

Christoph Mueller
University of Berne

«Turn it on! -- How the Swiss <CHancelbot> came to be»

PAPER presented at the EASST CONFERENCE 2002
in York (UK), August 1 2002
DRAFT

ABSTRACT (max 300 words):

Sooner or later, every community faces problems of maintaining its social structure. This is especially true for text based «online communities» on the Internet. The easy entry and exit options, the lack of formal membership and moderation, and the restriction of interaction to text makes it difficult to establish rules, to sanction unwanted behaviour, and to maintain continuity. As a consequence, a group may split up, or even disappear.

Our case study shows how in 1998, some members of Usenet newsgroups of the Swiss <ch.>-hierarchy requested the introduction of a technical agent, a «bot», with the special task to ensure a certain «group order». The so-called <CHancelbot> would be capable to cancel «faulty» messages, i.e. to delete them from the distributed newsgroup servers. Basically a short computer script, the <CHancelbot> is a very simple agent, with only little autonomy and no «personality». However, it was argued that it may be a useful tool helping to eliminate «duplicates» and «excessive crosspostings». Yet, as the bot would act for itself, it could make mistakes. After all, a cancelling agent is a «killerbot»! Given that the filtering software could be extended to delete messages containing certain keywords, censoring them thus by its *content*, the CHancelbot raises some important questions: Under what rules should anyone be allowed to *kill* messages of someone else? Should this task better be done by a machine or by human beings? What are the risks of delegating power to a machine?*

The discussion about the introduction of the <CHancelbot> reveals different views on the balance between technical on social means of maintaining social order. The study shows how the meaning and the task of the CHancelbot are negotiated between different users. Would the agent really help to solve the social problems of «misuse» of the newsgroups?

KEYWORDS: Internet, Agents, Sociology

CONTACT:

Christoph Müller, muellerc@soz.unibe.ch

DRAFT -- for updates, see
<http://www.soz.unibe.ch/forschung/ii/virt>

1) Introduction

My presentation is about the introduction of an automated agent called «CHancelbot» in a newsgroup of the Swiss hierarchy of the Usenet. The research is part of my PhD thesis studying online communication services, like chats and newsgroups. These services allow users to maintain existing social relations over long distances, as well as to build new contacts with people they probably never met face-by-face.

As the communication services I am studying are not formally moderated, participants are confronted with some specific problems in establishing and in maintaining a certain social «order». How do they manage these problems?

The question I am asking is not restricted to online communities. Often, decisions about how to use collective goods, as well as conflict resolution, is delegated to an external authority. As Elinor OSTROM has shown in her broad study on social groups facing problems of organizing the use of «common pool resources», social groups can be very successful in «governing the commons» without referring to an external «Leviathan». Ostrom defines some criteria for a successful «bottom-up» regulation, like public monitoring of the actions, differentiated modes of sanctioning, procedural rules adapted to local needs and conditions (*figure 1*).

In studying chats and newsgroups, I am not dealing with the same kind of collective goods Elinor Ostrom did, but as there is no central authority on the Usenet, some strategies of conflict resolution may be similar.

The data of the following presentation consists of the publically available interaction in a newsgroup of the Swiss hierarchy of the Usenet, called <ch.talk>. I am studying the content of about 30'000 messages collected between October 1997 and October 1999. In observing the interactional behavior of the participants, I am focussing on the strategies they apply in managing conflicts.

As I will show, the task of maintaining a social «order» is especially difficult to be fulfilled in text based online communities on the Internet. There are different reasons for this fact, based on particularities of computer-mediated communication (CMC):

- The easy entry and exit options,
- the lack of formal membership (e.g., passwords),
- anonymity / pseudonymity
- multiple identities
- no formal moderation,
- interaction is restricted to text only.
- no possibility for sanctioning the body.

These particularities may lead to a great amount of unwanted behavior - in any case, they make it difficult to establish rules, to sanction unwanted behaviour, and to maintain a stability and a continuity.

2) The Usenet as a research field

Figure 2 -- Distributed Usenet Network

The Usenet - or Netnews - network can be compared with a public «forum» for announcements and discussions. Messages are posted to one or more newsgroups, where they are archived for some days or weeks. Usenet is based on a non-hierarchical network of newsservers. If you send a message, it will be transmitted by your provider to your newsserver. The message then receives an ID number and is forwarded to the next newsserver, and so on. If you want to read Usenet messages, you ask your newsserver to send you the archived messages of the newsgroup you're interested in.

It is obvious that Internet is not just a *technical network*, linking computers and servers, but also a *social network*, linking people – or computers, as we can never be sure if the other party is human or if it is a robot, a so-called «bot».

Figure 3 - List of newsgroups (structure, hierarchy)

The Usenet network is non hierarchical in the sense that there is no central computer organizing the exchange of messages. However, there is a *topological hierarchy* in the organization of the different newsgroups: There are eight «big» categories (the «big eight», like rec.* or comp.*), as well as categories based on nations, like ch.*, de.* or fr.*. As sub-categories, there are newsgroups like ch.test, for technical testings. As I will explain later, there is an established democratic procedures to add a new group or to remove an existing newsgroup. [In this presentation, I will not discuss the special alt.*-hierarchy, which is more anarchic.]

Figure 4 - Example of NG-Messages (List) / message

The individual message sent to the Usenet is similar to an e-mail message. It consists of a «header» and of a «body». The header contains information about the date, the sender (From:), the receiving newsgroup(s), a subject line, a message ID, and other information.

In the case I am studying, messages should consist of text only, without attachments, while in other newsgroups, so-called «binaries» are exchanged as well, like pictures or sound files. The restrictions of this text-based, not formally moderated form of interaction makes it difficult to establish a binding social order and to sanction unwanted behavior, that is: to build and maintain a «group culture».

As the messages are archived, social monitoring is relatively easy - as long as participants use a consistent sender name (From:) - but sanctioning is very difficult. I will discuss these problems taking as an example the case of <ch.talk>, a newsgroup symbolically based in Switzerland.

3) Case study: Newsgroup <ch.talk>

The aim of ch.talk is formally fixed in a so-called *charter*. It is broadly defined as to discuss topics regarding Switzerland, e.g., discussions about Swiss politics. The charter is the result of a democratic procedure: In order to define the aim of a newsgroup or to propose a new group, every person can propose a formal «*Requests for Discussion*» (RfD). After a one-month period of discussion, the request may be revised and relaunched. Then, a «*Call for Votes*» (CfV) is carried out. Every Usenet user may give one vote about the proposal. The voting procedure has to be carried out by volunteers: someone has to collect the votes. Results are made *public* in the sense that the names of all voters are published, including their vote (yes or no).

Once a majority decided to introduce a new newsgroup, or to remove an existing newsgroup, the newsservers of the network have to be reconfigured. This is done by an encrypted «*control message*» someone has to send. So, like in the voting procedure itself, there are human beings involved, who have to be trusted. Finally, each administrator of a newsserver may decide individually, if he or she agrees to accept the <control message> that is: to add or to remove the newsgroup.

Although Switzerland has a tradition of direct democracy, this procedure is not typical for this nation only, but for the whole Usenet - it is like a very simple prototype of «e-voting». In a similar way, the *technical standards* of Usenet and Internet are not defined by some «general manager», nor by an international standardization agency (like ISO), but as a result of discussions among system operators and users, following the ideal of accepting the technically «one best way» (-- although this ideal is not always fulfilled in practice...).

The charter for Swiss newsgroups also defines some rules of conduct on what is allowed and what is not allowed in these newsgroups. For example, it is stated that «Binaries should not be posted to these groups», or «In general, cross-posting among ch.* newsgroups is strongly discouraged.» Speaking generally, the charter defines what *language* should or should not be used in the ch.*-newsgroups - that is: «social language» like german, french, english, as well as technical languages, like HTML, MIME, ISO-8859-1 or «Quoted-Printable».

figure 5 --> ch.-Charter / norms of conduct*

Other guidelines of conduct are formulated in a document called «*Frequently Asked Questions*» (FAQ) as well as in different «*Netiquettes*». Unlike the «official» charter, these documents have not been approved by a voting procedure. Therefore, they have less legitimacy. In a strict sense, they are nothing more than the result of the personal opinion of one user. In a broader sense, however, they claim to represent a «common sense» of social conduct on the Usenet, based on «tradition»: «This document describes the Usenet culture and customs that have developed over time. (...) All new users should read this document to acclimate themselves to Usenet.»

figure 6: Example of a Netiquette
(http://www.use-net.ch/netiquette_engl.html)

«Never forget that the person on the other side is human»
«Be brief», «Use descriptive titles»
«Avoid posting to multiple newsgroups»

The intention of these documents is to help new users («newbies») to find their way, or: to adapt their behavior to the «Usenet culture». -- we can regard this as a case of «configuring the user».

4) Mistakes and critics

Although there are some rules of conduct, this does not mean that they were not contested, nor that there were no mistakes committed. In the contrary: As the content analysis shows, there are a lot of mistakes and errors committed by users, and criticized by other users.

Often, mistakes are made by new users with badly configured software, or not knowing the rules of conduct, but sometimes they are committed by purpose, in order to provoke, to disturb or to send commercial e-mails (SPAM).

Some critics are focusing on the *content* of messages, e.g. messages posted to the wrong newsgroup, or the insulting style of a message -- that is: focussing on social standards of conduct. But in most cases, the critics are focusing on *formal* mistakes, ignoring technical standards of conduct.

figure 7: critics on content X critics on form

A lot of critics are about «*flooding*». This can refer to multiple posting of an identical message to one newsgroup (duplicates), or to a lot of different newsgroups (crossposting). Further, «flooding» also refers to messages containing «binaries», like pictures or sound files. And flooding can refer to «wrong quoting», like citing a whole previous message just to add «me too».

The main argument expressed against flooding is «technical» in the sense that these messages have a bad «signal-to-noise ratio» and that they are a «waste of bandwidth», that is: a waste of resources, i.e. of bytes.

Critics often act advocating for users who do not have the very last up-to-date technical equipment and who do not have a fast Internet connection. The main reason is to *maintain the accessibility* and the usability of the Usenet for every user.

5) Dimensions of sanctions

As there is no formal authority controlling the messages sent to the newsgroup, the participants have to organize themselves in maintaining an order. There are a lot of different possibilities sanctioning unwanted behavior, and I can give only a broad overview on different *dimensions of sanctions*: (figure 8)

- a) technical vs. communicative solutions
- b) public vs. private solutions
- c) internal vs. external solutions
- c) individual vs. collective solutions
- d) legitimate vs. illegitimate solutions

Most of the observed sanctions are *complaints*, sent in public to the newsgroup or by private e-mail to the author of the message. These appealing sanctions are both denouncing the authors as well as «educating» *all* participants, «setting the rules». However, as this strategy is only appealing, it does not by itself *impose* a sanction. There are only a few technical possibilities in order to sanction unwanted behavior. The most common is to ignore the sender by *filtering* his or her messages. On an individual level, this can be done easily, as most newsreader software offers this feature. But it can not easily be done on a collective level -- and it does not eradicate the problem -- the message still exists. There are solutions to this problem as well, but they are not considered as legitimate. The example I am discussing in the following presentation is about «*cancelling*» the message of another user. In order to allow every sender to correct a previous posting, the sender is allowed to send a «cancel message» which is -- like normal postings - distributed to all networked newsservers, marking the original message as «cancelled». It will not appear on the list anymore. One reason for this feature is to allow users to correct wrong configurations of their newsreader or newsservers, especially to delete duplicates. Cancelling messages is socially restricted to the sender of the original message (each sender can only delete his or her own posting), but it is not technically restricted: Using some tricks, it is possible to cancel the messages of other senders. However, this is considered as very unfriendly among Usenet users.

6) Sanctioning and role taking

Who is taking the role of sanctioning unwanted behavior in ch.talk? The observation shows that there is basically a «core» of active participants complaining about mistakes of other users. In doing so, they often present arguments and advices on how to do it better, linking to documents like the Charter, the FAQ or Netiquettes, as well as presenting exhaustive explanations for their claims of what is right and what is wrong. Of course, in doing this job, one can raise his or her prestige. On the other hand, these activists are exposing themselves to critique: They are insulted as «policemen», as «teachers» -- even as Faschists («blockwart»). It is crucial for them to *never ever* commit a mistake themselves, as they are constantly observed. It needs a lot to take the role of «educating» other users: Technical and social competence, integrity, a thick skin, resistance to insultings - and a lot of engagement, persistence, and endurance.

7) Automation: The CHancelbot

In July 1998, a user proposed to introduce a «bot» in order to technically support the task of maintaining a certain «group order». The technical agent would be capable to cancel «faulty» messages, i.e. to delete them from the newsgroup servers. Unlike other technical agents, this so-called <CHancelbot> is a very simple tool, basically a short computer script, with only little autonomy and no «personality» -- far away from any «artificial intelligence». The task of the <CHancelbot> would be to compare every new message with some defined criteria, using simple «if...then»-algorithms. If an error is detected, then a cancel command should be sent out to the other newsservers. Still, it was argued that it may be useful in helping to eliminate «duplicates» and «excessive crosspostings». However, as the bot would act for itself, it could make mistakes itself. After all, a cancelling bot is a «killerbot»!

On the other hand, there are some important advantages of such a cancelbot: First, it would be a relief for those participants taking the role of maintaining a group order -- not only in terms of time and work, but a *moral relief* as well, because the pressure on legitimizing the own action is externalized and moved over to a computer script. As a second advantage, it was argued that the bot would be taking decisions following strict rules -- acting in a democratic way in the sense that the same criteria is applied for all users, being less arbitrary and not based on emotions.

The discussion then focused on two questions: Should the CHancelbot only warn the author of a faulty message by an automated e-mail, or should it be allowed to effectively cancel messages?, and (b) What are the criteria for the definition of a faulty message?

To the first question: It was argued that appealing to the author would not effectively solve the problem. Further, it is often difficult to contact the authors, as they are using false sender addresses. However, as cancelling the message of someone else is considered as a very severe action, it needs a strong legitimation. In the discussion, legitimation is taken from technical standards, as defined in a RFC-1036 (e.g. a technically correct, verifiable sender address as a «Required Header line»), as well as from the official charter of the Swiss newsgroups (e.g., Binaries are not allowed). Soon, it became clear that the CHancelbot should in fact be able to cancel messages.

But what exactly should be cancelled? Every participant of the discussion agreed that the CHancelbot should look for «*duplicates*», that is: for messages posted two or more times to the same newsgroups, mostly because of a wrong configured newsreader or newsserver, or because users are impatient and do not consider the time lag between sending a message and having it visible on their newsserver.

There were only few discussions about cancelling messages containing «*binaries*». Although all participants agreed in not accepting binaries in ch.*-newsgroups, it is technically more difficult to exactly define what a «binary» is - e.g., if the volume of a message should be considered as criteria.

More excessive discussion were carried out regarding *crossposting*, i.e., sending an identical message to more than one newsgroup. In fact, crossposting is often an indicator for SPAM or UCE, unsolicited commercial e-mail. How many newsgroups

should be allowed? And what about cancelling messages sent to newsgroups outside the ch.*-hierarchy as well?

Other suggestions for criteria were messages containing HTML-code, messages with a false sender address, and messages containing certain patterns in the subject line, like «\$\$\$\$\$» or «Make Money Fast», in order to detect and to delete SPAM. These suggestions were heavily contested. It soon became a consensus that criteria should be strictly based on form, and not on the content of a message, even if it was SPAM. In sum, the CHancelbot was regarded as a useful tool to redeem the users from the «plague» of faulty messages. Participants now demanded to «*turn it on!*»

At this point, another user entered the discussion, strongly objecting. He argued that censoring was always a wrong solution, and he warned of George Orwell's «Big Brother», because the filtering software could be easily extended to delete messages containing certain keywords: «Today, messages are cancelled by its formal structure, tomorrow they will be censored by its content.»

In fact, the CHancelbot could be a very powerful censoring tool — and some of the participants decried the bot with aggressive words, making allusions to weapons and wars. At the same time, other participants underlined that the bot was not at all a «monster», but a impartial, neutral tool (eine «sachliche Maschine»), executing only the orders of what have been programmed by a human being. But this argument moved the critics to the question on why and how one should trust the operator. Who should be programming and controlling this bot? What «frame» should be established in order to restrict the power of the CHancelbot and of its operator, to «tame the monster»?

The first prototype of the CHancelbot was proposed by a technically competent user -- let's call him Luc. According to his own statements, Luc has been operating a similar «cancelbot» for some newsgroups of the french fr.*-hierarchy. For the Swiss newsgroups, Luc started a bot on his own server in a «*test mode*», without really cancelling messages. Every night, this bot automatically published a *report*, i.e. a list of messages fulfilling the defined criteria. (As we know from other innovations, «test modes» or «pilots» have become important procedures in introducing new technologies in general.)

The messages were stored during one week in a special «*repository*» on a public webpage. An overview showed soon, that the configuration considered mainly SPAM and other automatically generated commercial e-mail.

(Another important point concerning monitoring is the fact that the CHancelbot was programmed as «*open source*», so it was not a black box, but everybody could have a look at the code.)

Although some users asked Luc in public to «turn it on!», he did not agree, because: «As CHancelbot runs on my machine for now, and my machine is not willing to receive mass-complaints, I will only install rules which have been approved by a vote.»

It then still took some months to start the voting procedure. The first «Request for Discussion» consisted of two parts: One part was about the creation of a new group <ch.bulletin.chancelbot>, serving as a place to publish the activities of the bot (status reports), the other part was about the CHancelbot-charter. It ended up in being the longest charter of Swiss ch.*-newsgroups.

- In a preface, the charter restricted the activities to messages posted to at least one ch.*-newsgroup, and explicitly stated that «Articles must not be cancelled on the basis of opinions, views or facts they contain.»
- The CHancelbot would be allowed to cancel binaries, excessive crossposting and duplicates, following a precise technical definition.
- All activities are documented daily in a separate newsgroup <ch.bulletin.chancelbot>, and cancelled messages are kept on a public Web-repository during one week.
- The cancel control message contains an explanation why a certain posting was deleted.
- Further, the CHancelbot attempts to inform the sender of a message about the cancelling.
- Every cancelling command is marked, in order to allow each administrator of a newsserver to accept it or not.
- And finally, two human CHancelbot operators are to be elected by the readers of the ch.*-groups by at least a 2/3 yes majority.

It took quite a long time to establish the bot: Almost one year after the first public discussion, the CHancelbot was accepted with the great majority of 108 Yes to 7 No. (April, 4 1999). In Autumn, two «CHancelbot Operators» were elected.

8) Conclusion

To cancel the message of someone else is considered as a very severe action. Therefore, an automated cancelbot requires a strong legitimacy in order to be accepted. The introduction of the Swiss CHancelbot raises some important questions: Under what rules should anyone be allowed to *kill* messages of someone else? Should this task better be done by a machine or by human beings? What frame should be established to limitate the risks of delegating power to a machine and to an operator?

As we have seen, legitimacy for the Swiss CHancelbot is ideologically based in *technical rules and standards*. Further, legitimacy is taken from the presented *integrity of a person*, Luc, the first operator. And most important: Legitimacy is based in the formal *voting procedure*. The CHancelbot charter is not only defining the operational criteria of the bot itself, but six important principles:

- **locality**
«messages posted to at least one ch.*-newsgroup»

- **monitoring**
All activities are documented daily in a separate newsgroup, and cancelled messages are kept on a public Web-repository
- **argumentative rationality:**
The cancel control message contains an explanation.
- **information of the sender:**
CHancelbot attempts to inform the sender about the cancelling..
- **voluntaryness**
Every cancelling command is marked, in order to allow each administrator of a newsserver to accept it or not.
- **reversibility**
The rules as formulated in the charter may be changed by a common voting procedure.

It is important to note that these principles are not «technical», but social. Although the ideology of the participants is relying on technical arguments -- the CHancelbot is delegating cancelling power to a machine, strictly following technical rules, -- the frame established by the participants is mainly about social organization.

Technical principles served as a reference in the discussion, but social principles of self-organization were building the frame enabling the introduction of the CHancelbot.

1.