nodes) and correlations (links) between them, similar to the networks of language studied in [11].
3. Finally, we go a step further to order profile terms according to decreasing weight. We impose on that sense order on the relations between terms.

$$w_{ij} = \frac{fr_{ij}^2}{fr_i \cdot fr_j} \cdot \frac{1}{d} \tag{1}$$

The above three steps synthesise out of a set of interesting documents a cyclic term network that formulates a separate hierarchy for each general topic discussed in the documents. Figure 1(left) depicts a hypothetical network constructed from a set of documents about two overlapping topics. The two topics are reflected by two hierarchical sub-networks that share a small number of common terms. Each hierarchy can be identified by a term that is only connected to terms with lower weights (fig. 1(left): terms T1 and T2). This kind of "dominant" term can be used to identify the profile's "breadth", i.e. the number of general topics represented. A hierarchy's "depth" on the other hand corresponds to the number of terms with decreasing weight that are connected explicitly or implicitly to the dominant term. A topic of interest discussed in the majority of the user specified documents will be reflected by a hierarchy with larger depth. A hierarchy's depth is therefore a measure of a topic's importance within the profile.

Document evaluation is formulated as a directed spreading activation model. Given a document D, an initial energy of 1, is deposited with those profile terms that appear in D. In figure 1(right) activated terms are depicted by shadowed nodes. Subsequently, energy is disseminated sequentially, starting from the activated term with the smallest weight and moving up the weight order. If and only if, an activated term t_i is directly linked to another activated term t_j higher in the hierarchy, then an amount of energy E_{ij} is disseminated by t_i to t_j through the corresponding link. E_{ij} is defined by equation 2, where E_i^c is t_i's current energy, w_{ij} is the weight of the link between t_i and t_j, and A^h is the set of activated terms higher in the hierarchy that t_i is linked to. The purpose of the normalization parameter $\sum_{k \in A^h} w_{ik}$ is to ensure that a term does not disseminate more than its current energy. The current energy of term t_i is $E_i^c = 1 + \sum_{m \in A^l} E_{mi}$, where A^l is the set of activated terms lower in the hierarchy that t_i is linked to. After the end of the dissemination process the final energy of a term t_i is $E_i^f = E_i^c - \sum_{k \in A^h} E_{ik}$.

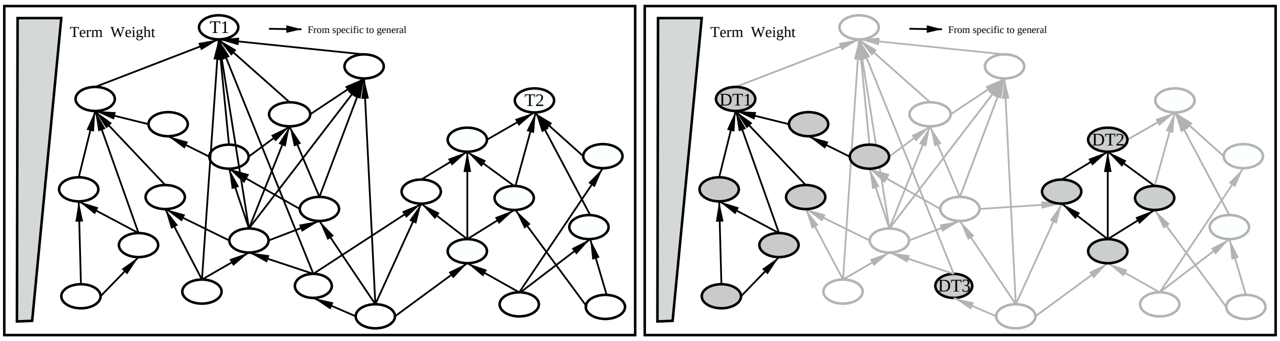

Figure 1. Hierarchical Term Network: (left) deactivated, (right) activated

$$E_{ij} = \begin{cases} E_i^c \cdot w_{ij} & \text{if } \sum_{k \in A^h} w_{ik} \leq 1 \\ E_i^c \cdot (\frac{w_{ij}}{\sum_{k \in A^h} w_{ik}}) & \text{if } \sum_{k \in A^h} w_{ik} > 1 \end{cases} \qquad (2)$$

Similarly to the complete non-activated profile, activated profile terms define subhierarchies for each topic of interest discussed in the document. The dominant terms $DT1$, $DT2$ and $DT3$ can be defined as those activated terms that didn't disseminate any energy (fig. 1(right)). The number of dominant terms measures the document's breadth b, i.e. the number of topics discussed in the document. For each dominant term the depth of the corresponding subhierarchy is equal to the number of activated terms from which energy was received. The document's depth d can thereafter be approximated as the number of activated terms that disseminated energy. Obviously, $b + d = a$, where a is the total number of activated terms.

Based on the above spreading activation model we experimented with a variety of document evaluation functions. Here we concentrate on the most successful of them. More specifically a document's relevance score S_D is calculated using equation 3, where A is the set of activated profile terms, NT the number of terms in the document, and w_i is the weight of an activated term t_i. The factor $log(1 + (b + d)/b)$ favors documents with large depth and small breadth.

$$S_D = \frac{\sum_{i \in A} w_i \cdot E_i^f}{log(NT)} \cdot log(1 + \frac{b+d}{b}) \qquad (3)$$

The above establish a non-linear document evaluation function that takes into account the term correlations and topic-subtopic relations that the hierarchical network represents. The total amount of energy that an activated subhierarchy contributes to a document's relevance, amounts to its depth, and the weight of the terms and links involved. A document's relevance increases if it activates profile terms that formulate connected subhierarchies with large depths, and not isolated profile terms (e.g. term $DT3$ in figure 1 (right)). In this way, Nootropia has the ability to represent multiple topics of interest with a single profile and evaluate documents accordingly. Experiments have shown that it outperforms a traditional linear approach to document evaluation when two topics of interest are represented by a single profile [14].

4. Profile Adaptation

To adapt Nootropia's multi-topic profile to a variety of changes in a user's interests, we introduce in this section a process comprising five deterministic, but interwoven, steps that collectively allow the profile to self-organise in response to user feedback.

4.1. Step 1: Extract Informative Terms

The number of documents that received user feedback may vary from one to many. Here we describe adaptation based on one relevant or non-relevant documents, but the process may be easily generalised to more than one document. More specifically, the first of the five steps involves the weighting and extraction of informative terms. Given a feedback document D, after stop word removal and stemming[2], the online version of RelDF (equation 4) is applied to weight each unique term t in the document. In equation 4, n is the number of documents in a general, baseline collection that contain t and N is the total number of documents in that baseline collection.

To extract the most informative terms some threshold is required. Here we experimented with a fixed threshold equal to 0.3. An adjustable threshold is also possible. The term extraction process results in a set of weighted terms, some of which may already appear in the profile and some may not.

$$RelDF_t^o = w_t^D = \frac{1}{20} - \frac{n}{N} \qquad (4)$$

2 common feature dimensionality techniques

4.2. Step 2: Update Profile Term Weights

The second step of the process concentrates on those extracted terms that already appear in the profile. For each such profile term t, an updated weight w'_t is calculated using equation 5. In the case of a relevant document D, w'_t is calculated by adding to the profile term's weight w_t, its weight w_t^D in the document. In the case of a nonrelevant document, w_t^D is subtracted from w_t. The weight of profile terms that don't appear in the extracted set remains unchanged ($w'_t = w_t$). No learning coefficient is used in any of the cases.

$$w'_t = \begin{cases} w_t + w_t^D & \text{if } D \text{ relevant} \\ w_t - w_t^D & \text{if } D \text{ nonrelevant} \\ w_t & \text{if } t \ni D \end{cases} \tag{5}$$

Subsequently, in the case of a relevant document, we sum up the additional weights that have been assigned to the profile terms and then substract this sum evenly from all profile terms. This process is expressed by equation 6, where NP is the number of profile terms. The opposite takes place in the case of a nonrelevant document. Therefore, given a profile with a specific set of terms, this last process assures that the overall weight of profile terms remains stable.

$$w''_t = \begin{cases} w'_t - \dfrac{\sum_{t \in D} w_t^D}{NP} & \text{if } D \text{ relevant} \\ w'_t + \dfrac{\sum_{t \in D} w_t^D}{NP} & \text{if } D \text{ nonrelevant} \end{cases} \tag{6}$$

The net-effect of the above process is an appropriate redistribution of profile term weights that causes a change in the hierarchy's ordering. For example, profile terms that have been extracted from a relevant document have their weight reinforced while the weight of the rest of the profile terms decreases. The reinforced terms climb higher in the hierarchy, while the rest of the terms fall lower.

4.3. Step 3: Remove Incompetent Profile Terms

A side-effect of the decrease in the weight of some profile terms, which is caused either implicitly in the case of a relevant document, or explicitly in the case of a nonrelevant, is that some of them "run out of weight". In our case this means that the weight of some terms becomes less than zero. In this third step, terms that run out of weight are purged from the profile together with all of their links to other terms. Therefore, terms that were mistakenly added to the profile or have become incompetent due to the changes in the user interests are removed from the profile. At the same time we sum up the initial weight, i.e. the weight with which a term had entered the profile (see next section), of the purged profile terms (equation 7). The reason for this will be explained in the next section.

$$W_{purged} = \sum_{t \text{ purged}} w_t^{init} \tag{7}$$

4.4. Step 4: Add New Terms

Having updated the weight of profile terms and removed incompetent terms, at this step, those terms that have been extracted from a relevant document D and do not already appear in the profile are added to the profile. The initial weight of each added term in the profile is equal to the term's weight in the document ($w_t^{init} = w_t^D$). With the addition of each new term the weight of existing profile terms is not altered and therefore the overall weight of profile terms increases. The number of terms that are added depends on the semantic novelty of the relevant document in relation to what is being already represented. A document about a topic that is not already covered by the profile will contribute a lot of new terms and vice versa. We should also stress, that the added terms do not replace terms that have been purged in the previous step. The number of profile terms is therefore not fixed, but rather changes dynamically according to user feedback.

Finally, after the new profile terms are added, we substract evenly from all profile terms the sum of the initial weights of those terms that have been purged in the previous step, as expressed by equation 8, where NP' is the number of profile terms after the addition of new terms. Practically, this is done to avoid the escalating of the overall weight of profile terms due to the addition of new weight with every new term. Its importance in terms of self-organisation however, is that it renders the profile open to the environment: weight (energy) flows through the profile. The amount of weight (energy) that every new term adds to the profile is removed from the profile when and if the term is purged.

$$w'''_t = w''_t - \frac{W_{purged}}{NP} \tag{8}$$

4.5. Step 5: Reestablish Links

In this fifth final step we turn to link updating. For this purpose we refer back to the second step of the profile generation process (sec. 3). We observe that the same link generation process that was applied to a set of interesting documents can also be applied in a per relevant document fashion. All of the parameters of equation 1 can be updated online for each relevant document D, using equations 9 to 11, where $dist_{ij}$ is the aggregate distance between terms t_i and t_j in the documents processed so far.

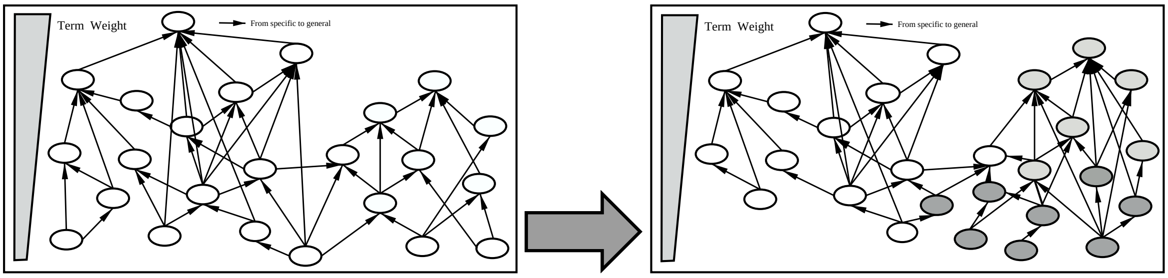

Figure 2. The effect of adaptation on the profile

$$fr'_i = fr_i + fr_i^D \quad (9)$$
$$fr'_{ij} = fr_{ij} + fr_{ij}^D \quad (10)$$
$$d'_{ij} = \frac{dist'_{ij}}{fr'_{ij}} = \frac{dist_{ij} + dist_{ij}^D}{fr'_{ij}} \quad (11)$$

So after adding the new terms, the relevant document is processed using a window of size 10 to identify links between profile terms and update the above parameters using the aforementioned equations. Once links have been established and the parameters updated, the weight of new links and the updated weight of existing links is calculated using the original equation 1 of section 3.

We should however note that it is practically inefficient to maintain in memory and keep updating the involved parameters for all of the documents processed so far. Although this has not been a problem for the experiments described bellow, one can either use normalisation or maintain in memory the frequencies and aggregate distance for a fixed number of the most recently processed documents.

4.6. Overview

These deterministic, interwoven, steps involve the weighting and extraction of informative terms from a feedback document, the updating of the weight of profile terms, the removal of incompetent terms, the addition of new terms and finally the identification and weighting of links. Figures 2 depicts the overall effect in the case of a profile about two topics of interest and a feedback document about one of them. The process causes an increase in the depth of the hierarchy corresponding to the topic discussed in the feedback document and a decline in the depth of the hierarchy corresponding to the topic that did not receive positive feedback. The process enables the profile to respond to feedback with structural modifications, which of course affect document evaluation.

We may now argue that Nootropia exhibits self-organising characteristics. During the second half of the 20th century various theoretical models were developed to account for self-organisation [4]. However, they all share three common characteristics:

1. Their interconnectedness renders self-organising systems non-linear.
2. Self-organising systems are open systems—energy and matter flow through the system—that operate far from equilibrium.
3. New structures and new modes of behavior are created in the self-organisation process.

Indeed, in Nootropia a hierarchical network of terms is used to establish non-linear evaluation of documents according to a user's multiple interests. During the above process weight (energy) flows through the profile with the addition and removal of terms, causing constant structural change. Nootropia is open to its environment, the user feedback, and operates far for equilibrium, never settling to a stable state. As a result of this constant structural modification new structures (hierarchies) and thus new modes of behaviour (document evaluation) are created. Nootropia complies with all three characteristics of self-organisation. The question that arises at this point is: can Nootropia adapt effectively through self-organisation to a variety of changes in a user's interests?

5. Experimental Evaluation

To answer the above question we present in this section some first experiments using virtual users which produced positive results. More specifically, we have established an experimental methodology using a variation of the 10th Text Retrieval Conference's (TREC-2001) routing subtask[3]. TREC-2001 adopts the Reuters Corpus Volume 1 (RCV1), an archive of 806,791 English language news stories that has recently been made freely available for research purposes[4]. The stories have been manually categorised accord-

3 For more details see: http://trec.nist.gov/data/t10_filtering/T10filter_guide.htm
4 http://about.reuters.com/researchandstandards/corpus/index.asp

Learn a new topic	
$l.1$	$R6/R21 \to R6/R21/R20$
$l.2$	$R41/R79 \to R41/R79/R58$
Forget a topic	
$f.1$	$R6/R21/R20 \to R6/R21$
$f.2$	$R41/R79/R58 \to R41/R79$
Penalise a topic	
$p.1$	$R6/R21/R20 \to R6/R21/\neg R20$
$p.2$	$R41/R79/R58 \to R41/R79/\neg R58$

Table 1. Simulated interest changes

ing to topic, region, and industry sector [16]. The TREC-2001 filtering track is based on 84 out of the 103 RCV1 topic categories. Furthermore, it divides RCV1 into 23,864 training stories and a test set comprising the rest of the stories

If we assume that changes in a user's interests are reflected by variations in the distribution of feedback documents about different topics, then we may simulate interest changes in the following way. Given the above classification, a virtual user's current interests may be defined as a set of classification topics (e.g. $R1/R2/R3$) [23]. A radical, long-term change of interest may then be simulated by removing or adding a topic to this set. For example if the user in no more interested in topic $R3$ then we may denote such a change as $R1/R2/R3 \to R1/R2$. Similarly, we present here results for three kinds of simulated interest change and two topic combinations (table 1).

The first kind of interest change involves an initial interest in two topics and then the emergence of a third topic of interest. As already described in the example, in the second case the user is no more interested in one of the initial three topics. The third kind of interest change is similar to the second with the difference that the user explicitly indicates the change of interest through negative feedback (denoted with "¬").

For each of the above changes a profile is initially adapted online according to the initial topics of interest. For that purpose we use a set of documents comprising the first 30 documents per topic in the training set. These training documents are ordered according to publication date and therefore their distribution is not homogeneous, but rather reflects the temporal variations in the publication date of documents about each topic. It reflects in that sense fast short-term changes in the virtual user's interests.

The same process is then followed using the first 30 training documents per topic in the set following the change of interest. Training documents that correspond to negated topics have been used as negative feedback. During this second adaptation phase the profile is used every five training documents to filter the complete test set. It is then evaluated on the basis of an ordered list of the best 3000 scoring documents, using the *Average Uninterpolated Precision* (AUP) measure. The AUP is defined as the sum of the precision value–i.e. percentage of filtered documents that are relevant–at each point in the list where a relevant document appears, divided by the total number of relevant documents. A separate AUP score was calculated for each topic.

Figures 3, 4 and 5 present the experimental results. Each graph represents for each topic the fluctuation of AUP score during the second adaptation phase. The AUP score of the topic to be learned, forgotten or penalised through negative feedback is depicted with a solid line while the rest of the topics with dashed line. Whenever required a second Y-axis has been used to account for large differences in the AUP score of the involved topics.

In the case of an emerging new topic of interest, figure 3(a) indicates a relative increase in the AUP score of the new topic R20. At the same time we may also observe symmetric fluctuations in the AUP for the existing two topics of interest R6 and R21, which reflect the aforementioned variations in the distribution of the corresponding documents in the training set. A more clear initial increase in the AUP score of the new topic R58 is depicted by figure 3(b). The subsequent drop corresponds to a drop in the density of relevant documents in the training set.

When the virtual user looses interest in one of the initial three topics, figure 4(a) reveals a drop in the AUP score of topic R20. Symmetric fluctuation in the scores of the persistent two topic R6 and R20 are again observed. The drop in the score of topic R58 is also obvious (fig. 4(b)). It is accompanied by an increase in the score for topic R79 which then drops symmetrically to an increase in the score for topic R41.

Finally, when the virtual user explicitly provides negative feedback on the no more interesting topic, figure 5(a) indicates a drop in the score for topic R20, that is clearly larger than in the case of figure 4(a). A similar observation can be made for topic R58 in figure 5(b). The corresponding line exhibits a steeper drop. Symmetric fluctuation in the score of the persistent topics are again apparent in both cases.

In summary, the above experimental results indicate Nootropia's ability to adapt in respond to variations in a stream of feedback documents. Through self-organisation, Nootropia appears to be able to learn an emerging topic of interest (fig. 3) and forget a topic that is no longer interesting (figures 4 and 5). In addition to these radical long-term changes, the results show that Nootropia can quickly reflect fast short-term variations in user interests. Therefore, both kinds of change are tackled with a single, multi-topic profile.

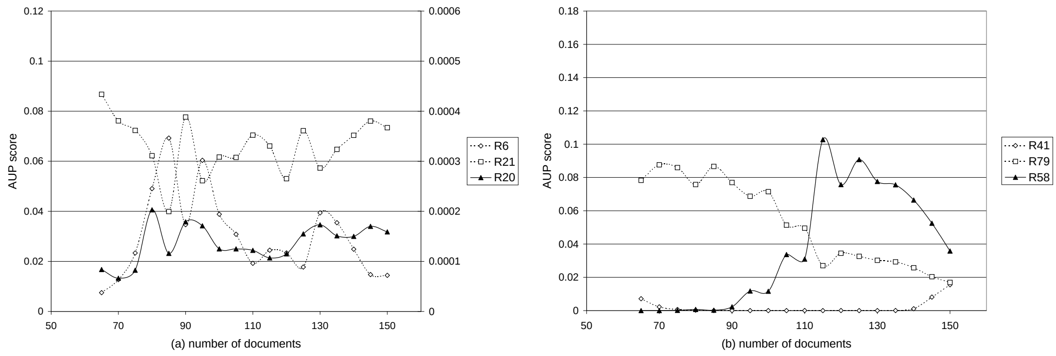

Figure 3. Learning a third topic of interest: a) $l.1$ **b)** $l.2$

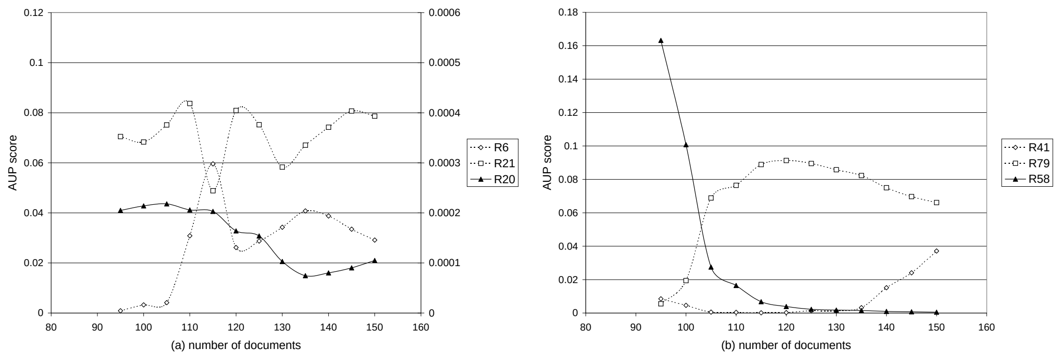

Figure 4. Forgetting the third topic of interest: a) $f.1$ **b)** $f.2$

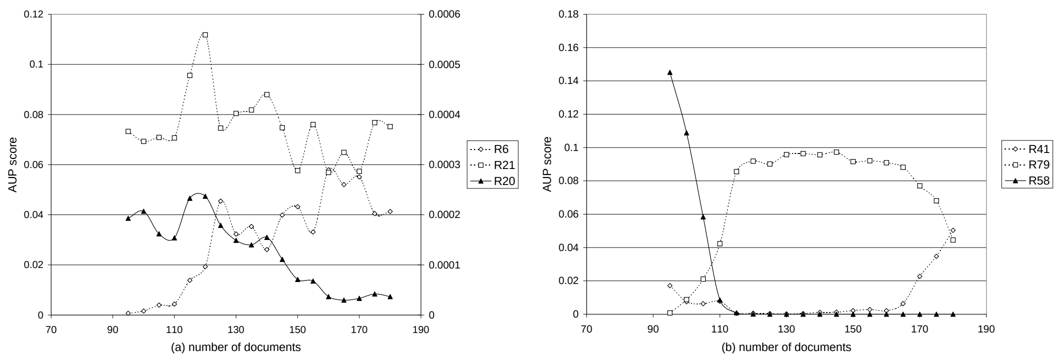

Figure 5. Negating the third topic of interest: a) $p.1$ **b)** $p.2$

6. Summary and Future Research

We have presented Nootropia, a self-organising information agent, capable of evaluating documents according a user's multiple and changing interests. We employed a hierarchical term network to represent more than one topics of interest with a single profile and established on that network non-linear document evaluation. We then introduced a process that adjusts the profile structurally in response to user feedback. The process renders Nootropia open to its environment, causing it to operate far from equilibrium. New hierarchies develop to reflect emerging topics of interest and existing hierarchies that represent unexcited topics can disintegrate and be purged from the profile. Such structural modifications affect document evaluation. Arguably Nootropia exhibits the three basic characteristics of self-organisation. Initial experiments show that through self-organisation, Nootropia may successfully adapt to a variety of interest changes. So instead of breaking the problem into multiple profiles and adaptation levels, we tackled it with a single, multi-topic and self-organising profile.

We have not focused on a specific application area. We are looking at applying Nootropia for various personalisation services, like the filtering of incoming documents, autonomous search, expert finding and collaborative filtering. We should also note at this point that although we concentrated on document evaluation, the same approach can in principle be applied to other media like image and audio, for which features can be automatically extracted. Such a research direction would further increase Nootropia's scope.

References

[1] Paul E. Baclace. Competitive agents for information filtering. *Communications of the ACM*, 35(12):50, 1992.

[2] D. Billsus and M. Pazzani. A hybrid user model for news story classification. In *7th International Conference on User Modeling*, Banff, Canada, June 1999.

[3] D. Billsus and M. Pazzani. A personal news agent that talks, learns and explains. In *3rd International Conference on Autonomous Agents*, Seattle, WA, 1999.

[4] F. Capra. *The Web of Life*. Harper Collins, 1996.

[5] W. Lam, S. Mukhopadhyay, J. Mostafa, and M. Palakal. Detection of shifts in user interests for personalized information filtering. In *19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 317–325, Zurich, Switzerland, 1996.

[6] P. Maes. Agents that reduce work and information overload. *Communications of the ACM*, 37(7):30–40, 1994.

[7] M. McElligott and H. Sorensen. An emergent approach to information filtering. *UCC Computer Science Journal*, 1(4), 1993.

[8] F. Menczer and R.K. Belew. Adaptive information agents in distributed textual environments. In *2nd International Conference on Autonomous Agents*, Minneapolis, MN, 1998.

[9] F. Menczer and A. E. Monge. Scalable web search by adaptive online agents: An infospiders case study. In M. Klusch, editor, *Intelligent Information Agents: Agent-Based Information Discovery and Management on the Internet.*, pages 323–347. Springer-Verlag, 1999.

[10] J. Mostafa, S. Mukhopadhyay, M. Palakal, and W. Lam. A multilevel approach to intelligent information filtering: model, system, and evaluation. *ACM Transactions on Information Systems (TOIS)*, 15(4):368–399, 1997.

[11] A. E. Motter, A. P. S. de Moura, and P. Dasgupta. Topology of the conceptual network of language. *Physical Review E*, 65(065102(R)), 2002.

[12] A. Moukas and P. Maes. Amalthaea: An evolving multi-agent information filtering and discovery system for the www. *Autonomous Agents and Multi-Agent Systems.*, 1(1):59–88, 1998.

[13] N. Nanas, V. Uren, A. De Roeck, and J. Domingue. Building and applying a concept hierarchy representation of a user profile. In *26th Annual International ACM SIGIR International Conference on Research and Development in Information Retrieval*, pages 198–204. ACM press, 2003.

[14] Nikolaos Nanas, Victoria Uren, Anne de Roeck, and John Domingue. Multi-topic information filtering with a single user profile. In *3rd Hellenic Conference on Artificial Intelligence*, 2004.

[15] J. Rocchio. *Relevance Feedback in Information Retrieval*, chapter 14, pages 313–323. Prentice-Hall Inc., 1971.

[16] T. Rose, M. Stevenson, and M. Whitehead. The Reuters Corpus Volume 1 - from yesterday's news to tomorrow's language resources. In *3rd International Conference on Language Resources and Evaluation*, 2002.

[17] G. Salton and M. J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill Inc., 1983.

[18] R. Schapire, Y. Singer, and A. Singhal. Boosting and Rocchio applied to text filtering. In *21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1998.

[19] Y. Seo and B. Zhang. A reinforcement learning agent for personalized information filtering. In *Intelligent User Interfaces*, pages 248–251, New Orleans, LA, 2000.

[20] B. D. Sheth. *A Learning Approach to Personalized Information Filtering.* Master of Science, Massachusetts Institute of Technology, 1994.

[21] H. Sorensen, A. O' Riordan, and C. O' Riordan. Profiling with the informer text filtering agent. *Journal of Universal Computer Science*, 3(8):988–1006, 1997.

[22] D. H. Widyantoro, T. R. Ioerger, and J. Yen. An adaptive algorithm for learning changes in user interests. In *ACM/CIKM'99 Conference on Information and Knowledge Management*, Kansas City, MO, 1999.

[23] D. H. Widyantoro, T. R. Loerger, and J. Yen. Learning user interests dynamics with a three-descriptor representation. *JASIS*, 2000.

[24] W. Winiwarter. PEA - a personal email assistant with evolutionary adaptation. *Internation Journal of Information Technology*, 5(1), 1999.