THE ARCHITECTURES OF TRUST

- Supporting Cooperation in the Computer-Supported Community

The University of Copenhagen
Faculty of Humanities
The Department of Film and Media Studies
Supervising Professor: Frederik Stjernfelt

February, 2002

THIS DOCUMENT MAY BE FOUND AT
WWW.MEDIAS-RES.DK/JONAS/THE_ARCHITECTURES_OF_TRUST.PDF

FOR MORE INFORMATION VISIT WWW.GAME-RESEARCH.COM

MA THESIS
JONAS HEIDE SMITH
SMITH@GAME-RESEARCH.COM

”There is no doubt whatever about the influence of architecture and structure upon human character and action. We shape our buildings, and afterwards our buildings shape us.”

- Winston Churchill

Cover illustration: Dark Age of Camelot
(Mythic Entertainment, 2001)

This version: 18-Apr-02

Foreword

We should, of course, have known.

When, in 1999, I had the pleasure of joining a small group of Copenhagen-based computer game designers, we realized that producing a large-scale online role-playing game would be no walk in the park. We were acutely aware that the technical challenges loomed to Himalayan heights and that the stiff competition would demand an alarming degree of aesthetic creativity on our side. We never thought it would be easy. Still, we had no idea.

At the time, a small number of similar games were neatly dividing up the rapidly growing world market. While this posed an obvious obstacle – in the dot-com parlance of the time we would not be *first movers* – it also gave us the opportunity to learn from their mistakes. Eagerly we dug into the lively online user communities that had sprung up around the largest games, looking for clues as to how to improve the user experience and avoid the pitfalls of unstable data transmission. We found very little of that sort.

Sure, players did complain about the dreaded "lag", involuntary freeze-ups caused by data fallouts and difficulties of synchronization. But such complaints were dwarfed by one explosive issue. The players seemed obsessed with a phenomenon that had very little direct bearing on technical and aesthetic design, the issue of *player killers*. Some players were killing other players, and the victims didn't like being dead.

In other words: Players of cutting edge online interactive role-playing games representing a revolution in the history of entertainment were concerned that other players weren't being *nice*.

The discovery hit us hard. We had storytellers, we had programmers; would we now need to recruit sociologists? Probably, although we never did. To be fair we were in over our heads in many other aspects and the project was abandoned. The experience, however, had taught us one thing. When people interact through computers there can be no separating technological, social, and aesthetic issues – one aspect irrevocably affects the other.

How this relationship may be formally described and thereby made usable is the main concern of this thesis.

One other event motivated my choice of subject matter. At the time I was considering which theoretical framework could most adequately capture the dynamics of online interaction, I witnessed a doctorial defence. The candidate

convincingly presented her research on online interaction and described her findings in microsociological terms. It seemed entirely reasonable. She was then asked if she had considered other theoretical approaches. She answered that she had, in fact, considered an explanation inspired by theories of *rational choice*. And why didn't she pursue this venue? It was, she said, far too cynical.

Cynical. A strange term to find in the scientific vocabulary. The exact meaning in this context is not entirely obvious to me, but if being cynical disqualifies a theory we will have to discard a growing body of research on computer-supported communities. We'd probably also need to part with natural selection and the majority of economic theories. We may want to think again.

If theories of social psychology are not compatible with more "rational" theories then at least one side is wrong. Thankfully they are, it seems, compatible. The terminology may differ, but human beings don't change on the way from one academic faculty to another. In this thesis I'll argue that researchers – as well as online community members – will be better off cooperating.

For valuable comments and discussion, I am very grateful to Thomas Wedell-Wedellsborg, Peter Lauritsen, Sine Just, Kasper Kofod, Simon Egenfeldt-Nielsen, and Søren Friis Smith.

List of contents

Introduction

How do people behave in online interaction, and is it possible to further cooperation and trust through software design?

That is the basic question that this thesis seeks to answer. Although the question of how human behaviour is shaped has ancient roots, it has only recently gained relevance within software design. The reason for this new-found importance may be boiled down to network technology. Simple cables have dramatically altered our conceptions of the computer. The story of this conceptual revolution places the following arguments in their proper context and will serve as our starting point.

Computers were not invented to facilitate human interaction. Quite the contrary, they were envisioned as automatons capable of replacing humans in arduous tasks or indeed as capable of providing a desirable alternative to faulty human logic. Leibniz, in the mid-seventeenth century, expressed the rationalist belief that a "universal logic machine" would be able to perform the necessary calculations to arrive at fair solutions to even highly complex social and philosophical issues, thus providing the foundation for a more just society (Mayer, 1999:4). More modestly, the 19th century inventor and polymath Charles Babbage (1791-1871) achieved funding from the British Government for his visionary "Difference Engine", a machine capable of producing accurate arithmetical tables of obvious interest to both industry and military. The machine would then be able to replace human number-crunchers known at the time as "computers" (Babbage, 1832:99).

The computer-as-replacement vision remained firmly in place until the mid-twentieth-century. British mathematician Alan Turing broke new ground with his systematic unveiling of the foundation for generalised logic and famously argued that computers might one day achieve human-like intelligence by sheer force of calculation (Turing, 1950). Turing's methodical vision of artificial intelligence (AI), however, falls squarely within the theoretical tradition of Leibniz and Babbage. An important break with this tradition was started in 1945 by American scientist Vannevar Bush (1890-1974). Reflecting on scientist contributions to the war effort, Bush claimed that a new challenge now lay ahead - a challenge of organisation:

> "…our methods of transmitting and reviewing the results of research are generations old and by now totally inadequate for their purpose… Those who conscientiously attempt to keep abreast of current thought, even in restricted fields, by close and continuous reading might well shy away from an examination to show how much of the previous month's efforts could be produced on call…truly significant attainments become lost in the mass of the inconsequential." (Bush, 1945:3).

What science needed, according to Bush, was not more research but rather an effective means of organisation and retrieval. In other words what was needed was a system capable of assisting, not replacing, the human effort. Bush's article is often seen as introducing augmentation research, the systematic focus on "*applications that extend people's capacities to create, think, and communicate*." (Mayer, 1999:11). More specifically, Bush's idea of an association machine, the Memex, looks remarkably like a theoretical forerunner of today's Internet.

Such thinking conceptually joins the computer to a human user. Whereas the machines of Babbage and Turing were built (and sometimes programmed) and then left to do their superhuman calculations, a device such as the Memex is more like a mental infrastructure. It does very little on its own. Hence the need to focus systematically on the interaction between user and machine, a field known as Human Computer Interaction (HCI).

HCI is a theoretical counterpoint to AI research. Although philosophical approaches to AI do take human psychology and perception into account, AI may be thought of as dealing primarily with the inner workings of machines. HCI, on the other hand, focuses on hardware-external processes, specifically on those of representation, communication, human psychology, and design. In MIT psychologist J.C.R. Licklider's original vision of the 'symbiosis' of man and machine (see also Hafner & Lyon, 1996:27-39) he stated the hope that

> "...human brains and computing machines will be coupled together very tightly, and that the resulting partnership will think as no human brain has ever thought and process data in a way not approachable by the information-handling machines we know today." (Licklider, 1960/1999:60).

Slightly more modest statements tend to characterise modern day HCI, but the focus on enhancing human capabilities by thoughtful interaction design remains central.

To computer scientists of the 1960s it may have appeared that a wide range of professions was taking a keen interest in the machines. What had until recently been a highly technical endeavour was now being transformed into an interdisciplinary field. However, another conceptual challenge was under way and found expression in a 1968 article by Licklider and Robert Taylor (Licklider & Taylor, 1968/1999). The article began immodestly: "*In a few years, men will be able to communicate more effectively through a machine than face to face*." (Licklider & Taylor, 1968/1999:97). The emphasis was no longer on the inner workings of machines, nor the interaction of machine and user but rather on the interaction between people *through* computers. This marks the start of research into Computer Mediated Communication (CMC), which is the research tradition most clearly associated with this thesis. The history and present state of CMC will be addressed in the following

chapter. At this point, it is sufficient to note that within the modern conceptual history of the computer, the technical sciences (AI) were supplemented by the humanists (HCI) who were soon followed by the social scientists (CMC). From an academic viewpoint, then, the computer exists in a interdisciplinary field and while this has obvious advantages, it also means that there is some disagreement as to how the computer is best described. This thesis presents only one approach to CMC and on many important aspects of the discipline it has very little to say. On the subjects of AI and HCI, of course, it makes almost no comments at all. But an effort is made to point to bridges between academic categories.

Focus and demarcation

As has been implied, this thesis centers on the necessary design conditions for computer-supported cooperation. Social issues pertaining to online interaction are analysed on the basis of existing sociological theory with the specific aim of determining if there are analytically important differences between interaction in offline and online settings. This leads to a description of how knowledge of online dynamics may be used to further cooperation and trust in collaborative computing.

This, however, is no manual for software designers. It is not a list of *do's* and *don'ts* for the aspiring online community developer. Still, much effort will be put into phrasing the analysis in as accessible and concrete terms as possible. It is my goal to enhance our theoretical understanding of online dynamics, but if the analysis can be put to actual use as well, so much the better.

Theoretically, I draw upon sociology and most specifically the field of "social game theory". In this vein, I also rely upon sociologically inspired elements of media studies, such as "medium theory" and on both interpersonal and mass communication theory. From computer-specific research, the analysis draws inspiration from Computer Supported Cooperative Work (CSCW), general CMC, and studies of "adaptive structuration", that is, the relation between software design and software use (Preece, 2000:172).

Compared to what is perhaps the better part of research into "virtual communities" this thesis touches only sporadically upon issues from social psychology and only in a critical context in **Chapter three** and in the final chapter will I touch slightly upon the much-discussed wider implications of computer-supported interaction.

Problem statement

What is cooperation and how does it come about? Is online interaction different enough from traditional interaction to warrant a distinctive description of cooperation?

How will answering these questions enable us to understand and affect the dynamics of interaction in computer-supported social settings?

Structure of the argument

Chapter one concerns the characteristics and historical development of computer-supported communities. The goal is an adequate description of the phenomenon.

Chapter two describes the logic of cooperation from a sociological perspective. Classical theories of the state are introduced and then described in the language of game theory. The goal is a description of cooperation and of the preconditions for stable communities.

Chapter three provides a theoretical framework for computer-supported communities. It explains empirical results from computer-supported communities in the framework of sociological game theory as described in chapter two and goes on to describe the intense theoretical debate that CMC has provoked. Research from CSCW is used to pinpoint important differences between online and offline interaction. The goal is to determine if and how prerequisites for social stability in online interaction differ from those known from offline communities.

Chapter four describes findings from a survey establishing actual user preferences. 300 online gamers answered questions of trust, and evaluated proposed design principles. Respondents displayed a clear predilection for communication and expressed attitudes toward trust that are noteworthy when analysed in the light of previous chapters. The survey serves to test whether real CMC users would find the design principles distilled from previous chapters appealing.

Chapter five sums up the argument. This leads to suggestions for further research and to a brief discussion of the claims that have been made as to the general implications of CMC.

Terminology

Pronounced interdisciplinarity coupled with novelty and huge financial interests has turned cyberstudies into a terminological minefield. Rather acute disagreement surrounds even the most fundamental terms such as *virtual,* and *community,* not to mention the constant problem of CMC dictionaries: *interactivity*. Other troublesome words include *cyberspace*, *information superhighway, information society*, and *knowledge sharing*. Fortunately, clarity is enhanced by the emergence of textbooks,

influential conferences, and reviews. Until something resembling a general consensus is reached, however, it seems prudent to define the terms rather carefully. I use the following specialised terms:

> **CMC**: Refers to actual computer-mediated communication as well as to the study of this phenomenon. The meaning should in each case be clear from the text.
>
> **Community**: Refers to a group of people characterised by stable cooperation and frequent interaction. The term is used liberally to include 'gatherings of convenience' that would fall outside visions of *gemeinschaft* and natural local harmony. (Will be further described in **The issue of community**, page 56).
>
> **Game theory**: Used as a label for sociologically oriented theories about the success rates of various strategies in game-like settings and the theoretical implications hereof. Though traditionally an economic discipline, it is used here in its relevance for individual and collective action more generally. (Will be further described in **Chapter two**).
>
> **Interactivity**: Used in a broad sense as a measure of user control over form and contents of a medium product.
>
> **Sociability**: Used here as a measure of how well a software product supports (favourable) social interaction (following Preece, 2000).
>
> **Usability**: A measure of the user-friendliness of a software product. The discipline of usability testing may be seen as a practical application of HCI.
>
> **Virtuality:** The term (as well as the adjective 'virtual') is generally avoided as it makes unwarranted ontological assumptions. When used it should be understood as synonymous to *computer-supported* or *online*. (Will be further described in **Beyond alienation**, page 56).

Since CMC is not a unified field and many details may not be generally known, I make extensive use of explicatory footnotes and illustrations where appropriate.

Theoretical assumptions

Metatheory, I believe, is mainly interesting in action. Siding explicitly with an '–ism' is less valuable than clarifying each step taken in an argument, making sure that the relationship between premises, evidence and conclusion is as obvious as possible. In other words: The following argument may be wrong, but if so, I hope it to be obviously wrong (to those who know better).

Nevertheless, I do not wish to be seen as representing the (mistaken) idea that there can be science, philosophy, or communication without *a priori* assumptions. The '–ism' that may be distilled from this thesis, then, is a modest reductionism. It is not the assumption that everything can be explained at lower levels but it is the idea that irreducibility is a very unfortunate *axiom*. If something can be explained by fewer,

simpler premises it would be a shame not to – if nothing else such parsimony would spare us great deals of confusion. This leads to an important distinction, since the term 'explain' is perhaps too broad. A low-level explanation of a given phenomenon can yield insights even if it does not account for every aspect of the phenomenon. Take human consciousness. This most complex phenomenon is a result of less complex processes. Whether natural selection alone is accountable, or we need various sorts of mindless help (random drift, non-random mutation, embryological constraints etc.) is beside the point. Understanding whatever simple, mindless processes has led to the development of consciousness (disregarding only magic and divine intervention) would be tremendously important to artificial intelligence and, of course, any theory of mind. But it would be important without explaining, in details, how human consciousness works, or how it feels to be conscious. Similarly, an analysis of cooperation and trust based on simple models can be important without accounting for the complex cognitive details of how actual people evaluate the trustworthiness of each other.

Having said this, I have already hinted at an underlying ideal of this thesis. I believe that interdisciplinarity is not just important, but in fact almost essential for two reasons. Firstly, at least on the subject of human beings, there are obvious advantages in drawing upon a variety of fields. Surely, perception and interpretation can be studied as if the human mind was some sort of black box (just like the heavens could be studied without knowing the nature of stars) but such studies would be strengthened by an understanding of why the results emerged. Secondly, the alternative to cooperation may very well be conflict. Like it or not, scientific fields overlap. Humanists are making claims about the natural sciences (or about their foundations) and natural scientists are making claims about very concrete aspects of human behaviour. Wasting time and effort on fighting this development is rather futile and occurs to the detriment of all combatants.

Avoiding interdisciplinarity, of course, can also be defended on more sober grounds. It is not without danger to venture outside one's field of expertise (as has been aptly documented by Sokal & Bricmont, 1998). But this is partially a question of mindset. A sober – and above all: modest – approach to adjacent fields is different from an arrogant excursion into foreign territory.

The architectures of trust

The idea that human action is a function of structure, systems, or architecture would be a controversial one in some sub-domains of sociology (Ritzer, 526pp). Nevertheless, such a notion lies behind most legislation and is certainly central to

any theory that would justify the existence of a state. The issue may be split up into several segments, taking literately the metaphor of architecture. Imagine a hypothetical building equipped, among other things, with a staircase and a kitchen. Is the individual forced to use the staircase? Surely not, he or she can jump out the windows or climb down the wall. Is he forced to cook his food in the kitchen? No, but it would be practical since that's where the oven is. However, over time and all things being equal, a majority of people are likely to take the stairs, sleep in the bedroom, and cook in the kitchen.

All things, of course, are not always equal. Disaster may strike in the kitchen forcing people to find alternative places to cook. But if unexpected things happen at random then, over time, things are equal.

Similarly, the relation between architecture and cooperation that is proposed here is not one of absolute regularity – it is not a deterministic theory. It cannot (necessarily) be applied to any existing computer-supported community to explain its present condition. Specific conditions may override the 'pull' of the system architecture. But retrospectively the theory may contribute to an explanation of the fate of the community and if we made an analysis of all such communities it would be able to statistically predict the odds that a new system would be visited by social strife (by whatever measure we set).

I should stress that this approach may come close to what Nancy Baym calls "*the mistake to view patterns in CMC as direct effects of the medium*." (Baym, 1996:149). However, although Baym makes her point well, I believe her conclusion is partly rhetorical. Through her analysis, she merely defines anything, which looks like determinism or reductionism as false by default (or she makes the rather obvious point that complete context-independent determinism is wrong, in which case one can only agree). Those who would argue that technologies or media do not have effects are either stating the obvious or facing the challenge of refuting works within medium theory such as Meyrowitz (1985) (see also McQuail, 1996:107-118).

This, then, is how 'architecture' is used in the framework of thesis. It is the structure that at any given time influences action, and which over time, given certain minimal assumptions, can predict the fate of a system (at least in terms of probability). I assume that the actions that we wish to induce are those that require trust[1]. Trust itself comes with many definitions, but one of the simplest will suffice: "*Trust is*

[1] Certainly, this need not always be so. There may be times when we wish to discourage cooperation: When criminals interact, or in market situations where cooperation between companies would equal cartels. Knowledge of important design principles remains important, of course, even if we wish to explicitly avoid them.

anticipated cooperation" (Burt & Knez, 1996:70). Cooperation itself is the topic of chapter two.

While, to some extent, she subscribes to what I will later describe as the Myth of Liberation, the clear vision of how architecture shapes online behaviour was first expressed by Judith Donath (Donath, 1996). She says it well:

> "…the future success of virtual communities depends on how well the tools for social interaction are designed. If they are poorly designed, the on-line world may feel like a vast concrete corporate plaza, with a few sterile benches… If the tools are well designed, the on-line world will not only be inhabited, but will be able to support a wide range of interactions and relationships, from close collaboration to casual people watching." (Donath, 1996:2).

CHAPTER ONE

A brief history of computer-supported interaction

To a large extent, computers have changed from being tools for calculation to being media for communication. This conceptual landslide has been outlined in the introduction and will in the following be supplemented by an account of actual historical experiences with people using computers for communication purposes. I divide this account into categories based upon conventional classification even if the fundamental technologies of two different phenomena may in fact be closely related. For reasons of volume, technical definitions, although important, are not addressed. To compensate, references to background literature are given wherever possible.
The aim of the present chapter is to be descriptive. Analytical and theoretical comments, although not entirely avoided, are kept to a minimum.

Experiences with online interaction

In the late 1970s, it became clear that something was very wrong. Wrong, at least, according to the original vision of the ARPANET designers. What they had diligently and explicitly built was a government-funded resource sharing network capable of connecting academic knowledge centres while being able to withstand severe damage due to its decentralised nature. Expensive mainframe time could be shared and crucial communication channels would be safe from even massive nuclear attack[2]. No one disputed the importance of these goals but nevertheless an entirely different purpose was obviously also being served. The network was awash with email. In the words of Internet chroniclers Katie Hafner and Matthew Lyon:

> "As cultural artifact, electronic mail belongs in a category somewhere between found art and lucky accidents. The ARPANET's creators didn't have a grand vision for the invention of an earth-circling message-handling system. But once the first couple of dozen nodes were installed, early users turned the system of linked computers into a personal as well as a professional communications tool." (Hafner & Lyon, 1995: 189).

On the brink of a decade that would see widespread proliferation of network access, this should not have come as a surprise. Plenty of unmistakable signals as to user preferences had been given. Within a few years of the 1969 founding of ARPANET, an electronic mailing list, run by and for science fiction fans within the programmer community, had sprung up (Pargman, 2000:28). Both form and content of this list ran contrary to the stated purpose of the network (see also Baym, 1996:146).

[2] The importance of this feature is disputed (Pargman, 2000:28; Hafner & Lyon, 1996:10). Part of this dispute may perhaps be attributed to social issues concerning working expressly for specifically military purposes in the American academic climate of the 1970s.

At the same time, electronic bulletin board systems (BBSs), were put to eager use among computer enthusiasts. Such systems would, as the term implies, provide a forum for sharing information, this information being automatically archived and structured in the process. Notably San Francisco programmer John James spearheaded the idea that such forums were capable of transforming social dynamics by careful consideration of design and suggestive use of metaphor (Stone, 1992:88). Inspired by this vision he named his new-built system *CommuniTree*.

The case of CommuniTree

"*We are as gods and might as well get good at it*." read the introduction to *CommuniTree #1* in 1978 in what A. R. Stone has labelled the "*technospiritual bumptiousness, full of the promise of the redemptive power of technology mixed with the easy, catch-all Eastern mysticism popular in upscale northern California*". (Stone, 1992:90).

True to its name, *CommuniTree* introduced the principles of 'branches' or 'threads' (see **Figure 1**), which proved a marked improvement upon earlier forms of BBSs that would only sort by posting time or provide crude search facilities.

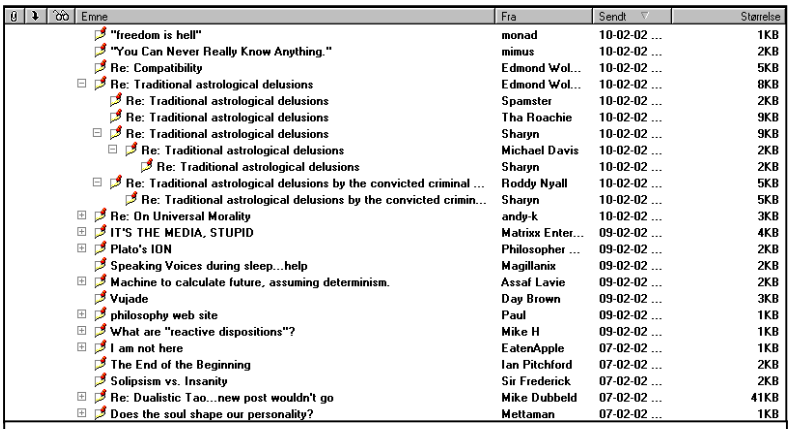

Figure 1: Present day Usenet newsgroup showing a threaded discussion.

The Tree flourished and its users, most of them nourishing optimistic notions about the democratic potential inherent in the system, enjoyed the opportunities for free and uncensored dialogue (see Reid, 1999:107). What has guaranteed *CommuniTree*'s fame, however, is not so much its virtues as its discouraging downfall. In 1982, four years after the system had opened its online gates, Apple Computer entered into an agreement with the American education system, hoping to gain a stable foothold in this lucrative market section. This led to a rather sudden increase in student access to personal computers and to network technology, which in turn led to a whole new user segment logging on to the always-open *CommuniTree* system.

True to its ideological underpinnings, no one could be denied access to the forum and there was no way of filtering incoming messages. Such faith in user responsibility may have inspired long-time guests, but many newcomers were not impressed. Rather they set forth *"with the linguistic proclivities of pubescent males*" (Stone, 1992:91) to turn *CommuniTree* into a private playground, thus flooding the system with obscene and – to original users – irrelevant posts. For three months, administrators attempted to counter the problem without compromising their principles. Then *CommuniTree* was choked to death.

BBSs as such, however, continued to be popular, albeit often in less tolerant forms, and remain widespread on the Internet, intranets and in particular on Usenet[3]. While categories are obviously not unambiguous, it has been estimated that the vast majority of messages on Usenet are on social or leisure subjects: "*In short, Usenet is used, above all, for social interaction on topics of personal rather than professional interest*." (Baym, 1996:147).

Electronic worlds; MUDs and their users

Non-professional human communication was not the only unexpected guest to call on computers and networks. Another phenomenon, way outside the bounds of officially desired use, was games. Programmer Steve Russel and a science fiction loving group of Harvard students had in 1962 introduced computer games to the world with the multiplayer action game known as *Spacewar* (Graetz, 1981). Ten years later, engineer Will Crowther, although busy ensuring the foundations of ARPANET, began work on a small program, which would simulate his favourite pastime, cave exploration, thus preferably keeping his children entertained. After undergoing major improvements, the game was widely circulated on the ARPANET under the name *Adventure* (also known as *Colossal Cave* and *ADVENT*) (Adams, 2000; Aarseth, 1997; Smith, 2000). In its (more or less) final form *Adventure* drew obvious inspiration from tabletop role-playing games popular at the time, especially TSR's *Dungeons & Dragons*. One thing was, however, very different. Crowther's game was decidedly single-player. Dedicated role-players missed the sense of joint exploration and playful interaction common to games like *Dungeons & Dragons*.

Roy Trubshaw and Richard Bartle of Essex University sought to address this problem by constructing what may be described as a multiplayer online version of

[3] Usenet is a network of servers providing access to user-submitted postings arranged in a topical threaded fashion. A topical discussion is known as s newsgroup (see **Figure 1**). Several hundred thousands of messages are posted to newsgroups each day (Smith, 1999:197). For sociological perspectives, see Gotved, 1999; Baym, 1996; Wellman, 1997. For a clear technical introduction see Smith, 1999.

Adventure (Pargman, 2000:30; Bartle, 1990). They named their creation *MUD* for 'Multi-User Dungeon'. *MUD* (usually referred to as *MUD1*) was a database consisting of objects and relations between objects. The system would allow users at distant terminals to log on and enter an online world in which they could interact with objects and other players[4].

For a wide range of reasons, MUDs have enjoyed considerable academic attention and are among the most well-described phenomena in CMC[5] (e.g. Turkle, 1995; Curtis, 1992: Pargman, 2000; Reid, 1999; Dibbel, 1999; Ito, 1996; Bartle, 1990).

Later MUDs or subtypes of MUDs such as MOOs (MUD Object Oriented) have downplayed or eliminated the explicit gaming focus by doing away with role-playing mainstays such as experience points and aggressive monsters. In 1988, for instance, Jim Aspnes of Carnegie Mellon University created the first "social MUD" called *TinyMud*. Other MUDs have served collaborative purposes for researchers and some have even been designed as therapeutic settings for overcoming traumas from sexual assault (Reid, 1999).

Most recently, MUDs have spawned a group of highly commercial offspring. Whereas MUDs have traditionally relied on purely textual interaction, these new so-called Massively Multiplayer Online Role-Playing Games (MMORPGs) such as *Ultima Online*, *EverQuest* and *Dark Age of Camelot* place high demands on user hardware. The basic principles are nevertheless fully comparable to those of *MUD1*.

Life in MUDs

Taken one by one, the history of most MUDs bear certain similarities to the history of *CommuniTree*. While genre, player demographics, and designer creativity have surely shaped the development of individual MUDs, noteworthy correlations between system architecture, social dynamics and subsequent development do stand out.

The following examples are not representative of all MUDs but serve to show how social life has taken shape in certain well-documented specimen. A theoretical explanation will be given in chapter three.

Perhaps the most well-known event in the history of MUDs was an act of rape. *LambdaMOO*, a social MUD hosted at Xerox PARC for research purposes, was a relatively peaceful place until the player behind a character called Dr. Bungle

[4] For a formal description of MUD characteristics see Curtis, 1992.

[5] Some researchers, however, have felt the desire to justify their interest in this game-like phenomenon leading to a "*bonanza of naming practices in the literature*" (Pargman, 2000:31). Most notably the "D" in MUD is sometimes taken to stand for 'domain' or 'dimension'.

managed to take control of several other characters. He used this power to commit (or to simulate, depending on one's ontological assumptions) a prolonged and highly public rape (Dibbel, 1999). This caused an outrage among the users, some of whom demanded that Bungle be immediately expelled. One popular opinion was that Bungle should be treated much as a real-life rapist would be. Obviously, however, he had not committed a crime that any real-life courtroom would take very seriously. Consequently, it was argued that he should receive in-game punishment (that is the *character*, not the *player*, should be punished) by being exiled or having his capabilities stripped away. But this was not altogether easy, since *LambdaMOO* had no laws and indeed no legal system as such. What the system did have was administrators known as wizards who had access to the very structure of the system. After a short period of time one of the wizards made up his mind and eliminated the Bungle character. The upset users were, however, not satisfied. Since *LambdaMOO* allowed users to build and create both characters and objects, many felt they had invested much time and effort into the system – that it was partly *their* system. The display of arbitrary power by the wizard made users worry for the safety of their characters in a system that wielded judgement without fair trials. The administration and the users then went on to establish a highly complex political system.

Much the same happened in the educational MUD *MicroMUSE* (Smith, 1999). Since the system was to be used by young students, it was generally understood that decidedly adult themes should be avoided. A character named Swagger decided against this consensus and built an 'orgasm room' into which he invited selected other users. For this he was punished arbitrarily causing an intense conflict between users and administration leading eventually to altered power structures (Smith, 1999:140).

Another large-scale virtual world to be visited by unexpected social strife was *Habitat*, a commercial graphical MUD developed by LucasFilm in the mid-1980s. Having optimistically designed the system architecture to enable 20.000 simultaneous user connections, the developers were humbled by early experience: "*By the time the population of our online community had reached around 50 we were in over our heads (and these 50 were 'insiders' who were prepared to be tolerant of holes and rough edges)*." (Morningstar & Farmer, 1990:9).

Some issues could be addressed by fixing bugs and streamlining data transmission procedures, whereas others took the form of genuine social dilemmas without obvious solutions and with clear philosophical implications. In particular, the issue of proper behaviour fuelled an intense debate between those who would attempt to eliminate all physical conflict at the system level and those who saw the opportunity

for virtual death and thievery as an integral part of *Habitat*'s appeal. Compromises were sought and methods of conflict resolution were implemented. These lead to town meetings through which officials, most notably the sheriff of Populopolis, were elected. Reflecting on this, the developers have noted:

> "For weeks the Sheriff was nothing but a figure head, though he was a respected figure and commanded a certain amount of moral authority. We were stumped about what powers to give him. Should he have the right to shoot anyone anywhere? Give him a more powerful gun? A magic wand to zap people off to jail? What about courts? Laws? Lawyers? … It was clear, however, that there are two basic camps [of players]: anarchy and government." (Morningstar & Farmer, 1990:11).

This division appears to affect most MUDs. Surveying social order in MUDs, Elisabeth Reid found that "*For some users, the possibility of 'playerkilling' adds depth and spice to the virtual world*" while for others "*playerkilling destroys the mental illusion in which they wish to immerse themselves*…" (Reid, 1999:123)[6].

Creative compromises have been sought to satisfy both camps. In *Habitat,* certain areas were designated as "safe", while others (the wilderness outside the city areas) held the risk, or possibility, of interplayer fighting. *Ultima Online* (Origin, 1997), the first successful large-scale MMORPG, employed a similar system, although emphasising verisimilitude by creating a narrative explanation for the peaceful urban areas: Anyone drawing weapons inside city limits would be attacked by the powerful police force. Nevertheless, the problem remains as is obvious from the website of the MMORPG *Dark Age of Camelot* (Mythic Entertainment, 2001, see **Figure 2**):

> "An unfortunate situation has arisen in several currently-available online games where some game players go out of their way to ruin the gaming experience for other players by killing them repeatedly, 'stealing' their monster kills, and generally making an nuisance of themselves. Camelot has several built-in methods for discouraging this behavior." (http://www.darkageofcamelot.com/faq/).

[6] Even in explicitly non-gaming settings trouble lurks. "*A single user of JennyMUSH [a sexual assault relief MUD] was able to subvert the delicate social balance of the system by using both technical and social means to enact anonymously what amounted to virtual rape*". (Reid, 1999:115).

Most notably, *Dark Age of Camelot* focuses on quests that require cooperative play and does not outlaw interplayer combat but rather limits it to fighting against inhabitants of other, rival countries.

Figure 2 - Dark Age of Camelot (Mythic Entertainment, 2001)
In this MMORPG, interplayer conflict is reduced by focusing on common "external" enemies.

The wired workplace and CSCW

Although flexible and varied, both Usenet and MUDs may be loosely grouped with other recreational forms of CMC such as Internet Relay Chat (IRC). More true to original visions of CMC, network technology such as Ethernet (Hafner & Lyon:250-251) in the 1980s inspired a focus on the potentials for enhancing flexible workstations by adding powerful communication capabilities[7]. The ambition was to create a virtual workspace for collaboration while moving beyond limitations of stand-alone applications such as email. This goal had led to the development of *groupware* (Baecker, 1995:741). Such systems often include a variety of well-known functions such as BBSs, chat, poll features, and calendars but improve on these by rendering data usable for a variety of purposes. For instance, a document may be "known" by the system to be in a certain stage of a project, which will determine editing privileges, notify users of deadlines, and enable the system to monitor

[7] More specifically Grudin, 1995 mentions that four conditions for interest in groupware were met: A) computation inexpensive enough to be available to all members of some groups; B) a technological infrastructure supporting communication and coordination, notably networks and associated software, C) a widening familiarity with computers, yielding groups willing to try the software; D) maturing single-user application domains that pushed developers to seek new ways to enhance and differentiate products.

whether appropriate resources are assigned to the project. The different programs used for such purposes are legion but among the most widespread are *Lotus Notes* (Shneiderman, 1998:14) and various Microsoft application packages. With increased Internet access, websites such as www.yahoogroups.com, www.groupcare.com and www.jay.net have made collaborative computing publicly and often freely available.

Groupware designers may have been inspired by Licklider's "*think as no human brain has ever thought*..." to assume that by eliminating obstacles for orderly, structured communication they were paving the way for a substantial efficiency increase. Now, compared to systems with lesser functionality, groupware obviously does offer advantages but in the mid-1980s there was no denying the obvious: Results were falling painfully short of expectations (Grudin, 1995:764; Baecker, 1995:741). Apparently, the mere bundling of existing software was not enough to ensure rapid synergy effects. This led to "*the recognition that it would be useful to bring together people from diverse disciplines – technical, social, economic – who shared an interest in issues of communication and coordination*." (Baecker, 1995:741). This was done under the heading Computer-Supported Cooperative Work (CSCW).

One influential CSCW case study highlights organisational hindrances to the successful implementation of *Lotus Notes* in a large corporation (Orlikowski, 1996). Through a semi-ethnographic approach, the researcher found a huge disparity between management expectations and employee perceptions. Considering the newly purchased software a boon to knowledge sharing in a multinational workplace, management expected an efficiency increase. The employees, however, did not know what to make of the new software (was it email? was it a word-processor?) and when they found out saw no obvious incentive to share their knowledge:

> "The competitive individualism – which reinforces individual effort and ability, and does not support cooperation or sharing of expertise – is countercultural to the underlying premise of groupware technologies such as Notes." (Orlikowski, 1996:184).

The conclusions drawn from this study are that user expectations of new systems can impact dramatically on actual use (see also Baym, 1996:158) and that collaborative computing is vulnerable to traditional problems of cooperation (which are the subject of **Chapter two**).

CSCW has also shown that even if groupware is integrated into an organisation, disappointments abound. If employed to facilitate teleworking or otherwise eliminate the need for physical co-presence, one will need to capture all relevant information within the system. In other words, one will have to represent formally the tacit knowledge inherent in most work processes while employing a precise (or at least

not wrong) model of human learning processes and the nature of communication. This has accentuated the gap between information, which may be handled by computers, and embodied knowledge (Brown & Duguid, 2000; Baecker, 1995:741; for a theoretical framework see Dreyfus & Dreyfus, 1986).

World Wide Web; from hypertext to social space

The Internet, as mentioned above, was conceived for practical professional purposes. The various applications that were continually added to the network appealed to wide-ranging user-groups and with increasingly affordable access, the number of users grew rapidly throughout the 1980s. The most influential addition to the network, however, occurred in the early 1990s. Physicist Tim Berners-Lee, reacting to the Babel of programming languages that were used in research communities, envisioned an invention that would facilitate access to information while overcoming the costly barriers of system incompatibilities. For this purpose, he constructed and advocated the basic structure of the World Wide Web (Berners-Lee, 2000). By rendering possible the low-cost sharing of information enhanced by free-sprawling association in the form of hyperlinks and a certain degree of self-organisation, World Wide Web (WWW) soon proved to be a major force in promoting private Internet access.

Challenging the boundaries of what was originally a means of representing data, the basic facilities of hypertext were soon supplemented by advanced database features and measures that would enhance options for user-website interactivity. This paved the way for a focus on e-commerce and forums that would allow user-user interaction.

One driving force behind such forums was simple economics. In an advertisement-driven economy, the basic currency is attention, measured on WWW in the form of user visits or ‘sessions’. Providing present and future users with motivations for such visits means the large-scale production of valuable information or ‘content’ (Nielsen, 2000:99). Seeing the chance to outsource this time-consuming task to users, many commercial websites have erected sophisticated user sections often under the rather hopeful heading of ‘community’. While inviting free content such sections are also seen as adding ‘stickiness’ to a website by inviting an investment by the user which will ideally prevent him or her from abandoning the website:

> “…online communities are spawning by the dozen on dot-com sites. Owners of these sites believe that online communities serve the same function as the sweet smell of baking cakes does in a pastry shop. Both evoke images of comfort, warmth, happiness and maybe even trust.” (Preece, 2000:17).

Some websites blatantly avoid the prohibitive costs of content production by offering nothing more than frameworks for interaction. Obvious specimen include online groupware (see page 22) but webmail services such as www.hotmail.com, dating services such as www.dating.com, and online marketplaces such as www.ebay.com employ similar business strategies.

The short history of WWW, then, mimics that of the Internet as such although with a clear commercial spin. While being designed to alleviate a concrete need, its flexibility has opened the gates for a variety of different activities.

Casual conversations; IRC and instant messaging

From its undramatic, and somewhat poorly documented, origins in a Finnish computer science lab in 1988 (Hamman, 1997) Internet Relay Chat (IRC) soon became a much-beloved topic in CMC research. IRC brings synchronous communication to computers by allowing users to type text, which is instantly displayed on the screens of other participants. As with traditional email, IRC uses a client-server setup in which users at terminals communicate through a widespread network of servers. Conversations are divided into 'channels' comparable to Usenet newsgroups (for technical details see Pioch, 1997).

Important to the initial proliferation of IRC was the Gulf War, which sparked the need for fast, informal, and uncensored communication between geographically separated parties ('historical' IRC communication can be found at www.2meta.com/chats/university). Having gained momentum, IRC use grew rapidly, resulting in 15000 visits to servers in October 1995 - a figure, which was doubled by April 1997 (Hamman, 1997). As is the case with MUDs, however, what has attracted academic attention is not the popularity of the phenomenon (neither are main-stream phenomena). Indeed the percentage of Internet users who have ever logged into an IRC channel is not likely to be very large. Rather, social scientists have been attracted by the obvious implications for self-representation and observations of highly intricate communicative conventions that reach far beyond the use of emoticons such as the 'smiley': :-) (e.g. Reid, 1991; Bechar-Israeli, 1995; Paulillo, 1999).

Later developments would indicate that the relatively limited popularity of IRC might be attributed to usability issues. Although itself somewhat technical, the program ICQ developed by Israeli Mirabilis Ltd. in 1996 very rapidly proved highly successful, heralding the start of so-called instant messaging. With very modest traditional marketing efforts the service quickly reached one million users and by February 1998 had 7,5 million registered users[8] (Einstein, 1998, see also Dutton, 2000). At heart this type of software combines IRC functionality with 'buddy lists', providing real-time information on which friends, acquaintances or colleagues are online and available for chat (see **Figure 3**).

Figure 3: Buddy list window of Trillian, an instant messenger. Icons indicate whether or not friends are online.

The need for hierarchically maintained 'channels' is thus eliminated by user-maintained 'personal channels'. Features such as file-transfer, SMS and mobile phone integration, and online gaming functionality are constantly added in the ongoing competition for user desktops. Between them America Online (who now own ICQ), Microsoft and Yahoo are thought to have had more than 68 million different users on their services in the month of July 2001 (Ostrom, 2001)[9].

Instant messaging thus seems to have improved significantly upon IRC by enhancing accessibility and by increasing social presence (Preece, 2000:150) compared to more diffuse online discussion settings.

Summary and theoretical perspectives

The ARPANET, a distributed packet-switching network designed to serve both military and academic interests, was almost immediately exploited for purposes of non-professional communication. Catering specifically to such needs, electronic bulleting board systems were constructed, often with the intention of providing forums for uncensored dialogue on a variety of issues ranging from the political to the sexual. Early experience, however, showed that boundless freedom as a hardcoded principle in system architecture might jeopardise the robustness of the

[8] Instant messengers thrive from the fact that the value for a specific user depends on how many friends and acquaintances use the very same software. If a user wishes to communicate through the service he or she will often have to promote it first.

[9] These figures should be regarded with healthy scepticism (or rather: It is not obvious what they signify). Both America Online and Microsoft automatically provide instant messenger accounts to users of other services, thus artificially boosting numbers.

forum. Much the same has been made painfully clear to MUD designers who often appear genuinely surprised when their virtual worlds are visited by social strife caused by diversity of player ambitions and problems of cooperation. Such problems also contradict the gospels of knowledge sharing, preached by proponents of groupware: "*Look in much of the management literature of the late 1990s and you could easily believe that faltering business plans need only embrace knowledge to be saved*" (Brown & Duguid, 2000:118).

Groupware design lessons echo experience from AI research that by digitally reproducing human communication and perception we are struck by the inadequacy of our intuition. Superficially, human language seems deceptively simple (even children quickly learn) and at the surface level work, learning, and social interaction seem separate issues. Nevertheless, the shortcomings of even sophisticated tools for collaborative computing emphasise the importance of tacit knowledge and the need for more realistic models of human communication. Many informal but important work routines and processes are inaccessible (and hidden) to workers who are not physically present in the workplace.

The development of the World Wide Web very much echoes that of the Internet as such, although developer motives are often more manifestly financial. Considering the popularity of content-heavy news sites one may certainly overstate the case that communication and 'connectivity' is all-important, but few successful sites have ignored the fact that interaction is an important element in website success.

The development from IRC to instant messaging indicates that the desire for communication is far from quenched. It also highlights the facts that slight concessions to human interaction preferences may have a large effect and that usability issues can be crucial to the popularity of a service.

Even the most unambiguously task-oriented tools known to computer science have been subverted into channels for banter, jokes, games, and social interaction in general. What is curious, however, is how the history of modern communication seems to be the history of how this revelation has hit system designers, apparently nurturing entirely misconceived ideas of human preferences. Commenting in 1983 on early commercial computer networks one analyst wrote: *"…what Compuserve and the Source apparently didn't realize when they first put together their potpourri of consumer goods is that people are not crying for airline schedules and biorhythms, but to talk to one another*." (Carpenter quoted in Baym, 1996:146).

From a commercial viewpoint, huge investments into content-rich services seem surprisingly misguided considering early twentieth century experiences with telephone use. While one may attempt to combat "useless calls" from a moral

position (as did the city of Chicago in 1909) it is hard to refuse the evidence that social uses are a prerequisite for high media penetration (Odlyzko, 2001).

Summing up, two perspectives seem clear. One is that people seek communication through virtually every possible medium. The second is that notions of conflict-free interaction are destined to be challenged even in digital environments. It is the last perspective that will be explored in the following chapters.

CHAPTER TWO

The logic of cooperation

Modern society is a highly improbable phenomenon. Imagine, if you will, a physically unconnected aggregation of thinking agents without omniscient central authority working in concert to maintain complex patterns of mostly orderly and seemingly co-ordinated activity in a multitude of settings. Imagine also that these agents, even if not entirely selfish, perceive any sort of contradiction between personal and collective interests and so at times feel tempted to 'free-ride' on the efforts of others. How, in such a scenario, would it be possible to walk in relative safety on roads build by others wearing clothes made by others? This question lies at the heart of sociology and is often abbreviated to 'How is society possible?'.

Influential theories of the state have grappled with this problem, but before examining their conclusions, it is important to note that one may in fact be justified in turning to a more modest issue. This is the issue of how any sort of cooperation comes about or simply: How do two individuals manage to work together? Solving this more simple mystery does not explain the dynamics of modern nation states but if it can't be answered surely it would be foolhardy to tackle the more immodest subject.

In the following, I describe a short range of classical theories that most explicitly deal with the issue of cooperation. Writings within 'social contract theory' are emphasized since they are the first to systematically ponder the foundations of human interaction in a way that has clear implications for modern theories of cooperation. They are complemented by the work of Adam Smith, in which we find the most convincing concept of spontaneous social order derived from the economics of cooperation. Subsequently, I introduce concepts from game theory, which allow a formalisation of the earlier theories and a generalisation of the logic of cooperation.

Social contract theory

What is the state and how do we justify its existence? Such considerations were central to the thinking of Thomas Hobbes (1588-1679), John Locke (1632-1704) and Jean-Jacques Rousseau (1712-1778), the most historically important proponents of social contract theory. While proposing highly different constitutional ideals, they all agreed that the state was a necessity springing from human nature. Without it, life would not be tolerable or would at least be vulnerable to exploitation and cooperation deficiencies. People, in this system of thought, acknowledge their condition and enter into an agreement in which they surrender a part of their autonomy to an impartial entity. This entity, the state, by representing the collective will of the people is able

to legitimately restrict the freedom of individuals thus in fact increasing said individual's freedom from constant strife or external threats. The agreement, the social contract, is then a hypothetical construct; there is no actual document or historical act of entering[10]. Rather there is the philosophical conclusion that the contract is what people 'would have made' or may be 'thought to want'. While this is surely a weak point of any theory seeking to justify the state the ontological status that each author applied to the contract need not concern us further here. Rather the following will focus on their concepts of cooperation.

Thomas Hobbes; the necessity of sovereigns

Man, thought Hobbes, if not downright evil is purely and unconditionally selfish. In his treatise on *Leviathan* written in 1651 in an effort to come to terms with rapid social and philosophical changes, he states that *"...I put for a general inclination of all mankind, a perpetual and restless desire of power after power, that ceaseth only in death*." (Hobbes, 1651/1997:80).

The book, inspired by classical philosophy and fuelled by Hobbes' newfound rationalism has many such observations on human nature that bear the mark of a somewhat primitive behaviourism. He does, however, put forth the question of how cooperation is possible. His argument runs like this: Man sees obvious incentives for cooperation but fears to enter into any obligation if he has no solid guarantee that the other person will live up to his part of the bargain. The word of the other person means nothing if that person is not somehow forced to live up to it. Furthermore knowing my neighbour's lust for power, I realize that he also knows that I share this ambition. Surely, being of somewhat equal strength physical conflict would be dangerous (and better avoided if possible) but since I have no way of assuring him of my good will, conflict, or at least tension, is unavoidable. The only hope for peaceful coexistence is a powerful guarantor. In Hobbes' words:

> "...during the time men live without a common power to keep them in awe, they are in that condition which is called war; and such a war, as is for every man, against every man." (Hobbes, 1651/1997:100).

At this point Hobbes' argument is somewhat subtler than it is sometimes portrayed. While he does emphasize that human life without a strong central power is "*solitary, poor, nasty, brutish, and short*" his vision of the 'natural' (that is original, state-less) condition is not one of all-out physical war. Rather it is a condition where there is a

[10] Constitutions represent a formalisation of the agreement but the act of agreeing remains hypothetical. There is nowhere external to the constitution from which to enter into the community of law. One is born into the contract and while one may leave a national jurisdiction one may not claim to be outside the contract while inside the nation.

known disposition to fight "*during all the time there is no assurance to the contrary*" leaving "*no place for industry; because the fruit thereof is uncertain*" (Hobbes, 1651/1997:100).

A powerful sovereign punishing those who dishonour their obligations is thus a precondition for growth and social order and the reason that I surrender myself to the mercy of the state is my certainty that you do the same. The contract thus reads:

> "I authorize and give up my right of governing myself, to this man, or to this assembly of men, on this condition, that thou give up thy right to him, and authorize all his actions in like manner." (Hobbes, 1651/1997:132).

John Locke; the limits of government

Locke does not have much to say about the actual logic of cooperation, and I will not dwell upon the details of his analysis. Suffice to say that whereas Hobbes advocates autocratic, and all-powerful, government based on his conception of the natural condition as unbearable, Locke sees man as inherently virtuous. There are, however, problems of common resource distribution, destructive passions and of partiality in an altogether ungoverned population underlining the need for a state. Locke's analysis presented in his *Second Treatise of Government* (Locke, 1690/1952) focuses on God-given rights of individuals based on 'natural law', that is laws that must be taken to apply to all creatures of God. The natural condition

> "though this be a state of liberty, yet it is not a state of licence; though man in that state have an uncontrollable liberty to dispose of his person or possessions, yet he has not liberty to destroy himself, or so much as any creature in his possession, but where some nobler use than its bare preservation calls for it." (Locke, 1690/1952:5)

Since all men are bound by natural law, complete freedom is not something we have to give away. Hence the state can never be all-powerful but rather must base itself upon the confidence of the people – it must, in short, take the form of a liberal constitutional state.

Thus, through law fulfilling strict requirements, society becomes more just and through the state it becomes possible to bear the cost of common goods, such as a national defence, that no one individual would find possible to bear alone.

Jean-Jacques Rousseau; the will of the people

Much like Hobbes, Rousseau stresses that by surrendering our freedom to the state our situation is improved. Unlike Hobbes (Hobbes, 1651/1997:110), however, Rousseau finds no legitimacy in power:

> "A brigand surprises me at the edge of a wood: must I not merely surrender my purse on compulsion; but, even if I could withhold it, am I in conscience bound to give it up? For certainly the pistol he holds is also a power.

> Let us then admit that force does not create right, and that we are obliged to obey only legitimate powers." (Rousseau, 1762:7-8).

Rather Rousseau seeks the theoretical foundations of democracy, although without emphasising the limits of the common will that Locke found so important[11]. More relevant to the issue of cooperation Rousseau emphasises that certain modern developments have created the need for formal community in which individual motives are surrendered. The human race needs to act with greater force, only achievable "*where several persons come together*" (Rousseau, 1762:11).

Thus, Rousseau does not justify the state on the grounds that the natural condition is unbearable, but rather on the conclusion that the whole is greater than the sum of the parts. By joining the collective, all individuals become as one and there can be no conflict as there is only common interest. As has often been noted there is more than a hint of totalitarianism in such thinking but there is also, as we shall see in the following, an unrealistic notion of the degree to which individual motives can be surrendered in a stable manner.

Adam Smith; spontaneous cooperation

Sparing little attention for constitutional issues of legitimacy Adam Smith (1723-1790) devoted his *An Enquiry into the Nature and Causes of the Wealth of Nations* to the concept of social order and its relation to human nature and the division of labour. The state, in Smith's thinking, has the important role of securing such services as will not be produced by individuals acting in their personal interest; most importantly national defence and the judicial system. What is interesting here, however, is Smith's concept of spontaneous, unintentional cooperation. By following personal interests, people achieve unplanned improvements of the common condition. The individual, seeking personal gain is "*led by an invisible hand to promote an end which was no part of his intention*" and of equal importance: "*By pursuing his own interest he frequently promotes that of the society more effectually than when he really intends to promote it*." (Smith, 1776/1993:292).

This synergy effect does depend on the nurturing of certain personality traits and is vulnerable to insecurities of social unrest caused for instance by poverty, but it is still fundamentally different from Hobbes' idea that such cooperation could only arise in the presence of a strong guarantor. Hobbes' missing ingredient is the division of labour. In Smith's analysis, it is specialisation that makes market behaviour attractive

[11] Specifically, Rousseau imagines a process that Locke would have found appalling: "…*the total alienation of each associate, together with all his rights, to the whole community*." (Rousseau, 1762:11).

even to a hypothetical selfish actor. The baker, from self-interest not kindness, will produce bread and will sell it with a profit. The customer, from self-interest not kindness, will buy the bread and still be better off than if he or she had invested time and money in baking at home: "*It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner, but from their regard to their own interest*." (Smith, 1776/1993:22). Thus, cooperation emerges out of self-interest.
While Smith does not neglect the need to limit harmful effects of free competition he also, as is hinted above, warns against the construction of abstract systems that do not take fundamental human nature into account. Like the social contract theorists, then, Smith bases his analysis on somewhat crude and arbitrary conceptions of 'human nature'. How many assumptions about human nature we actually need to back an analysis such as Smith's, will be discussed at the end of this chapter.

Modern theories of cooperation

It is a testament to the unpredictable evolution of social theory that an inaccessible, almost cryptic, branch of mathematics known uninformatively as 'game theory' should be able to spark such controversy. Game theory in itself claims to say very little about complex social issues outside of economics, and indeed some would claim that it does not. However, with the addition of only the slightest of assumptions – most notably that of Darwinian natural selection – game theory suddenly seems capable of shedding light upon a huge variety of social issues. If appearing out of nowhere, 'theories of everything' should of course be regarded with the utmost scepticism, and surely there are those who have overstated the potentials of game theory. In the following, however, I shall present the case that it is not new at all but merely a formalisation of what has been stated in various forms since Hobbes. The aim is to arrive at a theoretical framework for cooperation that is applicable to computer-supported interaction.

The history of game theory

Mathematician John Von Neumann is normally accredited with inventing game theory (although others had previously worked on similar issues). His article *Zur Theorie der Gesellschaftsspiele* appearing in a mathematical journal in 1928 was among the first to formally describe game strategies and their relation to game types (Gates & Humes, 1997:2). Seeing potential for a radical strengthening of theoretical foundations within economics, von Neumann and Oskar Morgenstern in 1944 synthesised their perspectives in the book *Theory of Games and Economic Behaviour*

describing most importantly how economic transactions may be thought of as games (see Gates & Humes, 1997:2).

'Game' in this perspective is merely shorthand for any situation in which several agents seek to maximise their own utility by the application of strategies. Thus, *Chess* is a game and so is the stock market. What distinguishes the concept from (very) traditional economics is the importance placed on perception. To determine what an agent is likely to do, one must consider his perception of others, his perception of others' perception of himself and so on. This causes an infinite chain of interdependent variables the complexity of which can only be managed by applying strategies. Which type of strategy is optimal depends on game type and the strategies of other players.

Game theory, then, is concerned with how to achieve the best possible result from a game in which the rules are known and where the other player (we stick to two-player games for now) is expected to do his best.

It takes little effort to see how such a discipline would strike the interest of war-planners. Indeed, during World War II von Neumann's principles were employed in planning bombing raids over Japan (Poundstone, 1993:68). Meanwhile von Neumann himself worked on the less than innocent Manhattan Project and soon game theory would become a favourite pastime of the controversial American RAND Corporation. Public perception of RAND had it as a gathering of asocial intellectuals plotting on how to win the cold war, a perception not entirely untrue (Poundstone, 1993:83-99). Little surprise, then, that game theory was widely perceived as the epitome of dangerous rationalist sophistry.

Safely outside the searchlight of critical public opinion, however, game theory was applied to a variety of less-controversial issues, such as car insurance and marriage.

Such applications would not have been possible but for one strangely persistent topic emerging clearly from the complex mathematics of von Neumann's writings: the Prisoner's Dilemma. All other game theory virtues aside, it is this simple, hypothetical game that has turned the heads of almost every textbook editor within philosophy, political science and biology (Ostrom, 1990:5; see also Dunleavy & O'Leary, 1987:79-80). I shall describe the Prisoner's Dilemma (PD) in detail below, but for now it is sufficient to say that it is a formal description of what happens when two rational players need to trust one another in order to achieve the best overall result while faced with a temptation to free-ride, or less formally: a "*situation where you are tempted to do something, but know it would be a great mistake if everybody did the same thing*" (Ridley, 1996:56).

Game theory should be humbled by the fact that PD, despite its simplicity, led theorists to conclusions that needed highly inelegant theoretical scaffoldings. Considering our original question of how cooperation might arise, an odd answer was given: it could not; at least not among agents who knew what they were doing. To certain theorists, if people chose to co-operate in PD, they were not playing the game right.

What empirical observations of cooperation could not drive home was forcefully underlined in the late 1970s, when political scientist Robert Axelrod showed that if cooperation was indeed an irrational inclination it was one shared by computers. What Axelrod had in essence discovered was that the rationality of the PD changes if the game is played more than once. In a repeated PD, it pays to support the common cause. In other words: Cynical cold war abstractions could powerfully support the argument that being 'nice'[12] was a very rational decision; it was a message of hope. Within certain limits, greed was a sufficient motivator for cooperation. Axelrod had convincingly argued that when it came to the necessary conditions for cooperation Adam Smith was right and Hobbes was wrong.

The Prisoner's Dilemma

Bob and Alice are arrested, charged with bank-robbery and isolated. They are then both given two options: Each may either confess or deny the crime. If neither confesses, the police can only convict them of a lesser crime. If one confesses and the other denies the confessor is rewarded with instant freedom and the other is given a harsh sentence. If both confess they are both punished but with a slight reduction for confessing.

This is the Prisoner's Dilemma. The narrative framework, however, only serves to confuse the basic issue of the game, which is that Bob and Alice may choose to either co-operate or defect. On the whole, they will be better off co-operating, but seen through the eyes of either there is a dilemma as illustrated in **Figure 4**.

[12] That is "acting nicely", not necessarily motivated by altruism.

		Alice: Cooperates	Alice: Defects
Bob	Cooperates	**Bob=3, Alice=3** *Bob and Alice*: Reward for mutual cooperation	**Bob=0, Alice=5** *Bob*: Sucker's payoff *Alice*: Temptation to defect
Bob	Defects	**Bob=5, Alice=0** *Bob*: Temptation to defect *Alice*: Sucker's payoff	**Bob=1, Alice=1** *Bob and Alice*: Punishment for mutual defection

Figure 4 - The Prisoner's Dilemma.
Numbers indicate points given (the more points the better). The overall maximum is achieved through cooperation (6 points) but the individual maximum (5 points) is achieved through defecting while the other player co-operates. In essence, points need only follow the relative ranking showed.
Based on: Axelrod, 1984:8.

Assuming a single round and a desire to win, what will Bob do? He will consider Alice's move, which is either *co-operate* or *defect*. If he believes that Alice will co-operate he should defect (5 points rather than 3). If he believes that Alice will defect he should also defect (1 point rather than 0). Bob defects. And based on the same analysis so does Alice (Axelrod, 1984:9).

If the two players could only trust each other, they would, taken together, achieve a better result. This separates the PD from most recreational games. In *Chess*, for instance, there is no incentive for cooperation. In game theoretical terms *Chess* is zero-sum, one player wins as the other player loses, while PD is non-zero-sum, the total number of points depends upon player behaviour (see Gates & Humes, 1997:1-5).

A series of one-shot PDs is the natural condition imagined by Hobbes. His solution is to dissolve the dilemma by changing the score. Essentially what the state does is do away with the temptation to defect by lowering the payoff for defecting against a co-operating other player. If Alice sells Bob her bread and Bob cheats Alice of her payment, he will be punished by law. Bob behaves and the problem is solved[13].

Implications of the Prisoner's Dilemma

What does this abstraction tell us? The PD is just a formal way of stating what is perhaps *the* challenge of actual cooperation. The welfare state, for instance, is a non-zero-sum game. By pooling resources, by paying taxes, it becomes possible to procure public goods such as roads, hospitals, and the education system. If all citizens cooperate, the collective is better off. From an individual point of view,

[13] Naturally whether or not Bob behaves depends on his perception of the score. If he is poor and starving to death the threat of punishment is not likely to deter him.

however, it would be more attractive to enjoy collective goods without contributing (that is, the individual has an incentive to defect). A high level of cooperation is achieved by constructing institutions that punish defectors such as the police and the judicial system. Importantly, these systems do not mainly punish but by their very existence ensure would-be-cooperators that they are not taken advantage of[14]. Although not in itself supportive of Hobbes' argument, the dynamics of the welfare state reflect his cynical prescription. Supporting cooperation by the introduction of a neutral guarantor is clearly possible in a wide range of situations[15].

But far from always. The Cold War is an instructive example. We may, in this case, treat the superpowers as players since national and individual interests coincide on the issue of avoiding nuclear attack. The unfortunate situation was that both super powers had an incentive to attack, as long as they would not themselves suffer (if nothing else this incentive was the possible freeing of tremendous resources spent on armament, espionage and diplomacy). Mutual cooperation was a goal, while it is, of course, debatable whether the sucker's payoff would be worse than mutual defection (see also Axelrod, 1984:181).

Each day of the Cold War may then be seen as a Prisoner's Dilemma. No all-powerful guarantor existed, so no Hobbesian solution was within reach. Still neither 'player' ever defected, so somehow the scores must have been changed. In PD terms the key to cooperation was early warning. If defecting on any given day meant that the other player would get an extra turn in which to retaliate, the temptation to defect was eliminated. Cooperation had, under specific conditions, arisen from sheer self-interest.

It would seem that the PD does not predict constant defection after all.

The shadow of the future

Posing the question in strategic terms political scientist Robert Axelrod repeated the age-old question: "*Under what conditions will cooperation emerge in a world of egoists without central authority?*" (Axelrod, 1984:3). Axelrod's ambition was to determine which strategy would do best in a series of repeated PDs. To answer this he set up a computer tournament challenging experts in game theory to each submit a strategy that would do better than all other strategies. Contestants were highly

[14] In this perspective, it is my trust in the taxation authorities that ensures my trust in the cooperation of the person living across the street, whom I have never met.

[15] While possible, the result is often a far stride from the collective harmony imagined by both Hobbes and Rousseau as has been emphasised by the 'public choice' school of political thinking (Dunleavy & O'Leary, 1987:159-163; Stroup, 2000).

creative, yet among wonders of sophistication the simplest of strategies emerged successful: Tit-for-Tat. Tit-for-Tat (or the player using it) merely cooperates on the first round and then does whatever the other player did in the previous round. Thus, the tournament winner was 'nice'; it never instigated an aggression. In other words: the logic of the PD had been dramatically changed.

This is best explained by returning to Bob and Alice. In this case, they are neighbours; that is they are likely to meet repeatedly. Suddenly the future enters into the equation. Bob rightly believes that Alice's strategy is Tit-for-Tat (she may have told him or he may have deduced it). If he defects on round one (5 points against Alice's 0) Alice will defect on round two (1 point each). Bob's only hope for forgiveness is to cooperate while Alice defects (0 points for Bob). Considering that cooperation will earn Bob 3 points in all rounds defecting in any round will make him worse off than constant cooperation[16]. While he could easily beat Alice by always defecting (he would lead by 5 points) beating the other player is not what PD is about. The rational player seeks to maximize his or her own score and should ideally not care how many points the other player gets. Tit-for-Tat in fact "*won the tournament, not by beating the other player, but by eliciting behavior from the other player which allowed both to do well*." (Axelrod, 1984:112).

Tit-for-Tat, however, has another claim to fame. As a strategy, it is collectively stable. It not only spreads in a population of adaptive players, it is also self-policing; no other strategy can invade it once it gets a hold.[17] So, if people in a group use this individually rational strategy, cooperation may be stable.

Exit the temptation to free-ride. In terms of systems, we have done away with a major threat to stability: the problem of suboptimisation (Heylighen, 1999). This is merely the observation that any system in which selection works on individual parts must have corresponding collective and individual interests to be stable. Without such correspondence, selection pressure will "urge" individuals to free-ride.

Conditions for spontaneous cooperation

Axelrod gives a thorough discussion of rational behaviour when faced with a PD (basically just an analysis of Tit-for-Tat's behaviour as it is perceived by others). Our

[16] This logic depends on the exact point distribution but for the present purposes we may ignore such complications.

[17] For a description of how Tit-for-Tat may get a foothold in a ruthlessly selfish population see Axelrod, 1984:98. Tit-for-Tat's collective stability has been refuted but the premise that being 'nice' is prudent in repeated PDs holds (Ridley, 1997:74-84; Zyvicki, 2000; Dawkins 1976/1989:216). Axelrod's response to the critique is outlined in chapter two of Axelrod, 1997.

primary concern, however, is how to design a system in which it is rational to cooperate. Following Axelrod's analysis, three necessary conditions are evident:

> **Repeated interaction**: Only when the shadow of the future influences an encounter does the logic of the PD change in favour of cooperation. Consequently, the players must be aware (or think) that the probability of future interaction is sufficiently large and important. In human social groups, this may be taken less literally to imply that Bob should be aware that he is dependent upon public perception and the good will of Alice in the future. If he is convincingly and publicly labelled as a defector his future interactions will suffer (for perspectives on gossip, see Burt & Knez, 1996:72).
>
> **Interaction history**: Information on previous interaction must be available. Bob must know how Alice acted in earlier interactions in order to assess the situation and plan his strategy.
>
> **Persistent, distinct identities:** In order for the above to be possible Bob must be able to recognize Alice. For humans and animals, this typically means devoting a section of the brain to facial recognition (Axelrod, 1984:100; see also Zyvicki, 2000).

From Axelrod's simple simulations, then, rather breathtaking implications arise. Here, we shall content ourselves with present day human beings but might the analysis in fact be applied to all levels of organisation sharing certain characteristics? Evolutionary biology, in particular, has given the issue serious consideration. The rationale applied in that work is worth touching briefly upon, since it leads to important tools for explaining more complex forms of cooperation.

Biological perspectives

Meet others openly but don't let people take advantage of you. If this seems like a reasonable social rule, it should come as no surprise. After all it's Tit-for-Tat, the most reasonable (rational) PD strategy. In the discussion above, I have intentionally refrained from putting the problem in genetic terms although the most wide-ranging implication of niceness sometimes being rational is evolutionary. I shall neither go into details nor base the coming analysis on evolutionary foundations but merely sketch the theoretical consequences of such an approach (which has received careful attention following Maynard Smith, 1976).

'Rationality' as used above is close to synonymous with 'fitness' in the Darwinian sense: that trait which is favoured by selection. If cooperation in some form is 'rational', it is also conducive to fitness. Thus, natural selection should favour Tit-for-Tat players: we should be nice. And indeed, to a large degree, we *are* nice in the special way used above. And not only we: fish swim in pods, beavers build dams,

wolves hunt in packs, and vampire bats famously share their food (Dawkins, 1976/1989:231).

Whether or not this should surprise us depends on our level of analysis. If the group or species is the unit, which undergoes selection, cooperation is an obvious Darwinian adaptation that needs no further explanation. But (most) modern biology does not look kindly upon theories of group selection. Groups don't replicate and therefore are not subject to selection. The unit that natural selection works upon is the gene, hence Richard Dawkins' oft-mentioned 'gene's eye view of natural selection' (Dawkins, 1976/1989; Dawkins, 1982/1999; see also Axelrod, 1984:89; Ridley, 1997:17-20).

For the following discussion, what is important is that the unit of analysis makes all the difference. Rationality (and fitness) in itself is nonsense; if the term should be employed at all, it is important to remember that something can only be rational relative to something else. What is rational at the level of the gene may not be rational at the level of the individual (e.g. self-sacrifice to save one's family). The fate of the gene is closely connected to the fate of its carrier, the organism, which is in many ways dependent upon the fate of the group and the fate of the species. It is, however, the lower level that matters in the evolutionary process (or so Dawkins and others have argued rather convincingly; objections, however, have been raised).

Reciprocal altruism

One important concept, which has direct consequences for this analysis, has received careful attention within biology. Reciprocal altruism by which "*an individual provides a benefit for another in exchange for a reciprocal benefit, or the expectation of a reciprocal benefit in the future*" (Zyvicki, 2000:17), poses an obvious challenge to any explanation that would put individual rationality first in all things. Axelrod's repeated PD underlines the conditions under which Bob and Alice will cooperate at the same time. However, if Bob is in danger of drowning or is begging in the street there is no mutual cooperation and we might assume that Alice should be indifferent to Bob's trouble. But the real world is filled with situations where people (and animals) do help each other without demanding immediate repayment. Either they're being truly altruistic[18] or they are, at some level, expecting to be paid back by someone or something. Biologist Robert Trivers (1971) showed

[18] True altruism is a tricky concept. If I take pleasure in the well-being of others helping them would be the selfish act of helping myself. If I am kind in the hope that I shall be rewarded in the afterlife, I am still arguably being selfish. Since my desire is what I want to do, I can only – contradictorily - be truly altruistic by mistake. The term is used here to loosely describe voluntary actions that are at no level beneficial to the agent.

how such behaviour could be explained within an evolutionary framework. The problem has also been tackled in sociology under the heading of 'generalised exchange' (Kollock, 1999:222).

For such a system to work, however, there must be easy recognition and good memory. Bob is not likely to lend Alice any money if she could be gone the next day or if he fears that either party might forget the debt. Reciprocal altruism is a vulnerable condition as it "*differs from other patterns of cooperation in that it is fraught with risk, depends on trust, and requires that individuals whose contributions fall short be shunned or punished, lest the whole system collapse.*" (de Wall quoted in Zyvicki, 2000:17). A group of people enjoying the fruits of reciprocal altruism will have a clear interest in stability. For the complex debt system to work over time, group boundaries must be well-defined. Too many new-comers or too many quitters undermine the fragile basis for long-term investments. In short: Non-zero sum situations provide the basis for cooperation but the important temporal distribution of favours is likely to be antithetical to generous access to group membership.

N-person games, free-riders, and the Tragedy of the Commons

Considering implications of the PD (page 35) I used the welfare state as an example. But in fact, applying the logic of simple two-player games to larger aggregations requires a leap of faith. Or rather: It requires an explanation using a few additional concepts. Turning to real-life groups what we want is responsible use of *common goods*[19]. We want to produce resources that are useful to all while making sure no-one free-rides on the investments of others. We want, as one traditional example goes, to build a lighthouse while making sure that those who benefit also contribute to its construction and maintenance. More pressingly, we want to make sure that no individual selfishly depletes a finite common resource. Following Garrett Hardin's example of shared access to grassing grounds; failure to achieve this goal is often said to result in a 'tragedy of the commons':

> "Therein is the tragedy. Each man is locked into a system that compels him to increase his herd without limit -- in a world that is limited. Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons. Freedom in a commons brings ruin to all." (Hardin, 1968:5).

Such statements have earned Hardin the expected wrath from less government friendly parties, some of whom claim that Hardin was just plain wrong (e.g. Ridley, 1997:231; see also Kollock & Smith, 1996:110). The debate, obviously, quickly

[19] A common, collective, or public good is any good that, if available to some group members, cannot feasibly be withheld from the others (following Olson 1965:14).

becomes highly political. As before, I shall acknowledge the Hobbesian solution but concentrate on systems with no external authority.

Reciprocal altruism sheds some light upon the dynamics of n-player games. There are at least two ways in which cooperation in larger groups may arise in the first place. One is the emergence of metanorms, norms that prescribe that norm-breakers must be punished (Axelrod, 1997:50-68). Alice observes that Bob defects on Eve, and Alice then defects on Bob. Bob's incentive to defect is slight. A second possible road to cooperation is social ostracism: Free-riders are labelled and exiled from future interactions. Bob is well aware that abusing the commons will mean that other villagers will refuse to deal with him in any way.

Both possibilities are only off-putting to the would-be defector under the shadow of the future and with some degree of surveillance. Labelling of free-riders may, of course, range from verbally warning others to actual branding, cutting off hair or hands and a host of other techniques used throughout human history.

Future interaction, knowledge of the past, and powers of recognition remain essential to the emergence of cooperation[20].

Problems of game theory

"*To reduce the complexity of life to a silly game is the kind of thing that gets economists a bad name*." writes Matt Ridley (Ridley, 1997:56) in an effort to do just that. The cynicism of Hobbes has earned him a fair amount of scorn within political thinking but such image problems are nothing compared to the deluge thrown at those who have championed biological perspectives on human behaviour. Quite strikingly, this fate has befallen so-called sociobiologists in the last decades following Wilson's seminal (and provoking) *Sociobiology* (1975/2000) and Dawkins' *The Selfish Gene* (Dawkins, 1976/1989). Wilson was instantly labelled "*a racist, a sexist, a capitalist imperialist*." while "*His book was characterized as a right-wing plot, a blueprint for the continued oppression of the oppressed*." (Wright, 1996:345). The historical explanations for such explosive reactions are tied to unfortunate forays into highly swampy common territory of the social and natural sciences. Most notably, social Darwinism represents a dark chapter in the history of interdisciplinarity, an experience that rimes well with Emile Durkheim's warning against drawing directly on biological explanations within sociology (Durkheim, 1895/2000). Of course, one should not ignore that biology since Wilson has

[20] For further discussions of n-player PDs see Ostrom, 1990; Ridley, 1997:230-233; Axelrod, 1997:40-68.

challenged the scientific division of labour by venturing into the domains of psychology and the social sciences, shooting heavily at existing paradigms (notably the much-reviled whipping-boy of Darwinists: The Standard Social Science Model; see Wright, 1994:6; Wilson, 1998:207; Dennett, 1995:490-491).

The debate about the merits of sociobiology is obviously beyond this thesis and I shall only comment briefly upon it towards the end of the last chapter. However, we need to establish whether or not game theory is theoretically sound as it applies to human behaviour. I will approach this problem on two levels. First, I shall discuss whether Axelrod's analysis holds on its own premises. This is followed by an analysis of the major theoretical assaults aimed at explaining human behaviour by assuming rationality in any form.

Internal problems

Accepting, for now, Axelrod's assumptions that his results have real-life implications, allows us to ignore his premises and critically review the validity of his study. The question, as mentioned, has not changed since Hobbes: "*Under what conditions will cooperation emerge in a world of egoists without central authority?*" (Axelrod, 1984:3). Whereas Hobbes' answer was that it would not, Axelrod shows that it is merely a question of set-up. But by his very answer he appears to inadvertently display his own weakness: the tournament rules. Axelrod, it seems, considers his tournament unbiased and 'neutral'. In this he is entirely justified; but only in the judicial sense that rules that apply to all are neutral. A law forbidding *all* citizens to sleep under bridges may fulfil all legislative requirements while in effect only apply to the very poor. Axelrod thus takes the constitutional abstraction that 'all strategies are created equal' to justify an arbitrary tournament framework. One might wonder whether the outstanding success of Tit-for-Tat is entirely (or even to some degree) a function of the tournament rules. An illustrative parallel issue may be raised as to the capabilities of the strategies. The sports connotations of 'tournament' perhaps make us swallow too readily the high degree of measurement accuracy and rules clarity. We needn't venture into the social world of real people to realize that few systems are so acutely transparent as Axelrod's computer arena. Now, the whole digital battlefield is only a model and as such it should not be criticised for failing to take every conceivable aspect into account. It would be critical, however, if the introduction of some simple, 'realistic' factor could seriously skew the results. What, for instance, would happen if there was the slightest risk of misinterpretation? What if Bob could not be absolutely sure that he had perceived Alice's action correctly, or

similarly if Alice could not be absolutely certain that she had done as she intended? Tit-for-Tat would go down in flames.

If Bob and Alice are both using Tit-for-Tat and there is a misinterpretation on round X, the result will be never-ending mutual punishment (as Axelrod was well aware, see Axelrod, 1984:38)[21].

Such objections have been raised and various rival strategies have been proposed, most notably 'Pavlov' and 'Firm-but-Fair' (Zyvicky, 2000; Ridley, 1997:98).

Axelrod, to his credit, goes out of his way to counter such criticism by eliminating the arbitrariness of the available strategies. By way of a 'genetic algorithm' he allows the strategies themselves to evolve. In this study he used "*a population of strictly random strategies (chosen from a huge universe of possible strategies that had no bias toward either cooperation or reciprocity), and let the population evolve using the genetic algorithm*." (Axelrod, 1997:11). The result was that: "*Within a very short time, the population evolved toward strategies that behaved very much like TIT FOR TAT, and were certainly achieving cooperation based upon reciprocity*." (Axelrod, 1997:11).

While such manoeuvres do not actually repel the criticism, what is most important is that no objection has undermined the important conclusion that under the conditions described above, cooperation is the rational choice.

External problems

Do real people understand the complex calculations of the PD and act to increase their own score? In an instance of eager eloquence, Hobbes lucidly put the issue in different terms: Only he who does not lock his front door has the right to even ask (Hobbes 1651/1997:100). The very act of protecting one's possessions shows an understanding and a rational response.

This polemic passage is impressively forceful as it ignores people's self-perception and verbal self-representation (both of which may be strategically inaccurate) thus reducing the question to one of rational behaviour compared with actual behaviour. Forceful as it may be, it unfortunately raises another question: what is rational behaviour?

The concept of 'rationality' is surrounded by a frightening amount of confusion and whether is has any real explanatory value on the issue of human behaviour may well be questioned (see also Smith, 2001). However, it is often used in one of two general ways. *Either* in a "phenomenological" (internal) sense, in which people analyze their

[21] Unless there is a subsequent misimplementation by either player leading to various cycles of suboptimal play.

situation and then consciously act to achieve the best result. *Or* in a ”behaviouristic” (external) sense, in which rational behaviour is an entirely external characteristic. Taking medicine which you think will cure you (but which is ineffective) is internally rational but not externally rational. Taking your next breath is only externally rational (unless you carefully consider other options). Thus, arguably, if something is *very* rational it is likely to be internalized even if one has to learn it (e.g. eating with a spoon).

In all this, the operative noun remains ‘confusion’. However, when the term is employed in this thesis it is in the external sense. Rational behaviour benefits the agent as to whatever selection applies.

Confusion aside, economics is sometimes criticised for attributing too much knowledge to the individual. In fact the scope of this criticism is often overrated since perfect knowledge is not a necessary condition for market systems to work; the individual needs only very limited knowledge in order to take part in complex production processes (the Canadian wood-chop does not need to know that the wood becomes German pencils). At its most direct, however, the criticism is illuminating. In Herbert Simon’s words:

> “…economists attribute to economic man a preposterously omniscient rationality. Economic man has a complete and consistent system of preferences that allows him always to choose among the alternatives open to him; he is always completely aware of what these alternatives are; there are no limits on the complexity of the computations he can perform in order to determine which alternative is best; probability calculations are neither frightening nor mysterious to him.” (Simon, 1945/1997:87).

Simon contrasts this with social psychological axioms that would reduce all human action and perception to external dynamics, thus completely abandoning individual rationality and arrives at a reasonable diagnosis: “*The social sciences suffer from acute schizophrenia in their treatment of rationality*.” (Simon, 1945/1997:87). Simon emphasizes that people are never omniscient and that the costs of seeking information must always be taken into account, thus advocating the notion of “bounded rationality”. While contributing impressively to a realistic conceptualisation, however, this does not answer the question of which unit should be targeted by analysis.

In one sense, the difference is merely one of levels. One may wish to study individual cars or one may wish to study traffic patterns. However, to *explain* traffic patterns in anything but probabilistic terms one will often need to understand the workings of individual cars where all choices (in the broadest sense) are made[22].

[22] This is not to say that studying one car will enable us to predict traffic patterns. Just as looking at one human being out of context would not enable us to predict actual social patterns. It may be worth

Before leaving this metaphor behind for good, it should be clear that this thesis is concerned with the relationship between infrastructure and individual driving. The infrastructure is likely to affect driving choices but does not entirely determine them.
Although enjoying common characteristics (they are both highly deterministic) individual and group rationality lead to different predictions. Group rationality can only be taken to mean that group members discard their personal interests in order to achieve a common goal. This position has been all but surrendered following the seminal analysis of Mancur Olson's *The Logic of Collective Action* (Olson, 1965; this corresponds well with developments in biology, see page 39). Opposition to high-level explanations is a fundamental premise in rational choice theory (e.g. Elster, 1989) a school of thought often associated with game theory.
This association is not quite fair. There is the important difference that game theory does not assume neither conscious nor unconscious rationality on the part of the individual. There is no prediction in Axelrod's theory that individuals will use Tit-for-Tat (unless one takes the evolutionary perspective). Nor does expanding the scope to include actual human behaviour rest upon the assumption that people are ruthlessly selfish; for the analysis to apply we need only the slightest tension between collective and individual interests (Axelrod, 1984:7). We find only the incontrovertible conclusion that those agents who *do* apply a strategy that corresponds favourably to the selection in the system will do best (and if adoption is possible the successful strategies will multiply).
Social game theory thus goes a long way towards bridging the gap between explanations of human action that would emphasise either selfishness or altruism. For a social animal with our cognitive capabilities, the dichotomy is close to unimportant: Either motivation gives the same result[23]. Game theory thereby dismisses a frequently heard statement that trust cannot be explained in purely instrumental terms (e.g. Tyler & Kramer, 1996:10). Perhaps it cannot, but proving this will at least entail falsifying Axelrod.
This leads to a challenging aspect of Axelrod's theory: In its basic form, it is not open to falsification. This is not as much a problem of the theory as an aspect of any logico/deductive argument. Again, as an example, we may draw upon the work of Darwin. If an organism were found that could not have evolved by natural selection, it would falsify a theory stating that all life on Earth has evolved by this principle.

noting that this has little bearing on the Durkheimian axiom that social phenomena can and should be studied "as facts" (Durkheim, 1895/2000:58).

[23] I do admit that this point belongs squarely on the "external" side of the rationality dichotomy mentioned above.

But it wouldn't falsify the general theory that complex things can over time come about through the process of random change combined with selection. Similarly, no observation can challenge the basic logic of repeated PDs – it is true by definition (thus one would have to attack the premises). This reflects on Axelrod in the way that he may be said too easily to dismiss his need for an underlying theory in order to generalise his analysis to human social life. It may, however, reflect even harder on this thesis. The explicatory link between game theory and human life *is* (to a large degree, but far from entirely) evolutionary biology. And since I have chosen not to rely on evolutionary arguments for reasons of volume, nor to make bold assumptions about human nature, the step from game theory to predictions of human action is not quite secure. However, for now I shall consider it a likely possibility that people do employ reciprocal behaviour. That this is a fair assumption in the case of CMC will be supported in **Chapter four**.

Returning to the realm of advantages, another dichotomy to be challenged by game theory is that of emotions as a counterpoint to rationality. Sticking with rationality in the external sense, emotions provide a useful tool in bargaining and commitment situations (Trivers, 1972). The threat of what would look – to a short-sighted calculating agent – like an overreaction is a powerful defection deterrent[24].

Axelrod's analysis has proven resistant to critical analysis. Nevertheless, Hobbes made a good point: The theory should not be ultimately judged on its own premises but by whether it corresponds with empirical reality. Premises aside, does game theory provide a credible explanatory framework for what we know of online interaction? This will be the topic of chapter three.

Summary and theoretical perspectives

If we believed man to be truly altruistic there would be nothing to explain, social life would be bliss. If we acknowledge even a slight tension between individual and collective interest and admit that collective action is non-zero-sum we are left with the Prisoner's Dilemma, a game theoretical formalisation of age-old political analyses.

In this light, the social contract theorists asked the right question: How is cooperation possible? By contradictorily reducing human interaction to a series of one-shot PDs, however, they arrived at the wrong conclusion that it could *only* come about through a powerful guarantor. They failed to see that the social world of human beings can

[24] Interestingly, such observations rhyme well with more recent claims that thinking and feeling are two sides of the same coin (e.g. Damasio, 1996).

function in much the same way as Hobbes' *Leviathan*, it ensures that actions have consequences. Cooperation, being nice, is the rational thing to do.

But only under specific circumstances. For cooperation to arise, the system must provide *repeated interaction*, *knowledge of interaction history* and *recognition capabilities*. For temporally distributed cooperation, reciprocal altruism, to function group boundaries must also be clear.

In this perspective, then, cooperation is a function of system architecture. Individuals may have all sorts of motives and values but if *any* member of a group is just slightly selfish and if the others are just slightly risk-averse, stable cooperation is dependent of the presence of the above characteristics.

CHAPTER THREE

Computer-supported cooperation

In chapter one I outlined major elements of contemporary and historical online interaction with an eye for conflict. Chapter two described the basics of cooperation between selfish agents without central authority, concluding that such a state depends on reconciling individual and collective interest. Specifically, this may be achieved by making the future matter to the individual, thus supporting reciprocity and mutual dependence.

In the following, I shall do two things. Firstly, I describe the observations of chapter one in the theoretical language of chapter two. The aim is to convincingly show that cooperation theories may be fruitfully applied to computer-supported interaction. Secondly, I draw upon communication theory and CSCW to determine whether there are analytically significant differences between physical and computer-supported cooperation.

Online interaction revisited

To tentatively establish whether game theory may shed light upon CMC phenomena, we must determine what the theory predicts. First of all, it predicts that exploitable systems will be visited by conflict. Furthermore, it opens the possibility (to be supported in **Chapter four**) that if people want interaction at all (and chapter one established this) they will want it to live up to the necessary conditions for reciprocity – they will want the possibility of recognition and a degree of stability.

Travelling back in time to the primal scene of online conflict, the fall of *CommuniTree* now seems less than surprising. Nor should the initial state of bliss that sparked such high hopes surprise us. It is, as I have mentioned, a common observation in online communities that pioneers are enthusiastic and the exclusivity of early computer users obviously worked much like any tight-knit lodge to strengthen inner ties. Furthermore, users were most likely recruited by friends whose reputation would then depend on the good behaviour of the newcomer and interaction between two given users was frequent due to the small size of the group[25]. Finally, it is worth remembering that there is no reason to expect anti-social

[25] Group size is an important issue in supporting cooperation. Generally speaking: "the larger the group, the less it will further its common interests" (Olson, 1965:36). Or rather, the harder it gets to fulfil the necessary conditions for recognition etc. For details, see Kollock & Smith, 1996:118.

behaviour in the first place. We should not expect that people set out to break things, only that things that may easily be broken over time *will* be broken.

But then came the saboteurs. The destructive students did not play the noble game of intellectual give and take, and there was no way to make them. Extreme personal freedom had destroyed the community. The lesson, however, was not wasted:

> "Within a few years there was a proliferation of on-line virtual communities of somewhat less visionary character but vastly superior message-handling capabilities – systems that allowed monitoring and disconnection of 'troublesome' participants… Thus, in practice, surveillance and control proved necessary adjuncts to maintaining order in the virtual community." (Stone, 1992:91).

When online communities, despite all experience, are often established in high libertarian spirits it may of course be contributed to the ambition of the designers. Few MUD creators set out to recreate the intricate workings of liberal democracy; they set out to have fun (and sometimes to make a living). Most designers probably also sense that trust is a two-way process. Build a totalitarian nightmare and users are unlikely to respond by unguarded altruism (if they come at all). In fact, studies show that an explicit focus on the risk of exploitation is harmful to cooperation, that trust requires a certain blindness. Economists, it can be argued, by spelling out the logic of the PD actually 'make' people behave anti-socially (for a survey and criticism see Frank & Gilovich & Regan, 1993; for perspectives on the relationship between guarded cooperation and trust see Meyerson & Weick & Kramer, 1996:189)[26].

Designers are not alone in preferring informal social bliss. Most users, of course, would also rather have everyone behave acceptably than be bugged down with the time-consuming rituals of due process. After the Dr. Bungle affair in *LambdaMOO* (see page 18) one victim commented:

> "I'm not calling for policies, trials, or better jails. I'm not sure what I'm calling for. Virtual castration, if I could manage it. Mostly [this type of thing] doesn't happen here. Mostly, perhaps, I thought it wouldn't happen to me. Mostly, I trust people here to conduct themselves with some veneer of civility. Mostly, I want his ass." (Dibbel, 1999:16).

The conclusion that dawns upon the online optimists is the cynical one of the contract theorists: Universal natural kindness would be both appealing and 'cheap', but as a basis of society, the assumption that everyone is nice is an invitation to trouble. When it comes to sensible behaviour in PD-like settings, the old proverb is not far off the mark: 'Trust in God, but keep your powder dry."

[26] Such criticism is routinely directed at theories that would challenge an intuitive/traditional image of the world, such as Darwin's theory of the origin of species (Wright, 1996). While it should not be ignored that a description may indirectly shape its object, the criticism goes to the *effect*, rather than the *coherence*, of the theory and thus is not primarily a scientific criticism.

Fortunately, most MUDs can change in the face of conflict. This is due to the fact that they are in a sense only masquerading as free-for-all limitless playgrounds. MUD designers may speak the rhetoric of anarchistic social experiment, but they do so from a rather privileged position, that of a god. The players have all the free will they can use, in no sense are they puppets on a string, but the natural laws of the digital world are entirely defined by the designer. And the power structures, of course, are in the code (this discrepancy between ideals and reality is analysed by Pargman, 2000:197-204). Players may enjoy a high degree of freedom, but at any given time they risk elimination at the hand of the managers (as indeed happened to Dr. Bungle). Besides, any sort of social change that relies on technical modifications needs the approval of the administration. Thus, in an important sense, the emergence of law in such systems is not so much a Hobbesian process of going from the natural condition to an autocracy but a Lockean one of going from God-given natural rights to a liberal state (although most often one where God remains central).

The apparent case that the privileges of power are hidden to the privileged designers is not our main concern. However, the fact that most MUDs are strictly hierarchical explains why they can survive even serious crises[27]. Major changes can only happen with the blessing of the administration. The experiences of *LambdaMOO*, *MicroMUSE*, *Habitat* and the MMORPGs is close to what game theory would have us expect. With limited groups of carefully initiated users, things work well. But with increased group size and low importance of future interaction (no privileges or investments were tied to the user profiles of the saboteurs) freedom was exploited. From this we should expect a call for increased surveillance and/or control, and indeed this was soon heard. The correlation between (inner) harmony and limited group size was even explicitly commented upon by *LambdaMOO* users who referred to the first months as "*a Golden Age – a time when MOOers lived in peace and productivity and had no need of rules or disciplinary structures…*" (Dibbel, 1999:204). Furthermore, alarmed at "*the explosive population boom… older residents bewailed the declining quality of Lambda life and called for an immediate locking of the gates against the newbie hordes.*" (Dibbel, 1999:96).

These users are fully aware of what is going on. Or if not fully conscious of the fundamental dynamics of reciprocal altruism at least they act as if they are, which is sufficient for present purposes. In this respect, they resemble those Internet regulars who showed resentment towards America Online customers who at a time where

[27] At times, administrator commands invented for administrative purposes may also be turned into punishment techniques (Pargman, 2000:243).

thought to be especially crude and devoid of 'netiquette' (Donath, 1999:36). MUD users may not describe their behaviour in game theoretic terms but as Hobbes commented, actions can speak for themselves.

Another suggestive pattern of action has been observed. "*Adventure MUD users tend to view each other with some suspicion… conflict is not an inherent facet of social MUDs as it is of adventure MUDs*." (Reid, 1999:129). There may be several reasons for this difference (and I shall describe another in the following under the heading of metaphor). It does, however, suggest that the zero-sum game element of adventure MUDs (they are more like *Chess* than PD, see page 35) inhibits cooperation.

Arguably MUDs present a case which is, if not in fact special or analytically distinctive, somewhat difficult. To avoid becoming lost in the wilderness of uncountable variables, some researchers have chosen to focus on more manageable setting such as newsgroups and IRC channels. In these systems, users exhibit strong commitment to protecting the integrity of identity (perhaps contrary to common belief). In an influential early study of IRC Elisabeth Reid noted:

> "One of the most sensitive issues amongst users is the question of nicknames… These aliases are chosen as the primary method by which a user is known to other users, and thus generally reflect some aspect of the user's personality or interests. It is common for users to prefer and consistently use one nickname… Names, then, as the primary personal interface on IRC, are of great importance. One of the greatest taboos, one that is upheld by the basic software design, is the use of another's chosen nickname.
>
> The illegitimate use of nicknames can cause anger on the part of their rightful users and sometimes deep feelings of guilt on the part of the perpetrators." (Reid, 1991:19).

A rather similar pattern of name protection can be detected in MUDs (Curtis, 1992). Although not working within an explicitly game theoretical framework, Reid's conclusion echoes Axelrod. While acknowledging that one motivation for IRC use is experimentation, she finds that "*Experimentation ceases to be acceptable when it threatens the delicate balance of trust that holds IRC together*" and if recognition were not enforced "*it would be impossible for a coherent community to emerge*." (Reid, 1991:20).

Those who would threaten stability, even if only from ignorance of local customs, quickly realize that social control occurs even in the allegedly anomic reaches of decentralist cyberspace. On Usenet, norm violators are pointed out as such, often ridiculed and sometimes ignored entirely, which may be the most severe form of punishment in a social setting (Baym, 1996:160; Gotved, 1999:190; for details on social control on Usenet see Smith & McLaughlin & Osborne, 1997). This is very much in line with the idea that people seek reciprocity and work to protect common goods from saboteurs or free-riders.

However, what is equally noteworthy of Usenet newsgroups is the fact that they mark the limits of a game theoretical explanation. When researchers have so little trouble finding sociologically striking forums it is of course because they are the ones that exist. A forum without interaction – or without a reasonable level of cooperation – is of little value to anyone and so is bound to be shut down or attract no attention. Some forums thrive and some die. Since all newsgroups are architecturally alike, the explanation for their diverse fates is bound to fall entirely outside the domain of game theory (at least as regards the question of design as applied here). It would be highly interesting to know what variables (e.g. user personalities, topic categories, early interaction characteristics etc.) determine which newsgroups become thriving communities and which become digital junk yards (such a study might build on the results of Jarvenpaa & Leidner, 1998).

Not only typical MUD development deserves the heading of wishful thinking. Groupware implementation, despite high hopes, shows a remarkably poor track record (see page 22). In part, at least, this failure may be attributed to exactly the kind of group level analysis that Mancur Olson warned against within the social sciences (Olson, 1965; see page 45 above). From the point of view of management, it may appear that all employees share the same goal, that of the company. That may well be the case to the extend that this goal coincides with personal goals, but it doesn't follow from this that the employees would take an active interest in furthering knowledge sharing. Just as it doesn't follow from the fact that everyone may benefit from a lighthouse that everyone is eager to support its construction. Certainly, if the employees perceive - as they did in Orlikowski's study (see page 22) - any sort of internal competition (if their relationship is more zero-sum) suboptimisation problems are even greater. As if this wasn't enough, in a business with rapid development and frequent job changes (the 'flexible workplace' that is often regarded as somehow connected to knowledge-sharing) we should not be surprised if reciprocity breaks down entirely since the shadow of the future is short or non-existent.

Looking beyond game theory, radical changes in communication patterns often challenge existing power structures. At a very basic level one might imagine a hypothetical employee whose main expertise is his experience with an arcane software product. He may not approve whole-heartedly of plans to replace the old software with new products – even if these feature 'objective' improvements in both usability and knowledge sharing.

The discussion above does not in any strict sense test the hypothesis that a game theoretical framework is the most enlightening of all. But it does indicate that a range of diverse observations may be tied together by such a perspective. Is this surprising? If we already knew that the theory seemed able to shed light upon the logic of traditional cooperation, was there any reason in the first place to doubt that it would also be applicable to online phenomena? It certainly *has* been argued, from various positions, that mediated interaction is different from physical interaction. There are also certain patterns in computer-mediated communication that seem to suggest that some things, at least, cannot be taken for granted when the interaction does not entail actual physical proximity. Such differences are discussed in the following.

Imagined differences: The twin myths of CMC

A discursive battle has been fought over online communication. At extremes, explanations have been framed in the ideological rhetoric of emancipatory modernism and anti-modernist conservatism (Smith, 2001b). While these ideological skirmishes are not interesting in themselves, a critical analysis of the two myths enables us to discard certain common-sense assumptions about the ways in which CMC differs from traditional interaction. Furthermore, such an analysis sheds light upon notions that still shape public debate and influence theory and research (although thankfully it seems that their influence is dwindling in the face of sobering empirical results).

The myths may be thought of as two radically different ways of looking at online interaction. The object is the same, but where one side sees *liberation,* the other sees *alienation* (in various forms these polar positions are discussed by Parks & Floyd, 1996; Rheingold, 1993/2000; Wellman & Gulia, 1999).

The myth of liberation

This myth is as follows: Man's freedom and capacity for reason is limited by physical and geographical boundaries. We are shackled by tradition and influenced by superficial evaluations that result in prejudice, discrimination, and inhibition, thus distorting free dialogue. Could we but transcend these limits conflict would be a thing of the past. Leaving our bodies behind to interact by pure text makes this possible.

Not surprisingly, no one theorist represents this myth in its extreme form[28]. Social psychological approaches focusing on identity in a feminist and/or postmodern

[28] Utopian reflections on Internet implications for the democratic process, rather than personal freedom, are the more common.

perspective, however, have come close. Feminist cybertheorist Sadie Plant would have it that:

> "The Internet promises women a network of lines on which to chatter, natter, work and play; virtuality brings a fluidity to identities which once had to be fixed… Complex systems and virtual worlds… [are important] because of the extent to which they undermine both the world-view and the material reality of two thousand years of patriarchal control." (Plant, 1996:170).

Sherry Turkle, whose *Life on the Screen* (1995) has caused much discussion, presents a variation of this argument, claiming that the human condition has become more 'postmodern' and that identity play on the Internet helps people cope with this change (an idea shared by Murray, 1997:283).

The myth of alienation

On the contrary, the myth of alienation holds that true human feeling and empathy builds on physical co-presence. New technologies, and media in particular, remove us from this original condition, thus at best providing placebo interaction. People interacting online may be, or become, antisocial creatures and what they construct can never compete with "real" community.

Proponents of such views often subscribe to a sharp distinction between the real and the virtual. Thus, online communities can only be illusions of community. Michael Heim sums up the position:

> "Technology increasingly eliminates direct human interdependence… Even video conferencing adds only a simulation of face-to-face meeting, only a representation or an appearance of real meeting. The living, nonrepresentable face is the primal source of responsibility, the direct, warm link between private bodies. Without directly meeting others physically, our ethics languishes. Face-to-face communication, the fleshly bond between people, supports a longterm warmth and loyalty, a sense of obligation for which the computer-mediated communities have not yet been tested." (Heim, 1993:14).

It is interesting to note that the idea that face-to-face communication enjoys a particularly "real" status is shared even by some of those who focus strongly upon the arbitrary nature of much that seems real (e.g. Berger & Luckmann, 1966/89:28-29).

Beyond the myths

Even before examining the shaky theoretical foundations of these views, we may be fairly certain that they are both wrong. CMC, as I have shown, is not conflict-free social bliss where reason reigns. Neither is it a condition haunted by rampant disloyalty or lack of responsibility. Neither myth explains real, observable online interaction. The problems are in the premises.

Beyond liberation

The myth of liberation misunderstands the role and dynamics of language. It has this problem in common with a large number of software designers, making it worth going into. The mistaken idea is the classical one (McQuail, 1994:43) that language is a self-contained channel of information transfer. It is a notion that would preferably separate content from context by eliminating noise and would see meaning as an inherent property of the text itself.

The problems of such a notion are legion, and I will mention only a few. *Firstly*, language not only transfers information but is also used strategically and in efforts to establish boundaries (Donath, 1999:39). *Secondly*, language is underdetermined – there is no meaning without context (even a single word does not include its own definition and even such definitions depend on context). Furthermore one cannot 'keep out' context, since communication itself means entering into (or constructing) a shared structure (Baym, 1996:151). *Thirdly*, even when an understanding of the context is shared by communicators, communication remains the act of *encoding* (since a statement may be made in a number of ways) and *decoding* (for the same reason) (Hall, 1980). Even "simple" face-to-face one-way communication, then, entails the speaker's constructing a mental model of the listener and choosing his words on the basis of predicted interpretation while the listener applies the perceived to a similar model[29] (for a formal description see Burgoon, 1994:280-281). Anyone discarding certain terms when speaking to a child illustrates this.

Trying to minimise the influence of context may damage communication, but of equal importance is the observation that communicators merely find different ways of interpretation and in some cases attach added meaning to whatever social cues are available. Susan Herring notes that many researchers "*see in CMC a more equal access to information, empowering those who might otherwise be denied such information, and leading ultimately to a greater democratization of society*" (Herring, 1996:476). Her analysis of actual online interaction, however, leads to more sobering conclusions:

> "… male and female academic professionals do not participate equally in academic CMC. Rather, a small male minority dominates the discourse both in terms of amount of talk, and rhetorically, through self-promotional and adversarial strategies… Rather than being democratic, academic CMC is power-based and hierarchical." (Herring, 1996:486).

[29] The idea of a mental model should not be taken too literally. It certainly should not be thought to demand abstract conscious thought.

One of the cues that the academics in Herring's study seem to attach much meaning to is the user name. Newsgroup posts by users who "appear" female simply provoke different responses than the rest.

Furthermore, online communicators – as is apparent from the widespread use of emoticons and the like – manage to include cues and context in the form of body language even when limited to textual interaction. The body, if you will, is not left behind.

Beyond alienation

The myth of alienation misunderstands social reality. Whereas the body cannot be kept out of cyberspace, the world outside is not without virtual elements. Physical interaction is not fundamentally more 'real' than online interaction (Smith, 2001c). Any conversation – even if conducted face-to-face – is wrapped in the social reality of structure, which, at one level, is arbitrary. This social reality is the very context also ignored by the myth of liberation – that which frames the communication, thus defining its meaning. Metaphors, categories, and the very structures of discourse are neither neutral nor natural – at least not in the way imagined in the myth of alienation[30].

Such misunderstandings pave the way for a concept of community, which is both empirically and ontologically faulty. At the empirical level, computer-supported communities are critically compared to what is largely a fiction: "*Many of those fearful that virtual community is not 'real' community and that computer-supported cooperative work will create alienated workers are confusing the pastoralist myth of community and work for the present reality*." (Wellman, 1997:198). But not only are allegedly physical communities often in fact upheld at great distances by use of various technologies. Communities, like societies, do not exist as physical entities independent of human action and perception. Thus, *all* communities are in fact "virtual" (Jensen, 2000b).

The very use of the term "virtual" often seems grounded in a philosophically inconsistent boundary between the direct or physical and the mediated.

The issue of community

An issue that often travels with either myth is that of community. The concept has received such cascades of attention that I shall not dwell upon it at length (instead see Pargman, 2000:22-27&254-263; Gotved, 1999:13-19; Preece, 2000:14-22; Wellman,

[30] It is natural to categorise, and some metaphors may have biological foundations (Lakoff & Johnson, 1980) but dominant categories also reflect value choices (whether conscious or not).

1997:182-186; Kollock & Smith, 1999:16-17). However, its relevance for CMC was noted as early as 1968 by Licklider and Taylor in the article quoted on page 8. Later it was given a rudimentary definition by Howard Rheingold, who famously – if obscurely - claimed that:

> "Virtual communities are social aggregations that emerge from the Net when enough people carry on those public discussions long enough, with sufficient human feeling, to form webs of personal relationships in cyberspace." (Rheingold, 1993/2000:foreword).

To modern sociology, born in an effort to understand the effects of modernity, it has been a core issue. When Tönnies introduced his dualism of *gemeinschaft* and *gesellschaft* (Tönnies, 1887/1965) he was expressing the myth of alienation with modern urban interaction in the role of CMC. Tönnies' dichotomy is not incompatible with the analysis presented here. He is merely stating that small group size and solid community boundaries pave the way for relationships that are not instrumental in the narrow sense – pre-modern villagers lived in fine conditions for cooperation to arise and persist. However, such fertile grounds for trust came at a prize. The trade-off, of course, was one of security versus freedom. Whereas gemeinschaft is high on security (which is closely related to trust), gesellschaft is high on freedom. Thus, Tönnies' mistake was one of idealisation, not primarily one of false dichotomies.

Having a firm understanding of the two myths allows us to consider the discussion of whether community is possible online as closed (it is possible, but it might be difficult to establish as we shall see below) or consider it simply one of definition. In the latter case, we should note that the concept of community used in this thesis is a loose one that is not limited to notions of gemeinschaft.

Summary

Revolutionary perspectives on CMC have little relation to reality. While dramatic claims have held some appeal, they have proven quite unfruitful when applied to observable interaction and certainly suffer from severe theoretical inconsistencies. But this should not be taken to imply that the medium doesn't matter. It has persistently been observed how CMC is characterized by two apparently opposite phenomena: Communication problems and a high degree of helpfulness (or cooperation). To explore this apparent paradox we need to focus on real differences. This will be done under the headings of *bandwidth* and *public goods*.

Real differences: Low bandwidth and infinite public goods

Are people particularly aggressive online? Early studies found this to be the case, and researchers were not surprised. After all (it was believed) CMC eliminated social

cues and gestures, thus rendering communication crude and uniform. Such assumptions, sometimes referred to as the 'cues-filtered-out approach' (Baym, 1996:140), fit snugly into the framework of social presence theory and social context cues theory (Preece, 2000:151). It all added up to the conclusion that CMC was an impoverished social setting almost inviting aggression (this is often taken as a solid fact, e.g. Mann & Stewart, 2000:116).

In this thinking, of course, we see the contours of the twin myths – it is a 'this is very different' approach, and warning lights should go up accordingly. It should not surprise us that "*Claims that computer-mediated communication is characterized by impersonality, hostility, and nonsocial orientation… have been challenged repeatedly*." (Parks & Floyd, 1996:2). However, while the difference is not fundamental, the fact that low bandwidth does influence the efficiency of online collaboration (if only to make it 'somewhat lower') remains solid within CSCW (Brown & Duguid, 2000:226; Baecker, 1995:292). We just need to keep in mind that explanations of aggression and lack of inhibition in relatively anonymous settings do not need speculative supporting theories – it is widespread cooperation in such forums that would require an explanation.

In fact, such an explanation *is* called for. Jenny Preece, surveying the literature, notes that, "*People marvel at the unusually open, honest, and sometimes intimate nature of much online communication*". (Preece, 2000:83; see also Wellman & Gulia, 1999:175). However, Preece is guilty of a great understatement when limiting online helpfulness to open, honest, and intimate communication. And she is not alone. One fact about online behaviour appears to run so contrary to central axioms of the web economy that it is widely and intensely ignored: Not only do people not lust for professional content[31] (see page 26), they are prepared to spend considerable amounts of time and money producing content for others to use. The number of non-commercial websites provided for anyone to visit is staggering.

I agree with sociologist Peter Kollock: "*For a student of social order, what needs to be explained is not the amount of conflict but the great amount of sharing and cooperation that does occur in online communities*." (Kollock, 1999:220).

The problem of bandwidth

Bandwidth is normally defined as the maximum rate of data transfer in a given medium per time unit. As discussed above we should not be led to believe that computers filter out all social cues. However, low bandwidth does increase the

[31] The case against professional content as the driving force behind Internet penetration is made convincingly by Odlyzko, 2001.

communicator's powers of impression management (Goffman, 1959/90) – in an email message, you have more control over the signals you send than in face-to-face communication. It is easy to see how this affects trust: unfavourably. If Bob is wondering whether to trust Alice and only has her carefully constructed textual promise of trustworthiness, he is bound to be sceptical. So, how should Alice go about convincing Bob? Drawing upon concepts from biology, Donath (1999) distinguishes between 'assessment signals' and 'conventional signals'. Assessment signals are those that cannot (easily) be faked. For instance, fast and graceful swimming clearly signals that you swim rather well: "*Assessment signals are reliable, since sending an assessment signal requires that the sender possess the relevant trait*." (Donath. 1999:32). Telling someone that you are a good swimmer, however, is a conventional signal – it *might* be true.

Being able to test statements against a bodily reference is conducive to trust. Studies have shown that people are remarkably good at separating cooperators from defectors (in experimental PD studies) if given the chance of physical interaction before playing. Similarly, communication itself furthers trust[32] (Ridley, 1997:138&240; also Jensen et al., 2000:470).

Interestingly the link between bandwidth and trust is even clearer than one might dare to assume. In one study, subjects were faced with a series of PDs (Jensen et al., 2000). Each player dyad could either not communicate, had access to text chat, had access to a system that would read text chat out loud, or had access to voice communication.

The difference in collective pay-off was substantial (and statistically significant). Voice communication was the best inducer of trust[33]. Following up on player perception, the authors found that players

> "had a more positive image (likable and trustworthy) of those with whom they could communicate. In addition, people felt that their partners were more intelligent when they could communicate with them by voice." (Jensen et al., 2000:475).

In fact, we're seeing that the amount of what would, in terms of old-fashioned communication models, be labelled as 'noise' is proportional to the level of trust[34]

[32] It is not the act of communicating that furthers trust, of course, but the ability to establish that the other party or parties seem trustworthy by whatever cues are deemed important.

[33] Notably the difference between no communication and text chat was not statistically significant. Furthermore, text that was read aloud by the computer paved the way for better results than pure text. The latter observation calls for complementary explanations that are not given by the authors.

[34] I use 'noise' in a broad sense to include context, not as a technical label of dis-information. Some classical models would only label as 'noise' distortions that are introduced to the message when in transit. Such noise, of course, does not increase trustworthiness.

(see also Lewicki & Bunker, 1996:121). The less precision there is to the message, the more information it may convey about the transmitter.

Can this knowledge be of use in relation to systems architecture? Yes and no. Increasing bandwidth may not be cost-efficient and may entail other disadvantages. Synchronous voice communication brings back the coordination problems that asynchronous email solves. Furthermore, voice messages are harder to scan, index and search than text. However, the ideal of noise reduction is often upheld for the wrong reasons – for instance to allow a focus on 'content' over 'form'. From the perspective of Alice, trying to convince Bob that she is trustworthy, we should allow the use of cues such as portraits and elaborate signatures. In her signature, she may refer to any institution that has somehow vouched for her. For instance, Alice may be a PhD scholar and include a reference to her title and institution (for the exact same reason that people sometimes include their professional or academic title when signing letters). Preferably, of course, Alice would include a link to her page at the university website. In such a case, the issue of trustworthiness is lifted from Alice to the institution. The situation is comparable to financial transactions between individuals, where the bank system may assume the responsibility of vouching for either party. Such measures can be thought of as external verification – verification is obtained from outside the system.

We may, however, be able to devise a more sophisticated system of verification by focusing on assessment signals. Or rather, we can copy it, for indeed the problem has been solved. To see this, we need to forget about people. Imagine instead what the web looks like to a search engine, a system that surfs and indexes the web to be able to display search results in an algorithmic manner. Any page you encounter may be presented to you in an honest way – without inappropriate keywords or strategic content. However, it may also be written in a way that exploits knowledge of your particular algorithm, that is it may "tell" you that it is far more important than it really is (a common strategy employed to increase revenue from advertising). For years, this problem forced search engines to keep their algorithms secret and inconstant, making life difficult for honest programmers, thus reducing the value of the indexes. The solution, discovered by the search engine *Google*, was to create a non-exploitable algorithm: Don't listen to what page X tells you. Instead ask page Y and Z what they think of X and weigh their answers against their importance

(measured in the same way). The importance of page X is then a function of the number (and weight) of links to the page from other pages[35].

In human terms this means gossip, or rendering visible the social relations within a system. One method is to let users explicitly rate each other – or the help given by others. Instead of expensively approximating face-to-face interaction by increasing bandwidth, we would then be capitalising on the capacities of computers for complex, faithful, and fast data presentation. Instead of trying to mimic (the AI approach) we would be supporting (the HCI approach). The successful trading website E-bay employs this principle (**Figure 5**).

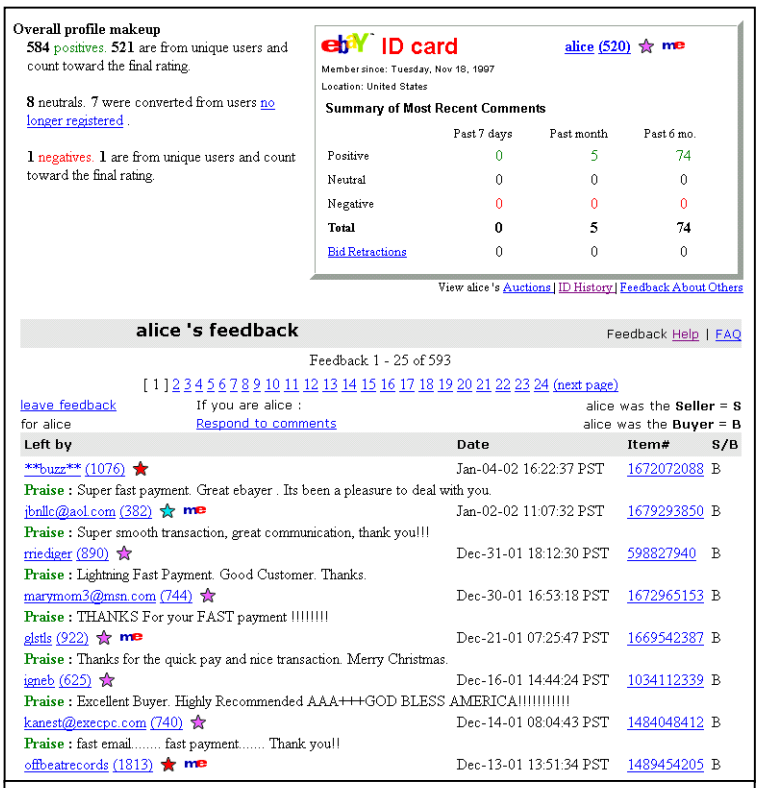

Figure 5 – "Reputation manager" at www.ebay.com.
Those who have interacted with Alice have attached their evaluation to her profile for others to scrutinize.
Without imposing technical restrictions such a system ensures identity persistence. While anyone may change their user name, only users with a bad reputation would be inclined to do so. Thus, a user with no reputation is bound to be treated with suspicion.

Another possibility is to give users access to the implicit verdicts of friends. Systems that employ buddy lists (see **Figure 3**) may let users see the verdicts of their 'buddies'. For instance, Bob is wondering whether to trust Alice, but being on the buddy list of Eve, he is entitled to the information that Alice is also on the list. Since Bob trusts Eve and Eve trusts Alice, Bob can trust Alice. Such features provide internal verification.

By considering these measures, we have in fact taken one powerful step back from the classical PD. Signals of trustworthiness – and methods of evaluation – are important to avoid playing with defectors in the first place. Whereas Tit-for-Tat gives everyone at least one chance (thus "losing" to a strategy that always defects) risk-

[35] Of course, this algorithm is not entirely unexploitable (nor is the ranking method fully disclosed) but it comes close enough to serve its purpose.

adverse E-bay users need not take such chances. The value of being able to test would-be players is a good (if only partial) explanation for the observation that "communication furthers trust" (page 59).

The obvious problem of both external and internal verification is felt upon entering the system. You have to win the trust of others to attain any privileges. And if, for some reason, no one is willing to take a chance on a newcomer you can never get into the circle of trust. Again, trust and cooperation comes at a price – both diminished personal freedom and the need to guard the community gates (see also page 40).

Digital public goods

The tragedy of the commons came about due to overexploitation by selfish individuals. But information is that strange commodity that you can give away and still possess. Thus, digital common goods differ from many analogue ones by being infinite[36]. You cannot - or at least you have no reason to - overexploit a website or a carefully written expert response on Usenet. Unless you deliberately set out to perform digital sabotage, your behaviour on a website is bound to have a truly marginal effect on the possibilities of other users.

A very important parenthesis is needed here. The fact that digital public goods are infinite does not – as we have seen – entail that people can enjoy absolute personal freedom without affecting others. Allowing ourselves some poetic licence, we might actually conclude that the tragedy of the digital commons is a reversed version of the analogue one: The scarce resource online is in fact *lack* of information. This is what Peter Kollock and Marc Smith mean when they say that "*On the Usenet, the key common resource is not an open pasture, but bandwidth… a great concern on the Usenet is using the available bandwidth wisely, which is to say, refraining from posting unnecessary information*." (Kollock & Smith, 1996:115). It is also, indirectly, an idea proposed by George Lakoff: "*One of the good things about the computer is that it enables people to write more; one of the bad things about the computer is that it enables people to write more*…" (Lakoff, 1995:123).

The concept of bandwidth applied by Kollock and Smith is not an entirely technical one – the actual possibilities for data transfer may well, for all practical purposes, be endless. Remembering *CommuniTree,* what we want is sensible use of bandwidth, or *restraint*. Irresponsible use of a newsgroup would be asking questions without taking

[36] As we have seen above, analogue public goods may be either infinite (the lighthouse type) or finite (Hardin's commons, fishing grounds, or the environment as regards pollution issues). It is the last type, which is contrasted to digital goods in the following analysis.

the time to check if they have been answered recently and without respecting whatever communication norms render possible the practical use of the system[37].

The analogy between a finite analogue resource and sensibly used bandwidth is not, I believe, entirely solid. However, the social dilemma inherent in digital forums is obvious: Individuals may feel compelled to exploit the forum, but if everyone followed this impulse, there would soon be no resource to exploit.

Now, one issue of crucial importance is still unaccounted for. Turning back to non-commercial websites sprinkled generously across the Internet we should ask ourselves how they got there. After all, why bother? Why would any individual take it upon himself to produce a public good? The infinite nature of digital public goods is mostly of theoretical importance if people have no reason to make them. But, of course, if contributing can be shown to entail some advantage for the "altruist" – for instance an improved reputation among other users – we have our answer. In other words: "*If an individual is motivated in even a small way to benefit the group as a whole, the fact that digital public goods are purely nonrival [Bob's use doesn't affect Alice] can be a significant incentive to contribute toward the public good.*" (Kollock, 1999:225).

We may imagine that what stops us from spending considerable effort on producing finite public goods in the physical world of atoms is the cost-efficiency of such an effort (it's just too expensive, and the effect is just too insignificant). If that is the case, a quantitative change in the cost-benefit ratio may have significant qualitative effects. Is the individual motivated?

Reciprocal altruism tells us yes. In an environment, which lives up to the criteria for such behaviour (basically one of persistent and distinct identity and good memory) individuals may have a reason to give, even if they are individually rational. Unlikely behaviour, however, would be giving without public recognition (fully anonymous contribution, see Kollock, 1999:233)[38]. The point may be made within a social psychological framework but the explanations may be quite easily reconciled:

> "…the process of providing support and information on the Net is a means of expressing one's identity, particularly if technical expertise or supportive behavior is perceived as an integral part of one's self-identity. Helping others can increase self-esteem, respect from others, and status attainment." (Wellman & Gulia, 1999:177).

[37] Such norms may seem entirely arbitrary but like traffic rules, they make action possible.

[38] One may note that much "anonymous" charity work by the contributor buying a label of altruism (a sticker, a pin etc.). Entirely anonymous contributions do occur, of course, but they are hardly the norm.

Few find it appealing to make achievements, which the group finds impressive without calling attention to themselves. This, coincidentally, is a problem for hackers: "*By contributing to the group, a participant can get credit and bolster his or her standing… despite the risks of getting caught, many hackers keep the same alias for long periods of time so that others will see their clever work.*" (Preece, 2000:180). However helping out may make you feel and however you may describe your motivation, your action has a very real positive effect for you. Valued contributors are "repaid" by receiving faster help and more positive reactions to requests than others (e.g. Wellman & Gulia, 1999:178).
Low cost and high effect of contribution may very well explain the great amount of contribution to the common good evident on the Internet.

Guidance by metaphor

Focusing on differences should not make us forget that some design principles that apply equally to the physical world are often ignored. Importantly, the architecture – in a more aesthetic sense – may by its very connotations guide or "script" the user.
In the physical world, courthouses often look like courthouses. Might one not instead model a courthouse on a seaside hotel? One might, but it would probably be a bad idea for two reasons. Firstly, the standard form of courthouses was most likely conceived to accommodate the workings of a court. Secondly, to a degree people are likely to behave in a court-like manner if presented with well-known courthouse cues. It isn't that one cannot act like a sea resort tourist in a courthouse – just as one is not *forced* to follow the rules of traffic – but your social sense (or 'conventionalism') will urge you not to. We can loosely describe this as the power of metaphor.
Outside the realm of architecture, this theme has been explored by social psychology (where it is a neutral conclusion of course, not a design principle) but interestingly also within the field of interactive fiction. Designers of such fiction are faced with a problem with some similarity to ours, how to make people behave in a way conducive to the narrative without hitting the rather restricting walls of limited options (Smith, 2000). Including more options, or a larger vocabulary, is rarely a cost-efficient approach. Similarly, in the design of computer-supported communities we want users to behave in a certain manner without the use of actual force and preferably without the need for greatly increased bandwidth.
But how then do creators of interactive fiction make a user feel that the experience is realistic enough to be interesting? They use the power of metaphor to limit the range of actions that the user is *likely* to make. For example, when constructing a

computer-controlled character, such as a 'chatterbot', they don't call it "Person" but "Psychotic Girlfriend"[39]:

> "A successful chatterbot author must therefore script the interactor as well as the program, must establish a dramatic framework in which the human interactor knows what kind of things to say..." (Murray, 1997:219).

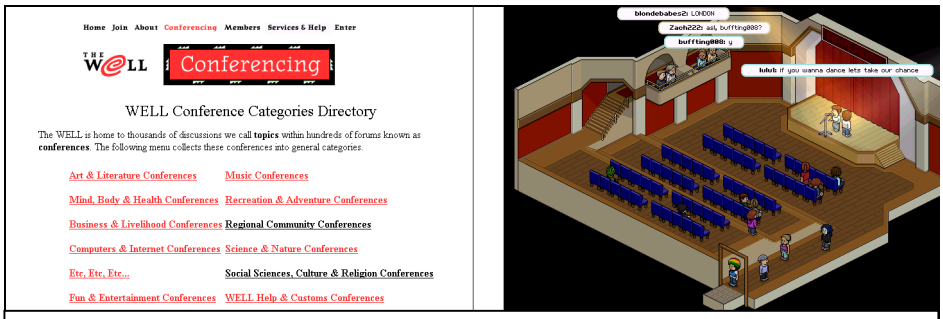

Figure 6 – *The Well* (left), a web conference system, and *Habbo Hotel* (right), a graphical chat system.
Whereas The Well (www.well.com) does very little to metaphorically guide the user, Habbo Hotel (www.habbohotel.com) makes use of special sections that may serve to cue the users (the theatre section is shown here).

In designing a forum for digital collaboration, then, we should make conscious use of metaphors that cue the desired forms of interaction (for perspectives on clear-cut roles and trust see Meyerson & Weick & Kramer, 1996:173). To say the least, this principle is not employed in many existing systems[40] (see **Figure 6**).

Summary and design principles

Computer-mediated communication is real in any meaningful sense of the word. Those who have been surprised by this may also have nurtured beliefs that the human condition changed dramatically, for better or worse, when transposed to a digital environment. Such notions are belied by accounts of actual CMC. The conclusion that theories of collective action – most particularly game theory – provide a credible explicatory framework for online interaction seems to indicate that online dynamics are not too different from those known from the physical world. Particularly, we may feel convinced that CMC is neither a modernist utopia of boundless freedom nor a cultural pessimistic nightmare of creeping alienation.

The special patterns of behaviour that do emerge in cyberspace may, at least partially, be explained with reference to bandwidth and the special status of digital

[39] Such a creature has the added advantage of being able to respond incoherently without drawing attention to its limited understanding.

[40] Probably because developers consider it 'noise'. But of course some systems are constructed for such a broad variety of purposes that any one metaphor might be inappropriate. Finally, users may be too culturally varied to make the same of a metaphor.

collective goods. While CMC enables people to draw upon the advantages of fast and inexpensive asynchronous communication, low bandwidth may be a hindrance to trust. Whereas this effect may be countered by raising bandwidth, the working of computers invites us to consider other options – particularly various ways of evaluating signals of trustworthiness. If low bandwidth (in the special way used here) is a burden to cooperation, the infinite nature of digital public goods may well have an opposite effect. Altruism comes cheaply in cyberspace.

Having considered differences, we should not overlook similarities. Careful use of metaphor may allow us to subtly cue a user as to the desired form of interaction. This may, without brutally limiting options, have the effect of synchronising interaction. By limiting the range of acceptable actions, coordination is facilitated.

What principles, then, should we employ when designing a system that inspires cooperation? First and foremost, actions must have consequences. An individual who either contributes or exploits must be either rewarded or punished. This condition requires that user identities be persistent and distinct and that individuals are likely to interact (in the broadest sense) again in the future.

Cooperating users build good reputations and thus will be highly reluctant to change their user names, but we may contribute to this effect by gradually attributing privileges (this is most obviously done in MUDs, but is similar to promotions in traditional organisations). Furthermore, we may provide support for such reputations by way of reputation managers that help fulfil the "good memory" condition for reciprocal altruism. Acknowledging that low bandwidth makes for poor "social visibility" in a system, we may employ resource-light mechanisms that illustrate social relations, such as reputation managers or buddy-lists.

As to communication, it is worth including options for self-expression. Without knowledge of the context in which a message is "decoded", communicators are forced to guess (am I talking to a child? A lawyer? Is he or she alone?). Without continuous high-bandwidth feedback, it is hard to adjust one's way of communication by judging how one is perceived. To avoid confusion as to who knows what, the system may at least present information about who have read a message (making possible the meaningful interpretation of a *lack* of response)[41]. Furthermore, features for self-representation should preferably allow the user to

[41] Like every other feature discussed, this one comes with a price, since surveillance in any form may influence action. If you cannot read a message without being seen – if responsibility comes with reading - certain work relations may make you not read it at all. A parallel issue concerns institutional memory or archives. If, say, everything in a national archive is immediately accessible to the public, much may not be archived at all. Thus, you may face the unpleasant trade-off between accessibility and integrity (see also Preece, 2000:103).

provide external verification by referring to whatever institutions have somehow "vouched" for him or her.
Finally, one may subtly cue or "script" the interactor by careful use of metaphor. By (more or less) implicitly referring to well-known "frames of conduct" such as a meeting room or a bar one may minimize confusion as to the preferred form of interaction.

These prescriptions for sensible design of digital collaborative environments all follow deductively from theory. In addition, they are well-established elements of actual systems (or environments) that are successful in this regard. What we have not established, however, is whether actual users would find such a system attractive if they had to choose among several alternatives. Theoretical advantages aside, do users of CMC agree that there is a problem? And would they agree on these prescriptions? The next chapter presents an attempt to answer these questions.

CHAPTER FOUR

Perceptions of trust

If we were to build a system that obeyed each and every design principle deduced above, could we be sure that users would want to use it? Obviously, having ventured into the real world, we would be faced with issues of marketing and software compatibility. Indeed, we would also need to tackle a plethora of sociability issues, which the poor performance of groupware has taught us not to ignore. Sure, our system might do well in experimental settings but a workplace of complex power structures is something else entirely. Basically, no – we cannot be sure.

Still, one thing we can do (without engaging in prohibitively expensive experiments) is to test whether actual users perceive the problems that previous chapters have emphasized. Furthermore, we can attempt to discover whether they in fact (claim to) employ the principles for trust management that would be expected of rational agents.

For this purpose, I have conducted a survey of online gamers who were asked to express the extent of their agreement or disagreement with a number of statements on trust and communication.

Online games and CMC

CMC serves different purposes for different people. Actual work processes are often considered more task-oriented than recreational interaction but this needs hardly always be the case. Discussion on a corporate intranet may certainly be far less narrowly goal-seeking than attempts to find suitable opponents for an online strategy game or teaming up to slay a rampaging dragon in a MMORPG. Online gaming itself – including both social MUDs and online Poker – spans all relevant spectrums. As regards design, then, it is meaningless to distinguish clearly between games and work; the dynamics are much the same.

Hence, I am not arguing that online gamers are in any (relevant) way special. The respondents are recruited in their capacity of CMC users, not because they happen to enjoy slaying digital monstrosities.

Methodology: do people know what they are doing?

Within media studies, it has proven most enlightening to focus on actual readers, viewers or users (McQuail, 1994:297-298). While maintaining a perspective on "micro" issues has perhaps resulted in a certain blindness to structure and large-scale phenomena, it has indisputably also served to eliminate simplistic notions of media effects and actual media use.

But can users actually be expected to respond truthfully to enquiries? There are two reasons why this may not be the case. Firstly, much knowledge cannot be verbalised at all. While we can ask someone *if* he or she can ride a bicycle, we cannot ask someone *how* he or she rides a bicycle[42]. Secondly, self-presentation may not be entirely truthful – some behavioural aspects may be censored (or just unknown) even to the respondent himself. For instance, media use perceived as 'vulgar' or just low cultural is tendentiously downplayed in favour of socially valued behaviour (e.g. Lewis, 1991:53).

A survey has the advantage of being low on social pressure. Only to the extent that social norms are truly internalized will they have a bearing on an entirely anonymous questionnaire, which was probably filled out in solitude. The method, however, suffers from the problem of verbalisation, which is particularly acute in the present case.

Interpreting a wide range of cues to arrive at an evaluation of trustworthiness is, at least partly, not a conscious process. Indeed, asking media users to be analytic about what is most likely a semi-conscious activity (such as using a website) is, at best, a risky approach. For this reason, Jakob Nielsen, whose catchy principles of usability (presented at www.useit.com) have practically defined the way usability studies are done, put the 'first rule of usability' to be "*Don't listen to users*" (Nielsen, 2001). Nielsen's point, of course, is the Hobbesian one that *actions* matter, not carefully constructed statements that may or may not be true.

However, there are a number of reasons why a survey may yield useful information. First and foremost, we are, as opposed to Nielsen, not primarily interested in how the users really act when inside a system. Indeed, we do want their perception of the truth since this is what they are likely to act *upon* when choosing between various systems. Much as we may think we know what is good for the users, designing anything would be senseless if the users did not to some degree share our perception.

Self-perception also matters on another account. If users consider themselves highly tolerant and altruistic inhabitants of wonderfully anarchic experiments, framing system features in the language of surveillance may lead to unfortunate results[43].

---

[42] Which, as we have seen, reflects negatively on groupware. Tacit knowledge does not get transmitted, to the extent that it cannot be formalised as data.

[43] Most likely it is never wise to employ such rhetoric (see the brief discussion of trust and cynicism on page 49). Some systems, however, are quite explicit. The peer-to-peer file-sharing program *Limewire*, for instance, lets users define "freeloaders" as people who share X number of files or less. Freeloaders are not permitted to copy files from the user's system.

Finally, we may be entitled to the assumption that if respondents are biased toward imprecise self-perception, their answers will lean towards altruism (Ridley, 1997:88). That is, they will think of themselves as kinder and less eager to engage in "cynical" evaluations than they truly are. Thus, if respondents claim to be guarded in their interaction we should not expect them to be underestimating on this account.
Importantly, this last assumption only means that any significantly guarded responses are rather solid.

Practical approach

The survey was advertised, with an introductory text, at www.game-research.com between 5th of October, 2001 and 8th of January, 2002. In addition, respondents were recruited in a variety of newsgroups (for details on sampling, recruiting and analysis, see **Appendix A**). The questionnaire (**Appendix B,** questions are referred to in the following as Q1-Q21) itself was web-based and besides basic demographic questions consisted mostly of closed questions in which respondents were asked to rate statements such as "*Communication/chat with other players is an appealing part of online gaming*." Results were analysed for statistical significance within single questions (could the outcome be a coincidence?) and between questions (for instance, do respondents who value communication/chat also find that users should have persistent user names?). Significance, here, is measured at the level of $p \leq 0.05$[44].

Respondent demographics

Respondents were, not surprisingly, overwhelmingly male (91,7%)[45]. 42,5% were in their twenties, while the mean age was 24,7 (see **Figure 7**). Whereas Americans constituted the largest group (42,4%), British respondents accounted for 17,9% of all responses. For further details, please consult **Appendix C**.

[44] For a result to be judged significant there has to be a chance of 5% or less for it to occur by coincidence.

[45] I do not consider this evidence of poor representativity in itself. There are, however, those who insist on being surprised by such numbers. Among these is the influential International Digital Software Association, who claims it as a fact that almost half of all computer game players are female (IDSA, 2001:7). While it is not entirely obvious what this should mean even if it were true, there are strong indications that the figure is misleading (Egenfeldt-Nielsen & Smith, 2000:19-23&64-68).

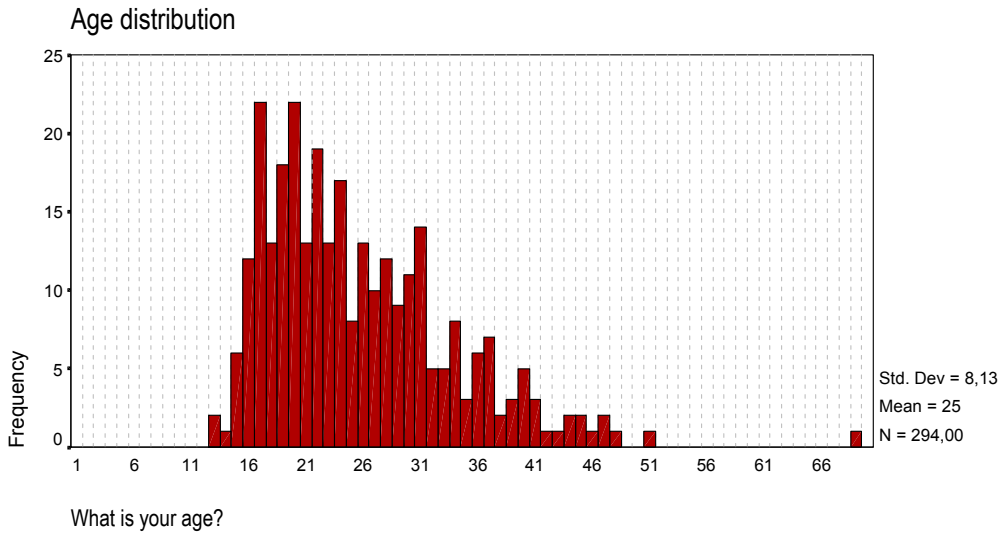

Figure 7 - Age distribution among respondents.

Is there a problem?

Online gamers, of course, are a motley crowd. Different game genres may present different problems of cooperation and as we saw earlier, different player types may have different concepts of fun (page 20). Furthermore, we may speculate that people who find online gaming worthwhile at all do not find the problems to be critical.

Figure 8, however, shows that respondents do think that saboteurs are a problem.

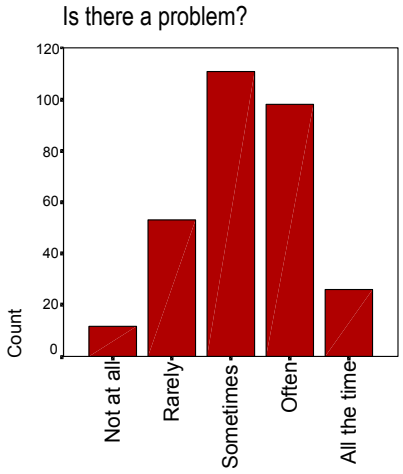

Figure 8 – Q6: To what degree do you find that online gaming is troubled by saboteurs (player killers, cheaters etc.)? (n=300)

Even if we consider the middle category "sometimes" as a statement of neutrality towards the issue (as is done throughout the following), a significant number (41,4%) reply that saboteurs are a problem "often" or "all the time".

Whereas traditional MMORPG player killers are likely to be the biggest issue, various sorts of tricksters may also plague strategy and action games. In such games, the very objective is often player-killing (and in team-based games the cooperating players may be treated as a unit since their interests are most often entirely coincident). Whereas there is no obvious way of cheating when these games have begun, there are a variety of gambler tricks that may tip the scales when discussing rules and set-up. Such start-up

chat-sessions might well be fertile ground for the study of signal assessment through textual interaction (for an example, see **Appendix D**).

Judging trustworthiness

Two strangers meet in an online gaming system. How do they determine whether to trust one another? Put slightly differently: What signals do they judge important? Here, we may be faced with obvious problems of verbalisation – it is likely that people just don't give it much conscious thought. Equally likely is the hypothesis that when asked to consider the issue, answers will lean towards ideals. Respondents may interpret the question "*How do you evaluate the trustworthiness of other players?*" as "*How should trustworthiness be evaluated?*".

Nevertheless, some very thought-provoking (and statistically significant) results were obtained from the questions asked on this account. First of all, user names are not uniformly taken as a statement of user trustworthiness (Q16)[46]. It might have been that particularly silly or vulgar user names were judged to signify low age or low respect for social norms but this seems not to be the case. If anything, a great number of respondents (24,4%) were adamant that this happened "not at all".

On the other hand, writing skills and apparent level of education is considered an important indicator (Q17). It might well be that paying attention to grammar and wording in general comes across as a commitment to the interaction. A communicator who is willing to spend time and effort on an exchange is likely to be serious about future commitment. It also means, of course, that good communicators (people who are used to textual interaction) have clear advantages when self-representation consists only of text.

Whereas form is important, actual statements and choice of subject matter appears to be even more crucial (Q21). Disregarding those who answer "sometimes" (29,5% of all) 81,4% of the remaining group claim to judge others on the basis of dialogue "often" or "all the time". This is hardly surprising. Value statements go to the heart of trust, and it would be strange not to take stock of extreme statements of egoism or altruism (although in some settings, one might be sceptical of the last sort).

[46] Here, as in the following, we are seeing only that there is no significant agreement on the subject. Some respondents *do* consider user names a good signal of trustworthiness.

Communication aside, how do users generally treat each other? In the light of chapter two, this may be rephrased in the language of strategies. What, then, is the strategy employed by the respondents? More than anything, it seems to be Tit-for-Tat.

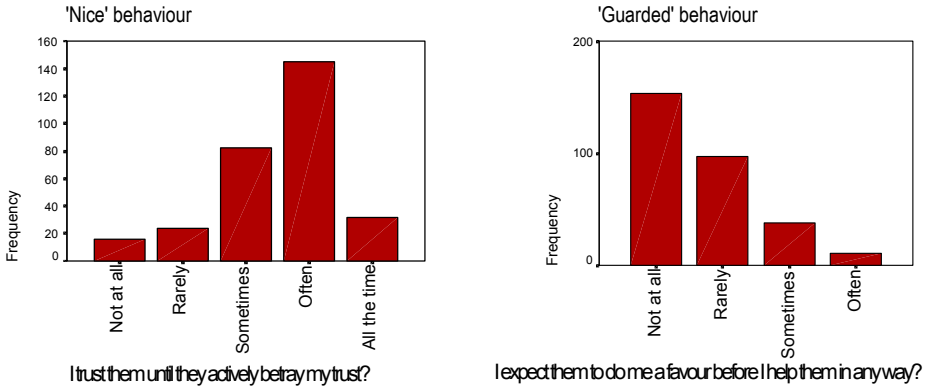

Figure 9 – Q19 & Q18: How respondents meet other users.
Most (58,6%) of all respondents trust until they are betrayed ("often" or "all the time). Even greater numbers claim not to demand favours before trusting (82,8% claim that this happens "not at all" or "rarely").
Note: No-one claimed guarded behaviour "all the time". Y-axis intervals are not identical.

As shown in **Figure 9**, respondents lean towards 'nice' behaviour and even more distinctly, they shy away from 'guarded' behaviour. If we assume that agreeing with "*I trust them until they actively betray my trust*" means that betrayal leads to broken trust and imminent defection, we have a rather clear allegiance to Tit-for-Tat (although we haven't measured the level of forgiveness; by not holding grudges real Tit-for-Tat "forgives" a "repenting" other strategy).

Of course, we should make sure that respondents do not contradict themselves. Theoretically, anyone taking the survey might claim both 'nice' and 'guarded' behaviour – or rather, correlations between being nice and not being guarded might be weak. For purposes of clarity, the twenty-five possible response categories (rarely-often, sometimes-rarely etc.) may be cut down to nine by grouping positive and negative answers. A crosstabulation of these is shown in **Figure 10**. While there are variations on the Tit-for-Tat theme, it is notable that the largest single category (155 respondents = 51,8%) is manifestly 'nice' while only one respondent (0,3%) is manifestly 'guarded'.

While suggestive (and highly compatible with previous chapters), the popularity of Tit-for-Tat style behaviour should be interpreted with caution. As noted above

statements of personal altruism may be idealisations or strategic projections (it would normally be quite irrational to say "I'm selfish", even if one considers it to be true)[47].

I expect them to do me a favour before I help them in any way? I trust them until they actively betray my trust? Crosstabulation

Count

		I trust them until they actively betray my trust?			
		No	Sometimes	Yes	Total
I expect them to do me a favour before I help them in any way?	No	29	66	155	250
	Sometimes	10	13	15	38
	Yes	1	3	7	11
Total		40	82	177	299

Figure 10 – Crosstabulation of behavioural questions with grouped results (Q18 and Q19). Please note that while these results are obviously anything but random, they are not significant according to the principles of **Appendix A**. It seems that those who answer 'yes' to Q18 also have a somewhat paradoxical tendency to answer 'yes' to Q19.

Design preferences

Cutting more directly to the issue of design elements, respondents were asked to rate some of the ideas summarised in **Chapter three**. There was quite striking agreement on the issues of internal verification and the importance of persistent identities – both of which we have seen to be necessary conditions for reciprocal altruism. On the issue of reputation managers, only 23,8% felt this to be a bad idea (26,2% were undecided). The numbers for persistent identity were very similar. Thus, while close to half did not express explicit fondness for such principles, those in favour clearly outnumber those against.

As might be expected from the brief overview in chapter one, the gamers find communication features to be important ends in themselves (see **Figure 11**). One might well be led to assume that online gamers wanted excitement not talk, but this would clearly be a mistaken dichotomy.

However, these gamers are not saying that playing is merely a path to communication. Other communication channels, both cheaper and more accessible, are available. A more reasonable, and perhaps quite obvious, conclusion would be that communicative facilities are among the reasons that respondents play online and not just content themselves with single-player offline entertainment.

[47] Respondents were also asked if they judged others by reputation. The question is ambiguous. It may well be that respondents would employ such measures if system designs allowed it. Consequently, the question is ignored here.

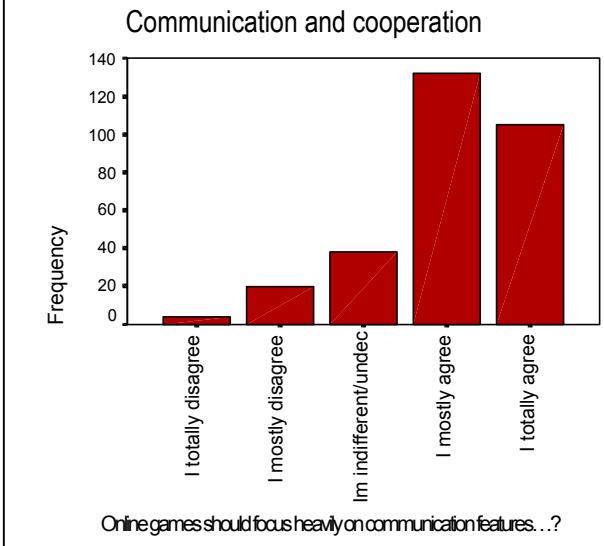

Figure 11 – Attitudes toward communication in online games.
These results are confirmed by other questions of the same type.

Outcome correlations

Examining the data more closely reveals some noteworthy relationships between outcomes. On the issue of reciprocity, respondents who value persistent identities (Q12) also claim to judge others by their writing skills (Q17) and by value statements (Q21). In addition, they feel that there should be strict limits to the number of players that are let into a game room (Q11) and that those who are let in should have to work to attain privileges (Q13). Interestingly, however, they are not particularly taken with the prospect of being able to label other users (Q14)[48].

Overall, whereas the survey revealed a range of clear attitudes towards specific issues, correlations between answers are not very strong. It may well be, that many respondents prefer to have few but strong techniques of trust management rather than a broad variety of options. Too many features may steal the focus from the game itself. This trade-off between trust-supportive techniques and simplicity is worth remembering when making actual design decisions.

Summary

Without assuming individual rationality and a rather sophisticated (if perhaps subconscious) understanding of social dynamics we had no reason to expect that online gamers would find the deduced design principles appealing. But not only do a significant number subscribe to the notion that online games are plagued by cooperation problems. They also claim to exercise Tit-for-Tat style behaviour and to desire features that support reciprocal altruism. Additionally it was found – or confirmed – that users are highly positive towards extended communication features and do not consider such activities to be superfluous or mere necessary evils.

[48] Please note what this means. A great number of those who prefer persistent identities do like the idea of labelling, but the correlation is not significant (as described in **Appendix A**).

Although the prime objective may be to play, communicating with other users cannot be considered a distinct issue – much would seem to indicate that it is part of the reason why people find pleasure in work and play.

CHAPTER FIVE
Conclusions and new perspectives

This thesis began by asking how people behave in online settings. A brief account of historical experience was given, leading to the initial summary that people, more than anything, seek other people, often contrary to the assumption of systems designers. Another assumption has taken a severe beating. Not only did CMC users not lust primarily for professional content, they were not all angels of altruism either. Nothing in the transition to Cyberspace eliminated the potential for social conflict rising from clashing interests.

Such conflicts – and ways to counter them – were examined in the light of social contract theory supplemented by the radically different analysis of Adam Smith. In the framework of game theory, it became clear that Smith had hit upon the logic of non-zero-sum games in which even selfish agents may be expected to cooperate. More specifically, the success of Tit-for-Tat in Robert Axelrod's tournaments emphasised how Hobbes' view had been unnecessarily gloomy. Criteria for rigorous duty ethics are not met, of course, but from a utilitarian, external perspective, good results may rise from careful attention to design principles.

The generalised logic of cooperation was applied to CMC, which served to debunk myths stating that online interaction is a source of either liberation or alienation. Rather, social life remains much the same online although differences can meaningfully be approached under the headings of bandwidth and digital public goods. The analysis led to design principles that supplied the necessary conditions for reciprocity – in the simple two-player sense of Tit-for-Tat or in the more complex variety known as reciprocal altruism.

Finally, self-perception and design preferences of online gamers were studied in the form of a survey. Results indicate a general allegiance to 'trusting' reciprocity and a predilection for the proposed principles of design.

Turning explicitly to the original questions, we can conclude the following. Cooperation is (or may be seen as) a situation where two or more individuals work together despite a short-term incentive to free-ride. If these individuals are just slightly selfish, the situation depends on reconciling individual and collective interests. Stated in such broad terms computer-mediated interaction does not call for a specific approach to cooperation. However, various differences between online and offline communication are observable and should be kept in mind.

From this vantage point, we may be able to affect the dynamics of online interaction within particular systems. Primarily we can see the need for features that support and encourage reciprocal altruism. These may take various forms, but can generally be thought of as features, which ensure that actions have consequences.

Perspectives and suggestions for further research

Having summed up the argument, the time has come to confess a transgression. I'll even point directly to the scene of the crime. On page 58, I used a phrase which I had decided to meticulously avoid; I hinted that there might be 'fundamental' differences between CMC and physical interaction. While I did not press the point, such a formulation points to exactly the sort of faulty reasoning that has plagued certain high-flying theories about the impact of computers on human life. But that's not all. I think the point can be made that the misguided search for fundamental characteristics continues to raise dense smokescreens within a disheartingly broad variety of the subfields in media studies (and probably beyond). The fault, of course, lies with Aristotle. Expanding upon Plato's realm of original ideas, Aristotle argued that objects in the world were separated by essential features – that there was an objective, indeed God-given, method of categorisation (Dennett, 1995:35-36).

Now, categorisation reduces complexity and is a necessity of our sensory apparatus as well as any analytical endeavour, but categorisation isn't right or true in any external sense. Believing otherwise means committing an essentialist fallacy. Having started down that dangerous road we will be on the lookout for those features that set our object of study *fundamentally* apart from others. Examples of such an approach are all around. Computers, it is sometimes argued, are fundamentally binary, have somehow absorbed their original military purposes as a fundamental feature, are automatons, are tools for communication. The hyperlink is routinely claimed to be the very "nature" of the Internet. The nature of e-mail is often said to be informal, semi-professional communication, and even the original non-graphical form of e-mail is sometimes described as inherently right (as opposed to e-mail with pictures etc.). Within interactive fiction, much effort has been put into examining whether games and narratives are fundamentally incompatible (Smith, 2000). And within film theory, arguments were championed that the essence of film was editing, was darkened (womb-like) rooms etc. (Andrew, 1976:11pp).

Unfortunately, calling attention to the fallacy is not always enough to ensure that it is avoided. Quite often, it seems to be committed for strategic reasons. Separating any subject as a worthy field of study often requires arguing that it is somehow distinct. Isn't film just theatre? Or perhaps it can just be studied as literature? The same could

be asked of interactive fiction. Durkheim, correctly (from a strategic point of view) sensing the need to justify studying social phenomena in ways not covered by psychology, philosophy, or biology, was probably well aware of the fallacy. Nevertheless, he argued vehemently that sociology was something fundamentally different from these known fields (he may also have been too optimistic about the objectivity of sociological classifications; Durkheim 1895/2000:77). His strategy worked, but it may certainly have erected false oppositions between sociology and other disciplines.

Whatever reason (or mix of reasons) CMC theorists may have to commit the same fallacy, it leads to an unfortunate focus on whatever feature is thought to make computers essentially different. And that means looking for revolution in all the wrong places.

But if CMC is somehow very important for human life, shouldn't we be able to see it clearly? No, and the very best argument is that Johannes Gutenberg died a relatively poor man (Hanebutt-Benz, 2000). The printing press is often seen as an agent of huge social change. If any event in cultural history deserves the label of revolution, it is arguably the move to a culture of writing. But it is doubtful that any contemporary thinker was able to see (or at least, to prove) that a revolution was taking place. Similarly, at one point in history an individual was born who would be the mother of the entire human race. We may not be sure as to when and where she lived, but the argument that she did is a logical one springing from the basic concepts of evolution[49]. She may have had several sisters. Indeed, it is likely that nothing about her made it possible for contemporary thinkers (if such there were) to single her out as special. Her greatness is entirely retrospective. The revolution was invisible and silent.

Without explicitly criticising the essentialist approach, Robert Kling (1996) informatively compares computers to cars. When introduced, cars were thought of as 'clean' as opposed to horses. Today, the case can be made that widespread use of cars have shaped modern cities and been important to the rise of suburbs. Since cars run on gas, they may even indirectly be important to explaining political processes in the Middle East. In the 1920s, however,

> "…it was difficult to predict what impacts the automobile would have simply by studying the uses to which automobiles were put, or by examining the major forms of social organization. If one had wanted to think ahead about what benefits cars would have, and what social dilemmas certain

[49] This woman is often known as Mitochondrial Eve since the mitochondria in all existing human cells are direct descendants of the mitochondria in Eve's cells (Dennett, 1995:97-99).

> styles of social organization around cars could produce, one would have had to be unusually prescient…" (Kling, 1996:17).

Throughout this thesis, I have drawn upon Darwin's theory of natural selection as an example of various points. Darwin, in fact, made a forceful attack on essentialism within his own field:

> "Naturalists try to arrange the species, genera, and families in each class, on what is called the Natural System... The ingenuity and utility of this system are indisputable. But many naturalists think that something more [than classification] is meant by the Natural System; they believe that it reveals the plan of the Creator…" (Darwin, 1859/1985:398).

But, of course, Darwin did more than demolish biological essentialism. Now, few issues have raised as much controversy – and indeed fanatism on both sides - as genetics and the possible merging of evolutionary biology and the social sciences. This is one of the reasons why some shy away from this attempt. But it is a pragmatic reason, and not one well-grounded in scientific conclusions. The debate centers on the question of if – or to what extent – evolutionary reasoning can explain human behaviour (for a thorough introduction to the sociobiology debate, see Segerstråle, 2000). Durkheim had reasons – both strategic and sober – to distance himself from the biology of his time. After all, although Darwin saw where his speculations were leading, he had not presented a theory, which would necessarily apply to human behaviour. Since he had an incomplete concept of heredity – he had no 'gene' – there was an obvious missing link (if you will) in any attempt to apply natural selection to the domain of human behaviour. But whereas this was true in Durkheim's time, it has changed as Darwin's theory has been continuously expanded and solidified. Indeed, for full-fledged sociobiologists it may seem that the burden of proof lies with anyone who would claim that natural selection is *not* relevant to animal (including human) behaviour at any level. If even the very basics of Darwin's theory is accepted (and surely, that is almost always the case) claims that human behaviour is somehow exempt from the principle that applies to every other feature of the biosphere smacks unpleasantly of medieval pre-materialism. It is an incoherent Darwinism-from-the-neck-down. On the other hand, there is the argument that human culture – empirically capable of overriding, or countering many biological 'imperatives' - is the one thing that sets us truly (fundamentally?) apart. But without further foundations (which may certainly be supplied) the culture argument is not noticeably more forceful than its predecessors.

What should at least be evident is that no scientific field is perfect. Corrections will have to be made; some paradigms will even have to fall. But unless one can point out actual mistakes it seems unconstructive to insist on watertight academic compartments in the study of social behaviour. It seems most unlikely that sociology,

anthropology and social psychology (not to mention communication theory and cognitive psychology) – to the large extent that they are compatible with empirical data – are in danger of anything but solidification from a biologically sound theoretical foundation.

While such a project is probably best undertaken as a constant commitment to explanations that are plausible on multiple levels, other, more concrete, studies will be able to shed light on the everyday patterns of computer-supported interaction. For instance, it might well be possible to quantitatively establish an understanding of the relation between group size and activity or coherence on discussion lists. This might help establish guidelines as to how quickly new members should be admitted to ensure that the sense of "groupishness" or community is not shattered. Clearly, qualitative effort would also be highly valuable in describing actual communication patterns and the way social norms arise and are maintained.

In this thesis, I have made a modest attempt – mostly through the application of game theory - to point to areas of common interest between communication studies, sociology, and biology. Through further integration of these perspectives, I believe, it will be possible to ground observations from Human-Computer Interaction soundly in theory thus paving the way for sensible – even pleasant – systems design within CMC. Not only will it be possible to avoid the digital concrete corporate plazas feared by Judith Donath (page 14) but, through an understanding of *real* differences, it will be possible to make enlightened choices as to how online interaction should be valued and understood as compared to offline interaction. Above all, I hope to have shown that the two forms are virtually the same. No pun intended.

Literature

A note on online resources: The contents of the World Wide Web are transitory. URLs change, move, or die without notice. To compensate, all online resources cited are included on the enclosed CD-ROM in PDF format. Page numbers in the text above refer to PDF files.

- Andrew, Dudley (1976). **The Major Film Theories**. Oxford: Oxford University Press.
- Axelrod, Robert (1984). **The Evolution of Co-operation**. London: Penguin Books.
- Axelrod, Robert (1997). **The Complexity of Cooperation – Agent-Based Models of Competition and Collaboration**. New Jersey: Princeton University Press.
- Babbage, Charles (1832). **The Economy of Machinery and Manufactures**. http://www.socsci.mcmaster.ca/~econ/ugcm/3ll3/babbage/.
- Baecker, R. M. (1995). Groupware and Computer-Supported Cooperative Work. In: Baecker, R. M. et al. (eds.). **Readings in Human-Computer Interaction**. San Francisco: Morgan Kaufman Publishers.
- Bartle, Richard (1990). **Interactive Multi-User Computer Games**. http://ig.cs.tu-berlin.de/ld/511/Reader/www/G/mudreport.txt .
- Bartle, Richard (1996). **Hearts, Clubs, Diamonds, Spades: Players Who Suit MUDs**. http://www.brandeis.edu/pubs/jove/HTML/v1/bartle.html.
- Baym, Nancy K. (1996). The Emergence of Community in Computer-Mediated Communication. In: Jones, Steve (ed.). **Computer-Mediated Communication and Community**. London: Sage Publications.
- Bechar-Israeli, Haya (1995). FROM <Bonehead> TO <cLoNehEAd>: Nicknames, Play, and Identity on Internet Relay Chat. **Journal of Computer-Mediated Communication**, vol. 1, no. 2. http://www.ascusc.org/jcmc/vol1/issue2/bechar.html.
- Berger, Peter L. & Luckmann, Thomas (1966/89). **The Social Construction of Reality – A Treatise on the Sociology of Knowledge**. London: Anchor Books.
- Berners-Lee, Tim (2000). **Weaving the Web**. London: Texere Publishing.
- Brown, John Seely & Duguid, Paul (2000). **The Social Life of Information**. Boston: Harvard Business School.
- Burgoon, M. et al. (1994). **Human Communication**. London: Sage Publications.
- Burt, Ronald S. & Knez, Marc (1996). Trust and Third-Party Gossip. In: Kramer, Roderick M. & Tyler, Tom R. (eds.). **Trust in Organizations – Frontiers of Theory and Research**. London: Sage Publications.
- Bush, Vannevar (1945). As We May Think. **The Atlantic Monthly**, Vol. 76, no. 1:101-108. http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm.
- Curtis, Pavel (1992). **Mudding: Social Phenomena in Text-Based Virtual Realities**. Proceedings of Directions and Implications of Advanced Computing, Berkeley, California. ftp://ftp.lambda.moo.mud.org/pub/MOO/papers/DIAC92.txt.
- Damasio, Antonio, R. (1996). **Descartes' Error – Emotion, Reason and the Human Brain**. Oxford: Papermac.
- Darwin, Charles (1859/1985). **The Origin of Species**. London: Penguin Books.
- Dawkins, Richard (1976/1989). **The Selfish Gene (2nd edition)**. Oxford: Oxford University Press.
- Dawkins, Richard (1982/1999). **The Extended Phenotype – The Long Reach of the Gene**. Oxford: Oxford University Press.

- Dennett, Daniel C. (1995). **Darwin's Dangerous Idea**. London: Penguin Books.
- Dibbel, Julian (1999). **My Tiny Life**. London: Fourth Estate.
- Donath, Judith (1996). **Inhabiting the Virtual City – The design of social environments for electronic communities**. Thesis, Massachusetts Institute of Technology.
- Donath, Judith S. (1999). Identity and deception in the virtual community. In: Kollock, Peter & Smith, Marc (eds.) (1999). **Communities in Cyberspace**. New York: Routledge.
- Dreyfus, Hubert & Dreyfus, Stuart (1986). **Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer.** New York: The Free Press.
- Dunleavy, Patrick & O'Leary, Brendan (1987). **Theories of the State – The Politics of Liberal Democracy**. London: The MacMillan Press.
- Durkheim, Emile (1895/2000). **Den sociologiske metodes regler**. Copenhagen: Hans Reitzels Forlag.
- Dutton, Jim (2000). Intrigue and Instant Messaging. **Matrix News**, vol. 10, no. 9. http://www.matrix.net/publications/mn/mn1009.pdf.
- Egenfeldt-Nielsen, Simon & Smith, Jonas Heide (2000). **Den digitale Leg – om børn og computerspil**. Copenhagen: Hans Reitzels Forlag.
- Einstein, David (17-02-1998). Chat for Free: Online Software lets you talk or message on the Internet**. San Francisco Chronicle**. http://www.sfgate.com/cgi-bin/article.cgi?file=/chronicle/archive/1998/02/17/BU105835.DTL.
- Elster, Jon (1989). **Nuts and Bolts for the Social Sciences**. Cambridge: Cambridge University Press.
- Frank, Robert H. & Gilovich, Thomas & Regan, Dennis (1993). **Does Studying Economics Inhibit Cooperation?** http://www.gnu.org/philosophy/economics_frank/.
- Gates, Scott & Humes, Brian D. (1997). **Games, Information, and Politics**. Michigan: The University of Michigan Press.
- Goffman, Erving (1959/90). **The Presentation of Self in Everyday Life**. London: Penguin Books.
- Gotved, Stine (2000). **Cybersociologi – det samme på en anden måde**. Copenhagen: The University of Copenhagen (thesis).
- Graetz, J. Martin (1981). The Origin of Spacewar. **Creative Computing**. No. 8. http://www.wheels.org/spacewar/creative/SpacewarOrigin.html.
- Grudin, Jonathan (1995). Groupware and social dynamics: Eight Challenges for developers. In: Baecker, R. M. et al. (eds.). **Readings in Human-Computer Interaction**. San Francisco: Morgan Kaufman Publishers.
- Hafner, Katie & Lyon, Matthew (1996). **Where Wizards Stay up Late – The Origins of the Internet**. New York: Touchstone.
- Hall, Stuart (1980). Encoding/Decoding. In: Hall et al. (eds.). **Culture, Media, Language**. London: Routledge.
- Hamman, Robin (1997). **History of the Internet, WWW, IRC, and MUDs**. http://www.socio.demon.co.uk/history.html.
- Hanebutt-Benz, Eva-Maria (2000). **Gutenberg and Mainz**. http://www.gutenberg.de/english/zeitgum.htm.
- Hardin, Garrett (1968). The Tragedy of the Commons. **Science**, no. 162. http://dieoff.org/page95.htm.
- Heim, Michael (1993). **The Erotic Ontology of Cyberspace**. http://www.rochester.edu/College/FS/Publications/HeimErotic.html.

- Herring, Susan C. (1996). Gender and Democracy in Computer-Mediated Communication. In: Kling, Rob (ed.). **Computerization and Controversy: Value Conflicts and Social Choices**. San Diego: Academic Press.
- Heylighen, Francis (1999). **The Problem of Suboptimization**. http://pespmc1.vub.ac.be/SUBOPTIM.html.
- Hobbes, Thomas (1651/1997). **Leviathan – Or the Matter, Forme and Power of a Commonwealth Ecclesiasticall and Civil**. New York: Touchstone.
- IDSA (2001). **State of the Industry – Report 2000-2001**. http://www.idsa.com/releases/SOTI2001.pdf.
- Ito, Mizuko (1996). Virtually Embodied – The Reality of Fantasy in A Multi-User Dungeon. In: Porter, David (ed.). **Internet Culture**. London: Routledge.
- Jarvenpaa, Sirkka L. & Leidner, Dorothy E. (1998). Communication and Trust in Global Virtual Teams. **Journal of Computer-Mediated Communication**, vol. 3, no. 4. http://www.ascusc.org/jcmc/vol3/issue4/jarvenpaa.html.
- Jensen, Carlos et al. (2000). **The Effect of Communication Modality on Cooperation in Online Environments**. Paper, CHI 2000 Conference Proceedings.
- Jensen, Klaus Bruhn (2000b). **We Have Always Been Virtual**. The University of Copenhagen: Working Paper, Global Media Cultures.
- Kling, Robert (1996). Computers as Tools and Social Systems: The Car-Computer Analogy. In: Kling, Rob (ed.). **Computerization and Controversy: Value Conflicts and Social Choices**. San Diego: Academic Press.
- Kollock, Peter & Smith, Marc (1996). Managing the Virtual Commons – Cooperation and Conflict in Computer Communities. In: Herring, Susan (ed.). **Computer-Mediated Communication: Linguistic, Social, and Crosscultural Perspectives**. Amsterdam: John Benjamin.
- Kollock, Peter & Smith, Marc (1999). Communities in cyberspace. In: Kollock, Peter & Smith, Marc (eds.) (1999). **Communities in Cyberspace**. New York: Routledge.
- Kollock, Peter (1999). The economies of online cooperation – Gifts and public goods in cyberspace. In: Kollock, Peter & Smith, Marc (eds.) (1999). **Communities in Cyberspace**. New York: Routledge.
- Lakoff, George & Johnson, Mark (1980). **Metaphors We Live By**. London: The University of Chicago Press.
- Lakoff, George (1995). Body, Brain, and Communication. In: Brook, James & Boal, Iain (eds.). **Resisting the Virtual Life – The Culture and Politics of Information**. New York: City Light Books.
- Lewicki, Roy J. & Bunker, Barbara Benedict (1996). Developing and Maintaining Trust in Work Relationships. In: Tyler, Tom R. & Kramer, Roderick M. (eds). **Trust in Organizations – Frontiers of Theory and Research**. London: Sage Publications.
- Lewis, Justin (1991). **The Ideological Octopus – An Exploration of Television and Its Audience**. London: Routledge.
- Licklider, J. C. R. & Taylor, Robert W. (1968). The Computer as Communication Device. In: Mayer, Paul A. (ed.) (1999). **Computer Media and Communication – A Reader**. Oxford: Oxford University Press.
- Licklider, J. C. R. (1960). Man-Computer Symbiosis. In: Mayer, Paul A. (ed.) (1999). **Computer Media and Communication – A Reader**. Oxford: Oxford University Press.
- Locke, John (1690/1952). **The Second Treatise of Government.** Indianapolis: Bobbs-Merrill Educational Publishing.

- Mann, Chris & Stewart, Fiona (2000). **Internet Communication and Qualitative Research.** London: Sage Publications.
- Mayer, Paul (ed.) (1999). **Computer Media and Communication**. Oxford: Oxford University Press.
- Maynard Smith, J. (1976). Evolution and the theory of games. In: Clutton-Brock, T. H. & Harvey, Paul H. (1978). **Readings in Sociobiology**. San Francisco: Freeman.
- McQuail, Denis (1994). **Mass Communication Theory (3rd edition)**. London: Sage Publications.
- Meyerson, Debra & Weick, Karl E. & Kramer, Roderick M. (1996) Swift Trust in Temporary Groups. In: Tyler, Tom R. & Kramer, Roderick M. (eds). **Trust in Organizations – Frontiers of Theory and Research**. London: Sage Publications.
- Meyrowitz, Joshua (1985). **No Sense of Place – The Impact of Electronic Media on Social Behavior**. Oxford: Oxford University Press.
- Morningstar, Chip & Farmer, Randall F. (1990). **The Lessons of Lucasfilm's Habitat**. http://www.ibiblio.org/pub/academic/communications/papers/habitat/lessons.rtf.
- Murray, Janet H. (1997). **Hamlet on the Holodeck – The Future of Narrative in Cyberspace**. Cambridge: The MIT Press.
- Nielsen, Jakob (2000). **Designing Web Usability**. Indianapolis: New Riders Publishing.
- Nielsen, Jakob (2001). **First Rule of Usability? Don't Listen to Users**. http://www.useit.com/alertbox/20010805.html.
- Odlyzko, Andrew (2001). Content is not King. **First Monday**, vol. 6, no, 2. http://www.firstmonday.dk/issues/issue6_2/odlyzko/index.html.
- Olson, Mancur (1965). **The Logic of Collective Action – Public Goods and the Theory of Groups**. London: Harvard University Press.
- Orlikowski, Wanda J. (1996). Learning from Notes: Organizational Issues in Groupware Implementation. In: Kling, Rob (ed.). **Computerization and Controversy: Value Conflicts and Social Choices**. San Diego: Academic Press.
- Ostrom, Elinor (1990). **Governing the Commons – The Evolution of Institutions for Collective Action**. Cambridge: Cambridge University Press.
- Ostrom, Mary Anne (01-09-2001). Instant messengers pick up speed. **Siliconvalley.com**. http://www.siliconvalley.com/docs/news/svtop/msg090201.htm.
- Pargman, Daniel (2000). **Code Begets Community – On Social and Technical Aspects of Managing a Virtual Community**. Linköbing: Linköping Universitet (thesis).
- Parks, Malcolm & Floyd, Kory (1996). Making Friends in Cyberspace. **Journal of Computer-Mediated Communication**, vol. 1, no. 4. http://www.ascusc.org/jcmc/vol1/issue4/parks.html.
- Paulillo, John (1999). The Virtual Speech Community: Social Network and Language Variation on IRC. **Journal of Computer-Mediated Communication**, vol. 4, no. 4. http://www.ascusc.org/jcmc/vol4/issue4/paolillo.html.
- Pioch, Nicolas (1997). **A Short IRC Primer**. http://www.irchelp.org/irchelp/ircprimer.html.
- Plant, Sadie (1996). On The Matrix: Cyberfeminist Simulations. In: Shields, Rob (ed.). **Cultures of Internet: Virtual Spaces, Real Histories, Living Bodies**. London: Sage.
- Poundstone, William (1993). **Prisoner's Dilemma**. London: Anchor Books.
- Preece, Jenny (2000). **Online Communities – Designing Usability, Supporting Sociability**. New York: John Wiley & Sons, Ltd.

- Reid, Elisabeth M. (1991). **Electropolis: Communication and Community On Internet Relay Chat**. Honours Dissertation: University of Melbourne. http://www.irchelp.org/irchelp/communication-research/academic/academic-reid-e-electropolis-1991.html.
- Reid, Elizabeth M. (1999). Hierarchy and power: social control in cyberspace. In: Kollock, Peter & Smith, Marc (eds.) (1999). **Communities in Cyberspace**. New York: Routledge.
- Rheingold, Howard (1993/2000). **The Virtual Community – Homesteading on the Electronic Frontier (revised edition)**. London: The MIT Press.
- Ridley, Matt (1997). **The Origins of Virtue**. London: Penguin Books.
- Ritzer, George (1996). **Sociological Theory (Fourth Edition)**. London: McGraw-Hill.
- Rousseau, Jean-Jacques (1762). **The Social Contract - Or Principles of Political Right**. http://www.constitution.org/jjr/socon.htm.
- Segerstråle, Ullica (2000). **Defenders of the Truth – The Sociobiology Debate**. Oxford: Oxford University Press.
- Shneiderman, Ben (1998). **Designing the User Interface: Strategies for Effective Human-Computer Interaction**. Boston: Addison-Wesley.
- Simon, Herbert (1945/1997). **Administrative Behaviour – A Study of Decision-Making Processes in Administrative Organizations (4th edition)**. New York: The Free Press.
- Smith, Adam (1776/1993). **An Enquiry into the Nature and Causes of the Wealth of Nations**. Oxford: Oxford University Press.
- Smith, Anna DuVal (1999). Problems of conflict management in virtual communites. In: Kollock, Peter & Smith, Marc (eds.) (1999). **Communities in Cyberspace**. New York: Routledge.
- Smith, Christine B. & McLaughlin, Margaret L & Osborne, Kerry K. (1997). Conduct Control on Usenet. **Journal of Computer-Mediated Communication**, vol. 2, no. 4. http://www.ascusc.org/jcmc/vol2/issue4/smith.html.
- Smith, Jonas Heide (2000). **The Road not Taken - The „how"s and „why"s of Interactive Fiction**. Paper, The Department of Film and Media Studies, The University of Copenhagen. http://www.akademiskopgavebank.dk/opgaver/roadnot.pdf.
- Smith, Jonas Heide (2001). **Den umyndiggjorte bruger – Usability, fornuft og indsigt i nødvendigheden**. Paper, The Department of Film and Media Studies, The University of Copenhagen. http://www.akademiskopgavebank.dk/opgaver/Den_umyndiggjorte_bruger.pdf.
- Smith, Jonas Heide (2001b). Den sociale maskine – perspektiver på computerunderstøttede fællesskaber. **Politologiske Studier**, vol. 4, no. 4.
- Smith, Jonas Heide (2001c). **Drømmer enhjørninger om fraktalgeometri? – Virtualitet i sprogfilosofisk perspektiv**. Paper, The Department of Film and Media Studies, The University of Copenhagen. http://www.akademiskopgavebank.dk/opgaver/Fraktalgeometri.pdf.
- Smith, Marc A. (1999). Invisible crowds in cyberspace: mapping the social structure of the Usenet. In: Kollock, Peter & Smith, Marc (eds.) (1999). **Communities in Cyberspace**. New York: Routledge.
- Sokal, Alan & Bricmont, Jean (1998). **Intellectual Impostures**. London: Profile Books.
- Stone, Alluequere Rosanne (1992). Will the real body please stand up? - Boundary Stories about Virtual Cultures. In: Benedikt, Michael (ed.). **Cyberspace: First Steps**. Cambridge: The MIT Press.
- Stroup, Richard L. (2000). Free Riders and Collective Action Revisited. **The Independent Review**, Spring 2000. http://www.independent.org/tii/media/pdf/TIR44_Stroup.pdf.
- Tönnies, Ferdinand (1887/1965). Gemeinschaft and Gesellschaft. In: Parsons, Talcott & Shils, Edward (eds). **Theories of Society**. New York: The Free Press.

- Trivers, Robert L. (1971). The evolution of reciprocal altruism. In: Clutton-Brock, T. H. & Harvey, Paul H. (1978). **Readings in Sociobiology**. San Francisco: Freeman.
- Turing, Alan (1950). Computing Machinery and Intelligence. **Mind**, no. 59:433-560. http://loebner.net/Prizef/TuringArticle.html.
- Turkle, Sherry (1995). **Life on the Screen – Identity in the Age of the Internet**. London: Phoenix.
- Tyler, Tom R. & Kramer, Roderick M. (1996). Whither Trust?. In: Tyler, Tom R. & Kramer, Roderick M. (eds). **Trust in Organizations – Frontiers of Theory and Research**. London: Sage Publications.
- Wellman, Barry & Gulia, Milena (1999). Virtual communities as communities – Net surfers don't ride alone. In: Kollock, Peter & Smith, Marc (eds.) (1999). **Communities in Cyberspace**. New York: Routledge.
- Wellman, Barry (1997). An Electronic Group is Virtually a Social Network. In: Kiesler, Sara (ed.). **Culture of the Internet**. New Jersey: Lawrence Erlbaum Associates.
- Wilson, Edward O. (1975/2000). **Sociobiology – The New Synthesis**. Harvard: Harvard University Press.
- Wilson, Edward O. (1998). **Consilience – The Unity of Knowledge**. London: Abacus.
- Wright, Robert (1996). **The Moral Animal**. London: Abacus.
- Zyvicki, Todd. J. (2000). Evolutionary Psychology and the Social Sciences. **Humane Studies Review**, vol. 13, no. 1, fall 2000. http://www.humanestudiesreview.org/fall2000/0900second2.html.

Abstract på dansk

Hvorledes opfører mennesker sig i forbindelse med computerunderstøttet kommunikation og er det muligt – gennem design – at påvirke denne opførsel?

Dette spørgsmål søges besvaret på baggrund af en gennemgang af de væsentligste erfaringer fra de seneste 30 års mange – mere eller mindre organiserede – computerunderstøttede fællesskaber. Denne gennemgang munder ud i en påpegelse af at computerbrugeres begejstring ved social (ikke-opgaveorienteret) interaktion ofte undervurderes samt at de beskrevne fællesskaber synes plaget af de samme samarbejdsproblemer som kendes fra offline-interaktion.

Disse samarbejdsproblemer sættes i en idéhistorisk sammenhæng og formaliseres ved hjælp af analyseredskaber hentet i spilteorien. Dernæst kombineres beskrivelse og teori og det konkluderes at den opstillede forklaringsramme er plausibel. Således bekræftet, vurderes omfanget af *reelle* forskelle mellem online- og offline-interaktion som mere begrænset end ofte antaget om end det påpeges at *båndbredde-forhold* og *digitale offentlige goders særlige karakter* afstedkommer en særlig dynamik.

På denne baggrund formuleres en række designprincipper. Disse principper tager udgangspunkt i at samarbejde uden central magtinstans nødvendiggør at individernes handlinger har konsekvenser. Endvidere vurderes det som et væsentligt princip at et system skal understøtte udsagn om personlig troværdighed gennem 'ekstern' eller 'intern' verifikation af disse udsagn samt at målrettet brug af æstetik og metaforik kan tilskynde brugerne til at påtage sig den ønskede interaktionsform.

For at vurdere om disse teoretisk udledte designprincipper vinder gehør hos de brugere som i praksis vælger imellem systemer er en gruppe online-computerspillere blevet bedt om at tilkendegive graden af deres tilslutning til principperne. Disse brugere finder at samarbejdsproblemerne er betydelige og synes positivt indstillet overfor funktioner, der – teoretisk set – burde kunne afhjælpe disse problemer.

Slutteligt opsummeres analysen og suppleres af en kort diskussion af mulighederne for – og gevinsterne ved – øget tværfaglighed i studiet af computerunderstøttede fællesskaber.

Appendix A

Details on sampling, recruitment, and analysis

Primarily, the survey was advertised at www.game-research.com, a website dedicated to computer game research (see **Figure 12**). Respondents were also recruited from the following Usenet newsgroups:

- alt.games.ultima-online
- alt.games.asherons-call
- rec.games.computer.ultima.online
- alt.games.everquest
- microsoft.public.games.zone
- alt.games.half-life.counterstrike

The message sent to these groups follows here:

> [Sincere apologies for cross-posting]
>
> At www.game-research.com we are conducting a survey on conflicts in online gaming.
>
> We'd be most grateful if you would spend 4-5 minutes taking the survey at
> http://survey.framfab.dk/survey/on_communication_and_trust/trust_online_gaming.html
>
> The aim is to determine which problems are the most dominant - and what should be done about them.
>
> Results will be posted at our site.
>
> Thanks.
>
> Sincerely,
> Jonas Smith
> www.game-research.com
> (about Game Research: http://www.game-research.com/about.asp)

A total of 302 people took the survey between 5th of October, 2001 and 7th of January, 2002. Since users, obviously, were free not to respond, those who did may be thought of as self-selected. From the population of all online gamers, the sample of 302 actually drew themselves out. The representativity of the sample, then, is dubious making it likely that the sample has distortions. It is particularly likely that respondents are not mainstream users, since their very

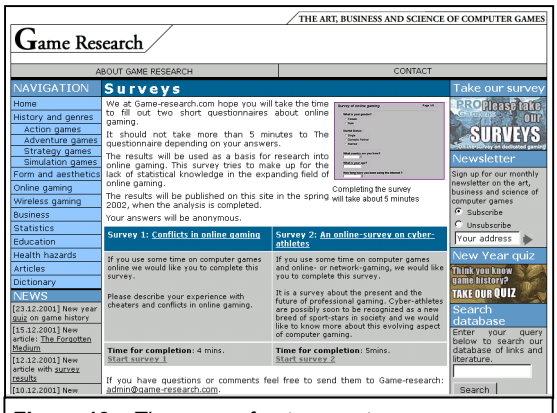

Figure 12 – The survey front page at www.game-research.com.

presence in the forums mentioned, places them firmly in the category of dedicated or 'hardcore' gamers. These respondents, however, will likely have a clear view of design features having assumably visited a variety of different systems.

Data analysis for single, ordinal questions

To test for statistical significance, single, ordinal questions (such as Q9) were submitted to the analysis described in the following.

Answers fall in five categories (1, 2, 3, 4, 5). Each of these categories has a probability (p_1, p_2, p_3, p_4, p_5) of occurring between 0 (no probability) and 1 (total probability). In addition, we can assert that their sum is 1 (we are certain that the outcome will fall within one of the five categories).

This is the model. Notably, we are not asserting that $p_1=p_2=p_3=p_4=p_5=20\%$ (that all outcomes are equally probable). Nor are we saying that $p_1<p_2$ *and* $p_4>p_5$ (that strong agreement or disagreement is less probable than mild versions).

Now we start looking for a null-hypothesis; a statement of what would convince us that the outcome distribution is not statistically significant. Within any particular question, we use the observed frequencies of outcomes. These five frequencies are divided by the total number of answers to arrive at probability percentages.

Having made no assumptions as to the actual structure of outcomes, we apply the null-hypothesis that the outcomes are symmetrical around category 3; thus, category 3 is disregarded. The null-hypothesis: $p_1=p_5$ *and* $p_2=p_4$.

Any outcome, which fails to satisfy this hypothesis, is tested for the probability that non-symmetry is coincidental. If this probability is less than 5%, the outcome is considered statistically significant.

It is important to note that the outcome category 1=90, category 2=10, category 3=20, category 4=80, category $_5$=10 would fail to satisfy the hypothesis (is it highly non-symmetrical). But such outcomes are not considered important here and will be disregarded.

Data analysis for outcome dependence

In some cases, it may be interesting to test whether answering one thing is correlated with answering another. For instance, do those respondents who answer negatively to "*I expect them to do me a favour before I help them in any way?*" answer positively to "*I trust them until they actively betray my trust?*".

To test for this, the following procedure is followed.

The five outcome categories are grouped into three: p_1 *and* $p_2 \rightarrow p_i$, $p_3 \rightarrow p_{ii}$, p_4 *and* $p_5 \rightarrow p_{iii}$. Thus, a respondent answering two questions may fall in one of nine categories. We then apply the same principles as in the test described above.

The null-hypothesis is best described by illustration. If answers in one of the nine categories are not dependent then the number of answers in the category is equal to the multiplied sums of the given row and the given column in the table below (Figure 13) divided by the total number of respondents.

The null-hypothesis: *Each probability equals (row total * column total) / total number of respondents*.

Thus, 'no dependence' would not necessarily equal any sort of even distribution within the nine categories. It merely means that the chance that a respondent who answers 'no' to one question also answers 'yes' to the other is not greater than for someone who answers 'yes' to the first to answer 'yes' to the other.

I'll not describe the actual math used in determining significance further here. The entire data set is included on the complementary CD-ROM (see **Appendix E**).

Communication/chat with other players is a necessary but not appealing part of online aming? * Communication/chat with other players is an appealing part of online gaming' Crosstabulation

Count

		Communication/chat with other players is an appealing part of online gaming?			
		No	Sometimes	Yes	Total
Communication/chat with other players is a necessary but not appealing part of online gaming?	No	3	3	215	221
	Sometimes	6	9	19	34
	Yes	12	8	25	45
Total		21	20	259	300

Figure 13 – Crosstabulation of Q8 and Q9. The outcome is statistically significant.

Appendix B

The questionnaire contained the following questions, distributed across six pages.

Survey on communication and trust in online games

1. **What is your age?** [Open]
2. **What is your gender?** [Two categories]
3. **What country are you from?** [Choice from list]
4. **How many hours do you play computer games ONLINE each week?** [Choice of interval from list]
5. **Which game do you play the most ONLINE?** [Open]
6. **To what degree do you find that online gaming is troubled by saboteurs (player killers, cheaters etc.)?** [Choice from list: *Not at all*, *rarely*, *sometimes*, *often*, *all the time*]
7. **What should be done to saboteurs?** [Choice from list: *Management should single-handedly throw them out right away, Management should single-handedly give them one or more warnings before throwing them out, Management should only take action if other players complain, Saboteurs should be ignored/left alone, Players should play a participant fee so they can be fined for transgressions*].

Please rate the following:

The following statements were all rated either *I totally disagree*, *I mostly disagree*, *I'm indifferent/undecided*, *I mostly agree* or *I totally agree.*

8. **Communication/chat with other players is a necessary but not appealing part of online gaming?**
9. **Communication/chat with other players is an appealing part of online gaming?**
10. **Management should try to let players work out their difficulties before stepping in?**
11. **There should be strict limits as to how many players are let into the same game world (or game room etc.)?**
12. **Players should be clearly connected to user names (user names should be permanent/persistent and/or hard to get)?**
13. **New players should have restricted powers within MUDs and roleplaying games until they've proven themselves in some way?**
14. **It should be possible to attach notes to other users about their reliability etc. and to make these notes available to friends/allies?**
15. **Online games should focus heavily on communication features enabling coorperation between players (pooling resources with allies, teaming up etc.)?**

How do you evaluate the trustworthiness of other players? Please rate the following statements

The following statements were all rated either *not at all*, *rarely*, *sometimes*, *often* or *all the time*.

16. I judge by the seriousness of their user names?

17. I judge them by their writing skills and apparent level of education?

18. I expect them to do me a favour before I help them in any way?

19. I trust them until they actively betray my trust?

20. I judge them by their reputation (eg. by asking others)?

21. I judge them on the basis of dialogue (value statements etc.)?

Appendix C

The respondents

No less than 91,7% of all respondents are male. Age distribution (see **Figure 7** above) shows that 74,5% are between 16 and 31 years of age with a total mean of 24,7.

A majority are either American or British with sizable groups responding from Denmark (9,6%) and Norway (5,6%) (see **Figure 14**).

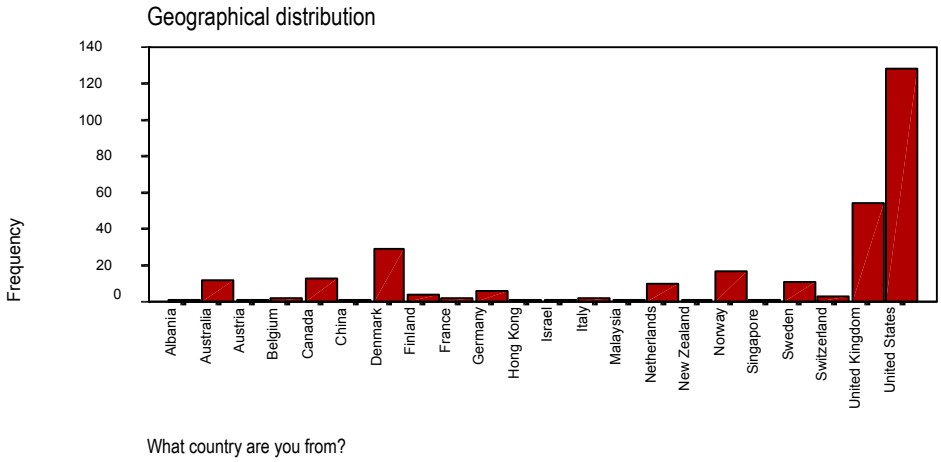

Figure 14 – Geographical distribution among respondents.

Close to half of the respondents play less than 12 hours a week, whereas less than 10% play more than 36 hours a week (see **Figure 15**). We may, however, note that

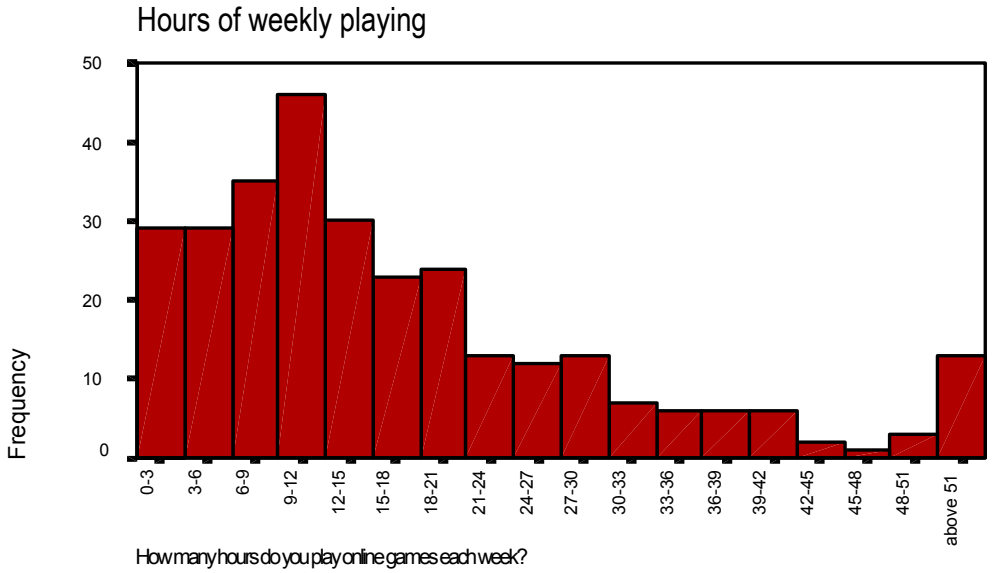

Figure 15 – Time spent by respondents on online games.

players are unlikely to have entirely accurate estimations of time spent on playing (and validity of answers may suffer from the fact that there will be different ways of categorizing game-related activities such as chatting in a game forum).
Correlation tests have not been made to determine relations between these variables and attitudes towards design principles. The data, however, was collected in a way that allows for such questions to be asked in the future.

Appendix D

Example of start-up chat session

In this example six players have gathered to play a strategy game on Microsoft's gaming forum www.zone.com.

The game is casual, and no official points are attributed to winners. Since players can just quit without loss of points, such settings often eliminate the need for careful evaluation of opponents. Nevertheless, as is seen below bobbob321 wants to know if there is "rushing" – that is, whether other players feel that early attacks are allowed. With no answer immediately forthcoming, takeshi108 concludes that bobbob321 shouldn't feel too safe ("It means that there perhaps rushing"). The player who administrates (hosts) the game, Odlakarab, agrees ("yes").

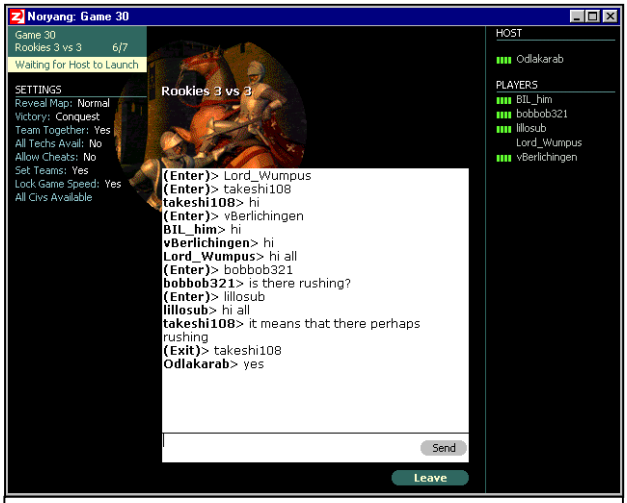

Figure 16 – Start-up chat-session before a game of *Age of Kings* on www.zone.com.

This particular system employs buddy lists and publicly visible profiles, but players cannot attach comments to each other. Neither external nor internal verification is supported, although various "spontaneous" user measures provide some compensation. Mainly, a number of users have established special groups (clans) with restricted membership. Members show their allegiance by employing systematically recognisable user names.

Appendix E

Contents of the complementary CD-ROM

All online resources are in PDF format.

File name	Description
Survey data	Complete data set (SPSS format)
Main survey frequencies	Frequency tables (PDF format)
Online resources/Babbage (1932)	Babbage, Charles (1832). **The Economy of Machinery and Manufactures**. http://www.socsci.mcmaster.ca/~econ/ugcm/3ll3/babbage/
Online resources/Bartle (1990)	Bartle, Richard (1990). **Interactive Multi-User Computer Games**. http://ig.cs.tu-berlin.de/ld/511/Reader/www/G/mudreport.txt
Online resources/Bartle (1996)	Bartle, Richard (1996). **Hearts, Clubs, Diamonds, Spades: Players Who Suit MUDs**. http://www.brandeis.edu/pubs/jove/HTML/v1/bartle.html
Online resources/Bechar-Israeli (1995)	Bechar-Israeli, Haya (1995). FROM <Bonehead> TO <cLoNehEAd>: Nicknames, Play, and Identity on Internet Relay Chat. **Journal of Computer-Mediated Communication**, vol. 1, no. 2. http://www.ascusc.org/jcmc/vol1/issue2/bechar.html
Online resources/Bush (1945)	Bush, Vannevar (1945). As We May Think. **The Atlantic Monthly**, Vol. 76, no. 1:101-108. http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm
Online resources/Curtis (1992)	Curtis, Pavel (1992). **Mudding: Social Phenomena in Text-Based Virtual Realities**. Proceedings of Directions and Implications of Advanced Computing, Berkeley, California. ftp://ftp.lambda.moo.mud.org/pub/MOO/papers/DIAC92.txt
Online resources/Dutton (2000)	Dutton, Jim (2000). Intrigue and Instant Messaging. **Matrix News**, vol. 10, no. 9. http://www.matrix.net/publications/mn/mn1009.pdf
Online resources/Einstein (1998)	Einstein, David (17-02-1998). Chat for Free: Online Software lets you talk or message on the Internet**. San Francisco Chronicle**. http://www.sfgate.com/cgi-bin/article.cgi?file=/chronicle/archive/1998/02/17/BU105835.DTL
Online resources/Frank & Thomas & Dennis (1993)	Frank, Robert H. & Gilovich, Thomas & Regan, Dennis (1993). **Does Studying Economics Inhibit Cooperation?** http://www.gnu.org/philosophy/economics_frank/
Online resources/Graetz (1981)	Graetz, J. Martin (1981). The Origin of Spacewar. **Creative Computing**. No. 8. http://www.wheels.org/spacewar/creative/SpacewarOrigin.html
Online resources/Hanebutt-Benz (2000)	Hanebutt-Benz, Eva-Maria (2000). **Gutenberg and Mainz**. http://www.gutenberg.de/english/zeitgum.htm.

Online resources/Hamman (1997)	Hamman, Robin (1997). **History of the Internet, WWW, IRC, and MUDs**. http://www.socio.demon.co.uk/history.html
Online resources/Hardin (1968)	Hardin, Garrett (1968). The Tragedy of the Commons. **Science**, no. 162. http://dieoff.org/page95.htm
Online resources/Heim (1993)	Heim, Michael (1993). **The Erotic Ontology of Cyberspace**. http://www.rochester.edu/College/FS/Publications/HeimEroticc.html
Online resources/Heylighen (1999)	Heylighen, Francis (1999). **The Problem of Suboptimization**. http://pespmc1.vub.ac.be/SUBOPTIM.html
Online resources/IDSA (2001)	IDSA (2001). **State of the Industry – Report 2000-2001**. http://www.idsa.com/releases/SOTI2001.pdf
Online resources/Jarvenpaa & Leidner (1998)	Jarvenpaa, Sirkka L. & Leidner, Dorothy E. (1998). Communication and Trust in Global Virtual Teams. **Journal of Computer-Mediated Communication**, vol. 3, no. 4. http://www.ascusc.org/jcmc/vol3/issue4/jarvenpaa.html
Online resources/Morningstar & Farmer (1990)	Morningstar, Chip & Farmer, Randall F. (1990). **The Lessons of Lucasfilm's Habitat**. http://www.ibiblio.org/pub/academic/communications/papers/habitat/lessons.rtf
Online resources/Nielsen (2001)	Nielsen, Jakob (2001). **First Rule of Usability? Don't Listen to Users**. http://www.useit.com/alertbox/20010805.html
Online resources/Odlyzko (2001)	Odlyzko, Andrew (2001). Content is not King. **First Monday**, vol. 6, no, 2. http://www.firstmonday.dk/issues/issue6_2/odlyzko/index.html
Online resources/Ostrom (2001))	Ostrom, Mary Anne (01-09-2001). Instant messengers pick up speed. **Siliconvalley.com**. http://www.siliconvalley.com/docs/news/svtop/msg090201.htm.
Online resources/Parks & Floyd (1996)	Parks, Malcolm & Floyd, Kory (1996). Making Friends in Cyberspace. **Journal of Computer-Mediated Communication**, vol. 1, no. 4. http://www.ascusc.org/jcmc/vol1/issue4/parks.html
Online resources/Paulillo (1999)	Paulillo, John (1999). The Virtual Speech Community: Social Network and Language Variation on IRC. **Journal of Computer-Mediated Communication**, vol. 4, no. 4. http://www.ascusc.org/jcmc/vol4/issue4/paolillo.html.
Online resources/Pioch (1997)	Pioch, Nicolas (1997). **A Short IRC Primer**. http://www.irchelp.org/irchelp/ircprimer.html
Online resources/Rousseau (1762)	Reid, Elisabeth M. (1991). **Electropolis: Communication and Community On Internet Relay Chat**. Honours Dissertation: University of Melbourne. http://www.irchelp.org/irchelp/communication-research/academic/academic-reid-e-electropolis-1991.html
Online resources/Smith &	Rousseau, Jean-Jacques (1762). **The Social Contract - Or**

McLaughlin & Osborne (1997)	**Principles of Political Right**. http://www.constitution.org/jjr/socon.htm
Online resources/Smith (2000)	Smith, Christine B. & McLaughlin, Margaret L & Osborne, Kerry K. (1997). Conduct Control on Usenet. **Journal of Computer-Mediated Communication**, vol. 2, no. 4. http://www.ascusc.org/jcmc/vol2/issue4/smith.html
Online resources/Smith (2001)	Smith, Jonas Heide (2001). **Den umyndiggjorte bruger – Usability, fornuft og indsigt i nødvendigheden**. Paper, The Department of Film and Media Studies, The University of Copenhagen. http://www.akademiskopgavebank.dk/opgaver/Den_umyndiggjorte_bruger.pdf
Online resources/Smith (2001c)	Smith, Jonas Heide (2001c). **Drømmer enhjørninger om fraktalgeometri? – Virtualitet i sprogfilosofisk perspektiv**. Paper, The Department of Film and Media Studies, The University of Copenhagen. http://www.akademiskopgavebank.dk/opgaver/Fraktalgeometri.pdf
Online resources/Stroup (2000)	Stroup, Richard L. (2000). Free Riders and Collective Action Revisited. **The Independent Review**, Spring 2000. http://www.independent.org/tii/media/pdf/TIR44_Stroup.pdf
Online resources/Turing (1950)	Turing, Alan (1950). Computing Machinery and Intelligence. **Mind**, no. 59:433-560. http://loebner.net/Prizef/TuringArticle.html
Online resources/Zywicki (2000)	Zyvicki, Todd. J. (2000). Evolutionary Psychology and the Social Sciences. **Humane Studies Review**, vol. 13, no. 1, fall 2000. http://www.humanestudiesreview.org/fall2000/0900second2.html