DOCUMENT RESUME

ED 466 580 IR 021 315

AUTHOR	Dron, Jon; Boyne, Chris; Mitchell, Richard
TITLE	Footpaths in the Stuff Swamp.
PUB DATE	2001-00-00
NOTE	7p.; In WebNet 2001: World Conference on the WWW and Internet Proceedings (Orlando, FL, October 23-27, 2001); see IR 021 310.
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Computer Mediated Communication; Computer Uses in Education; *Systems Development; Teacher Role; *Web Based Instruction; World Wide Web
IDENTIFIERS	Electronic Resources; *Learning Environments

ABSTRACT

This paper discusses stigmergy (i.e., the effect of communication through the environment) in relation to the Internet, especially with regard to World Wide Web-based learning. The paper begins by examining ways in which stigmergy occurs on the Web and then goes on to describe its use in the construction of a continually evolving system, CoFIND (Collaborative Filter In N Dimensions). CoFIND's purpose is to replace the role of a traditional teacher in structuring and selecting learning resources. It attempts to achieve this through a process of stigmergy and natural selection, leading to a degree of self-organization brought about through the independent actions and interactions of its individual users. The paper discusses the issues raised and explains some its successes and failings. (Contains 23 references.) (Author/MES)

Footpaths in the Stuff Swamp

Jon Dron (University of Brighton, UK)
Chris Boyne (University of Brighton, UK)
Richard Mitchell (Inferdata Corporation, US)

Abstract

In this paper we consider stigmergy (loosely, the effect of communication through the environment) in relation to the Internet, especially with regard to Web-based learning. We begin by examining ways in which stigmergy occurs on the Web, before going on to describe its use in the construction of our continually evolving system, CoFIND (Collaborative Filter In N Dimensions). CoFIND's purpose is to replace the role of a traditional teacher in structuring and selecting learning resources. It attempts to achieve this through a process of stigmergy and natural selection, leading to a degree of self-organisation brought about through the independent actions and interactions of its individual users. We discuss some of the issues raised and attempt to explain some of its success as well as account for its failings.

Stigmergy

Stigmergy is a word coined by Grassé to refer to systems such as those employed by termites when building mounds (Heylighen, 1999). Termites build mounds start by dropping mud randomly in a given area. The presence of a mud heap encourages other termites to drop their lumps of mud nearby, with larger mounds being more attractive than smaller ones. Adjacent mounds therefore tend to grow towards each other, so forming arch structures. Stigmergy plays a role in many self-organised systems. It explains the formation of ant-trails, for example. If an ant finds food, it leaves a trail of pheromones on its way back to the nest. When other ants encounter this trail, they are inclined to follow it. If they too find food, they too leave a trail of pheromones back to the nest. As more ants find food, so the scent grows stronger. Eventually the food runs out, no more ants leave a pheromone trail and the scent dissipates. The ant trail ceases to be. Stigmergy also plays a major role in human dealings. This can for example be seen in the spontaneous formation of footpaths in a forest, recessions caused by fluctuations in stock prices and the large crowds that gather round smaller crowds in a street (an effect well utilised by professional buskers). A number of simple individual acts which affect the environment in turn affect other people, directly or indirectly. Coordinated behaviour arises as a result of indirect communication processes.

Stigmergy and the search for useful Web pages

The Web is composed of a vast number of individuals interacting with Web sites and each other. Much of this interaction is recorded, which suggests the possibility that something akin to footpaths in a forest could form, although many of those records tend to be hidden in individual server logs. Crawford characterises the Web not as a super-highway but as " a swamp, albeit a swamp with many remarkable hillocks of well-organized, first-rate data and information." (Crawford, 1999). If some sense or pattern could be discerned within the stuff swamp, a system of footpaths leading to useful hillocks, then navigating the Web would be a simpler affair. There are two clear potential approaches to the building of such footpaths: by design, or by stigmergy.

Footpaths by design

Yahoo represents the pinnacle of achievement in the creation of footpaths by design. Its large group of researchers trawls through thousands of web sites every day, weeding out the poor and ill-conceived, categorising and organising the gems that they find. Despite a bizarre classification scheme which puts *Education* next to

Entertainment, Arts and Humanities next to *Reference* which in turn sits next to *Regional* (Hudon, 2000), Yahoo's power is that it is organised by beings who understand meaning in a human context. This distinguishes it from the various web crawlers whose algorithms mechanically extract some kind of meaning from the content of what they are indexing. There are disadvantages with designing footpaths, a big one being that an individual's view or that of a small group will not necessarily coincide with that of the general population of users. Even where categories are agreed upon, unless the community of researchers has a similar set of interests and goals to those of the users of the system, relevance may be low. Because of this, a niche market has developed for more specialised directories or subject gateways, created by groups of like-minded researchers with an interest in a given field. Educational technologists might use The World Wide Web Virtual Library of Education (http://tecfa.unige.ch/info-edu-comp.html) or the TIP database (http://www.gwu.edu/~tip/) for example.

Footpaths by emergence

Recent generations of Web search engines, such as Clever (http://www.almaden.ibm.com/cs/k53/clever.html), Alta Vista's Raging (http://www.raging.com), and Google (http://www.google.com) make explicit use of "footprints" in the form of what Kleinberg (1999) describes at Latent Human Annotation. Such systems iteratively mine the Web for clusters, based on an analysis of in-degree and out-degree in hubs and authorities. Implicit recommendations based on what is effectively a form of citation (links imply approval) are condensed and purified, resulting in emergent patterns indicating popular and probably relevant sites of interest. A similar idea underlies the seminal PHOAKS (People Helping One Another Know Stuff, http://www.phoaks.com), which parses newsgroup messages for URLs, then bases its categorisations on topics and themes which relate to them. The quantity of citations in a relevant category provides the rating of a resource as useful or not.

Stigmergy is exploited by the family of systems generally referred to as recommender systems, collaborative filters or social filters. PHOAKS and Google are both examples of implicit recommender systems, which take a form of existing behaviour and extrapolate or infer preferences from what people are already doing. Amazon takes a similar approach when recommending books based on matching your purchases with those of others. The other main class of recommender system is the explicit variety, exemplified by Firefly (Resnick & Varian, 1997). Here, users are actively polled for their opinions on items such as films, books or Web sites, and their opinions are matched with those of other users to provide recommendations. By taking explicit recommendations, areas of uncertainty found in implicit systems are removed. In an implicit system, backlinks to Web sites may well be used in unexpected contexts which do not indicate a genuine recommendation, such as "this a dreadful site." Similarly, Amazon may often be used to purchase books as gifts and thus implicit recommendations might not reflect a person's own taste. However, if the user has explicitly rated a set of books then we can assume likes and dislikes with a high degree of confidence. The disadvantage of explicit recommendations is the demand that they make on users, leading to the cold-start phenomenon: if there are no ratings, the system is useless, as there will be nothing to match one user with the next. Early users will thus achieve disappointing results and (in a negative feedback reaction) will cease to use the system, which will therefore never grow. None-the-less, collaborative filters show interesting stigmergic patterns. Imagine a greatly simplified system with only two matched users. If user Adam recommends books A, B, C and D, Bob recommends books B, C, D and E, because their ratings intersect each will be recommended to read a book which they have not already read. Assuming that each likes the book that has been recommended, then they in turn will recommend them, leading to a stigmergic reaction. Without any intentional act of communication, an organised cluster will have developed.

Stigmergic communities

Community interactions, both synchronous and asynchronous, can provide some indication of the paths their users are taking. For example, it is easy to spot more popular newsgroups by the number of postings. Like crowds that gather around a busker in the street, the greater the size of the crowd the more likely it is that others will join it. More subtly, there are often clues which may be gleaned from the subject lines and hierarchical thread patterns themselves. For example, flames tend to follow recognisable patterns and word usage which will either lead others to read those threads or (perhaps more often) to avoid them. Similarly in Chat rooms, people tend to be drawn to areas where there is the highest activity. Once there, quantity of words is often a broad indicator of the quality of a given discussion. Short, social messages may indicate less depth than longer, involved messages. There is therefore a stigmergic form of self-organisation that inhabits even the least structured forms of

discussion. However, its operation is typically crude and undiscerning. Only the broadest of patterns may be visible and, as a form of self-organisation, it is no smarter than the rule which makes a group of birds nest together.

Enhancing stigmergy to structure communication

Stigmergy is exploited by a number of communication systems to achieve subtler and easier forms of organisation than would be achievable by their more primitive counterparts. A superb example of such a system is ChatCircles (Donath et al, 1999). In this system, each participant is represented by a circle. The more a person participates, the larger the circle grows, shrinking with time like the pheromone trails left by ants. It is only possible to "hear" people (see what it is that they are typing) by moving your own circle close to them. Thus, drawn by larger circles and larger groups, clusters of activity form naturally as a result of participation in the process. Two cues are available to the user, the size of the circles (indicating activity) and the number of circles in a cluster (indicating levels of participation and interest). Neither cue is the result of an explicit attempt to communicate with other users; they are simply the result of rational behaviours within the system.

Odigo (http://www.odigo.com) takes another approach, a synchronous chat system that centres discussions around web pages and web sites. Visitors to particular pages can discuss shared interests or the contents of those pages. Users can discover "What's Hot Now" identifying those pages that are most in use by Odigo users in real time. This feature enhances the users' ability to identify clusters and areas of interest. Popular sites draw more users, which makes them more popular and so the self-reinforcing circle continues. A related approach is taken by uTOK (http://www.utok.com), which allows asynchronous interactions through the ability to leave notes relating to a given site. As the user browses the web, the small uTOK panel indicates whether any messages relating to that site have been left by other users. The value of a site is indicated by the number of interactions, with users being drawn to sites with larger discussions. uTOK combines this with a voting mechanism for notes that have been left, which gives some indication of the quality of a discussion as well as the quantity of contributions. This multi-dimensional approach allows for richer forms of stigmergy than are afforded by, say, simple clustering alone. Through the application of simple local rules, complex patterns that convey sophisticated information to users of such systems evolve and develop. uTOK takes this evolutionary metaphor to its logical conclusion, as low-rated messages are removed from the system. Thus, both the structure and content of a given cluster of knowledge are determined by individual local interactions.

CoFIND

Over the past three years we have been developing a Web-based system known as CoFIND (Collaborative Filter In N Dimensions, first reported in Dron et al, 1999) which seeks to organise itself through the combined individual actions of its users. In essence, CoFIND is a self-organised database of resources, akin to a subject gateway, with a small, relevant body of resources created and found collaboratively by its users. It is written using Microsoft's ASP with a database backend.

CoFIND utilises evolution, in a Darwinian sense. Through a process of voting, resources (mostly Web sites) compete by moving up or down lists of returned results. The landscape on which they compete is variegated, with individual niches determined by the use of *qualities*, the things that users value in a resource. Typical qualities might be "useful," "amusing," or "good for beginners." A differently ordered list of resources would be returned by users seeking *useful* resources than by those seeking *amusing* resources. Like resources, qualities compete, changing size and moving up or down the list according to how frequently they are selected and how often they are used to vote for resources. If they are not much used, they will become "extinct". Users also decide upon binary categories for the resources, characterised in CoFIND as *topics*. As resources, topics and qualities are created by the users themselves; the system orders itself and is regulated according to the needs of its participants.

CoFIND is designed to be used by a single cohort of adult learners with similar learning goals starting at a similar state of knowledge. If there is too much disparity between the needs of the users, then whatever shapes and patterns emerge will be chaotic. As Darwin recognised, evolution and adaptation works most effectively in small, parcellated populations such as those found on islands. Large populations with high levels of connectivity result

in most significant changes and mutations being cancelled out by a process of dilution and competition. By limiting the participants to a small and focussed cohort, CoFIND results in highly focussed subject gateways.

The intention behind CoFIND is to create a kind of group intelligence that takes on some of the roles of a teacher. One of these roles is to select and recommend source material and texts for students to read, matched to the needs of those students. Through the use of qualities to rate resources and topics to categorise them, students are able to achieve this goal themselves, without central guidance. We have achieved a measure of success in this goal (Dron et al, 1999, Dron et al, 2000). However, a number of obstacles still remain, most notable of which are those of encouraging sufficient participation to overcome cold-start problems, and of achieving meaningfully self-organised results. Lately, we have been attempting to partially address these problems by developing CoFIND along more stigmergic lines, allowing richer patterns to develop than are possible through evolution alone.

Stigmergy and CoFIND

The first place we sought to introduce stigmergy was in the selection of topics, a solution that arose as a result identifying the wickedness of the problem of generating collaborative paths. We wished not only for users to be able to select appropriate resources for their learning needs, but (as a teacher would do) to select them in an appropriate order. However, the dependency of one topic on another, and the need to understand one resource before embarking on the next is an intractable problem. In most teaching there is a narrative thrust. The appropriate order in which tasks should be attempted is dependent on what has come before and what comes after. It would not be appropriate to average-out, concatenate, or otherwise join sets of elements from one path with those of another, any more than it would make sense in most novels to jumble up the order of the chapters. The meaning of a set of resources is intimately connected to the order in which they appear. Our solution was not to create a single path by combining others, but to attempt to shape the route through a subject area in real time, to collaboratively beat out the path through the topics as we go.

Stigmergy and topics

Topics are generated by the users of a CoFIND system. They may be entered by any user into one of four distinct, but undifferentiated, screen areas. This number was decided on purely pragmatic grounds, calculated according to how many words could be squeezed onto one average 800x600 screen and the assumption that an appropriate number of topics in any sector would be around seven, which is approximately the number of items that can easily be assimilated at a glance (Miller, 1956). Within a given area, topics are in competition with each other. Every time a topic is selected, it gains emphasis by getting larger, whilst other topics in the same area get smaller. The algorithm is balanced so that this change is not too dramatic, allowing for a certain amount of stickiness before topics start to fade away.

Figure 1 shows a use of CoFIND within a traditionally taught classroom setting.

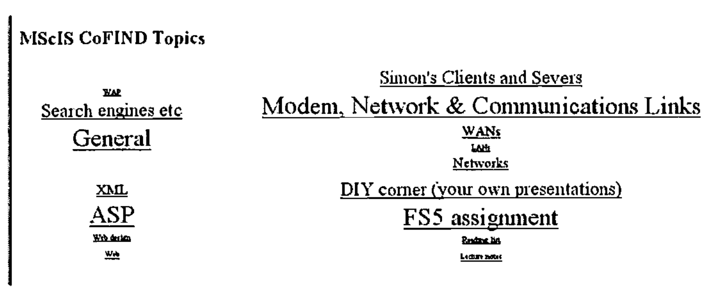

Figure 1: Simon's Clients and Severs

The prominent topics are mostly closely related to the development of the course at the time of this snapshot, with a project on ASP and a strong interest in XML in personal projects, for example. *Simon's Clients and Severs* [sic], however, is an anomaly. It was claimed to be a mistake by the student who added the topic, caused by a

misunderstanding of how to add some of his own work to the system. In fact, this is probably untrue, because the student repeatedly clicked the topic label over a period of a few minutes immediately after adding it, thereby increasing its relative sizeand prominence. It is (even when spelt correctly) a topic of only marginal interest in the context of the course. Despite this, the topic got to be quite popular, visited by twelve different students over two months. This compares with only ten visitors over a similar period to *ASP*, a topic which was being covered in class and thus of greater intrinsic interest. It is thus tempting to conclude that usage was being driven by stigmergy, kick-started by the student who added the topic in the first place.

Stigmergy and qualities

The second explicit use of stigmergy in CoFIND lies in the quality selection mechanism. Like topics, quality labels vary in size according to usage. For a quality, the size of its label is related to the number of times it has been used to rate resources. In addition, qualities are displayed in an order which is determined by an algorithm which primarily considers the of number of times they are selected, tempered by the number of times they are used to rate resources, and a novelty rating which decays rapidly as they age. It is thus quite common in the early stages of quality evolution to find smaller labels closer to the top of the list than larger ones, as shown in figure 2.

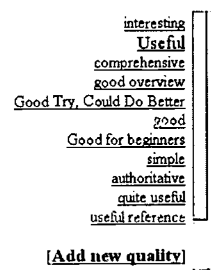

Figure 2: Stigmergy and qualities

Stigmergy in CoFIND is intended to enhance and emphasise patterns of behaviour. By providing different cues to the user, a more complex set of interactions can be elicited. The size of the quality indicates the relative number of explicitly rated resources which it will lead to, whilst position is an indicator of the relative perceived usefulness of the quality, gleaned largely from implicit usage statistics- users click the qualities which they believe will be more useful.

Users are encouraged to select a quality by its position and/or by its size. This process typically results in those qualities at the top of the list also having the largest labels as, once selected, a given quality will then be used to rate resources. The fact that this synchronisation does not occur straight away introduces a latency which magnifies the overall stigmergic effect, following a pattern like that of Senge's Beer Game (Senge, 1993). In the Beer Game, retailer, wholesaler and manufacturer are caught up in a vicious cycle of supply and demand which arises from the latencies within the system. Retailers, responding to a consistent slight increase in demand for a certain brand of beer, increase their orders from the wholesalers. The wholesalers increase orders from the brewer. Because of inherent latencies, partly due to delivery cycles and partly due to how long it takes to brew beer, orders are not immediately met, so everyone perceives increased demand and an inability to meet it. To compensate, the retailers increase orders to cope with the backlog, the wholesalers follow suit a little while later, the brewer brews more beer to cope. These effects are magnified as they rise through the system. By the time the backlogs of orders are met, all players in the system have a massive over-stocking of beer, orders are cancelled and everyone suffers accordingly. Within CoFIND, the analogous positive feedback loop is actually beneficial, leading to increased polarisation of the use of qualities, hence to a richer structure.

Discussion

The effect of stigmergy in CoFIND is used to create structure in a set of resources. By encouraging positive feedback loops, clustering occurs more readily than were the system simply to make use of qualities and ratings to decide which resources are valuable.

Our intention is to replace one role of the teacher with a collaborative group mind, a function that, by our own criteria, CoFIND is performing effectively. There is doubt, however, that this group mind can match that of even

the dullest of professors. Whilst it is true that in every one of the ten or so instances of CoFIND so far implemented students at all levels have found and rated a set of resources far better than those already discovered by the teacher alone, there are problems with the ways in which structure has developed. The example of *Simon's Clients and Severs* is a good illustration of how the system may easily be subverted by a determined anarchist or self-publicist. It would be easy enough to programmatically prevent such an occurrence in future, but the fact that the ploy was successful shows that stigmergy can as easily lead to useless or even harmful categorisations as to those which have pedagogic benefits.

Another unresolved issue lies in the related problems of overcoming resistance to using the system and the cold-start phenomenon. The former problem is partly one of interface design, an issue that is slowly being dealt with. However, it is also related to the latter problem, as the system is only useful once there are many resources, topics, qualities and ratings. Although the use of qualities should help to more effectively discern useful resources, the overhead in providing ratings for more than one quality makes the start even colder than that experienced by most other collaborative filters.

CoFIND's problems are exacerbated by the fact that the system is currently only being tested within a conventional University teaching environment. This means that the students always have alternative sources of knowledge to CoFIND, that they are motivated by the need to pass examinations rather than to learn about a given topic, and that there is a limit to the time they can allocate to what is perceived as a fairly peripheral activity. It is significant that the most prolific use of the system so far has been when it is used to provide formative peer assessment of the students' own web pages. Like any computer system, it does not exist in isolation and, as an adjunct to existing working systems, offers insufficient extra value to encourage widespread successful use.

In conclusion

We have begun to show the potential for stigmergic systems in structuring a web-based learning environment and identified some techniques for providing richness in that structure. Although we have identified some fruitful directions, the potential is still waiting to be fully exploited. Part of the reason for this lies in failure to consider the system as a whole, partly in the inherent flaws in collaborative filtering technologies. It is, however, apparent that the Web gives enormous potential for the development of group minds, composed of yet distinct from their constituent individuals and, given time and appropriate environments, that these minds may one day rival those of traditional teachers as a means of creating, distilling, structuring and disseminating knowledge.

References

Crawford, Walt. (1999) The Card Catalog and Other Digital Controversies. *American Libraries* 30:1 (Jan 1999), 52

Dron, Jon, Mitchell, Richard, Siviter, Phil & Boyne, Chris. (1999) CoFIND- an experiment in n-dimensional collaborative filtering. *WebNet 99*, AACE, Hawaii , 301-306

Dron, Jon, Mitchell, Richard, Siviter, Phil & Boyne, Chris. (2000) CoFIND: steps towards a self-organising learning environment. *WebNet 2000*, AACE, San Antonio, 146-151

Heylighen F. (1999): Collective Intelligence and its Implementation on the Web: algorithms to develop a collective mental map. *Computational and Mathematical Theory of Organizations* 5(3), 253-280.

Hudon, Michèle (2000) Innovation and Tradition in Knowledge Organization Schemes on the Internet, or, Finding One's way in the Virtual Library. *Sixth International ISKO Conference*, 10-13 July 2000, Toronto, Canada, Würsburg: Ergon Verlag, 35-40

Kleinberg, Jon M. (1998) Authoritative Sources in a Hyperlinked Environment. *Proceedings of the 9th ACM-SIAM Symposium on Discrete Algorithms*, ed. H. Karloff (SIAM/ACM -SIGACT, 1998)

Miller, George A. (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *The Psychological Review*, Vol 63, 81-97

Resnick, Paul, Varian, Hal (1997) Recommender Systems. *Communications of the ACM*, March 1997, Volume 40, Number 3

Senge, Peter M. (1993) *The Fifth Discipline- the art and practice of the learning organization.* Chatham: Century Business

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").