Trust in a Cryptographic Economy and Digital Security Deposits: Protocols and Policies

by

Joseph M. Reagle Jr.
B.S. Computer Science,
University of Maryland Baltimore County, 1994
Submitted to the Department of
Electrical Engineering and Computer Science
and the Technology and Policy Program
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Technology and Policy
at the
Massachusetts Institute of Technology
May 10, 1996
© 1996 Massachusetts Institute of Technology
All Rights Reserved

Signature of Author

\[\text{Department of Electrical Engineering and Computer Science} \]
\[\text{May 7, 1996} \]

Certified by

Dr. Lee W. McKnight
Thesis Supervisor, Lecturer, Technology and Policy Program

Accepted by

Professor Richard de Neufville
Chairman, Technology and Policy Program

Accepted by

Professor Frederic R. Morgenthaler
Chairman, Committee on Graduate Students

JUL 16 1996
ARCHIVES
LIBRARIES
Trust in a Cryptographic Economy and Digital Security Deposits: Protocols and Policies

by

Joseph M. Reagle Jr.

B.S. Computer Science, University of Maryland Baltimore County, 1994

Submitted to the Department of Electrical Engineering and Computer Science and the Technology and Policy Program on May 10, 1996 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Technology Policy

ABSTRACT

The richness and complexity of actions an Internet user may perform may soon match, or exceed, the capabilities of that person's interactions in the physical world. Transactions involving information retrieval and processing for medical, financial, professional, or entertainment purposes will exist upon a -- hopefully -- secure infrastructure. However, even if all underlying protocols are sound, this does not ensure that transactions in this environment are free of risk. Methods for managing the amount of risk one takes, and the amount of trust one extends to others, are still required.

Historically, formal trust relationships are represented by financial and legal instruments. A contractual obligation contingent on the recovery of a security deposit demonstrates both the "encoding" of the relationship, and the incentives for compliance with (or the lack of betrayal of) that relationship. This thesis hypothesizes that many of the contemporary instruments for dealing with trust can be implemented in digital form -- with perhaps greater efficacy.

To accomplish this, I first focus on the concept of trust: what is trust, and how is trust represented and evaluated in the real world. I then examine how trust is represented on today's Internet, and how can trust be managed in such an environment in the near future. I argue that trust will be managed in three ways, by: a) using traditional methods like credit scoring, financial instruments, and accounting for risk in the cost of the service, b) developing cryptographic protocols which limit the need for trust, and c) displacing the risk and the direction of the trust relationships from one agent to another. Furthermore, I provide an analysis of a security deposit protocol which accomplishes (c).

I conclude by focusing on policy makers' confusion of the historical instance of a financial or trust management instrument (tool) with the operational qualities of such tools. I also address how this can affect the development of efficient tools -- and consequently the market which would be dependent upon them.

Thesis Supervisor: Dr. Lee McKnight
Title: Lecturer,
Technology and Policy Program
ACKNOWLEDGEMENTS

I thank my academic advisor and thesis supervisor Dr. Lee McKnight for his guidance and support. I particularly appreciate the ability to pursue a – perhaps – atypical topic, but one for which I have a great deal of enthusiasm. Thanks to the other members and students of the Research Program on Communications Policy for providing the environment for me to pursue this work.

I also thank all of those that, through real or digital interaction, provided useful comments, inspiration, insights, time, and references to useful literature including fellow RPCP'er Joseph Bailey. Particularly, I thank Prof. Ron Rivest and fellow TPP student Brett Leida who shared some of their valuable time so I could describe a theory, enable me to better the theory, and clarify its exposition. Amongst those that I extend my digital thanks to are the members of the cypherpunks mailing list, particularly Wei Dai, and to all of those that contribute signal to the signal/noise ratio, including (but not exclusively) Timothy May, Hal Finney, Laurence Lundblade, Perry Metzger, Robert Hettinga, Carl Ellison, Robert Jueneman, Michael Froomkin and Jeff Weinstein.
# TABLE OF CONTENTS

1. INTRODUCTION ........................................................................................................ 12

2. WHAT IS TRUST? ........................................................................................................ 15
  2.1 TRUST AS TRUTH AND BELIEF ........................................................................... 16
  2.2 A THEORY OF TRUST .......................................................................................... 22
  2.3 DECISION ANALYSIS ........................................................................................... 25
    2.3.1 The Value of Credit Information ...................................................................... 25
    2.3.1.1 Expectation with No Information ............................................................... 26
    2.3.1.2 Expectation with Extended Information ..................................................... 27
    2.3.1.3 Expectation with Perfect Information ...................................................... 27
    2.3.1.4 Expectation with Sample Information ...................................................... 28
    2.3.2 Trust Algorithms as an Expression of Preference .......................................... 29
  2.4 TRUST AS COMMERCE ......................................................................................... 31
  2.5 APPENDIX A: CREDIT RATING AS A TEST ......................................................... 36

3. TRUST AND FINANCIAL INSTRUMENTS ................................................................. 37
  3.1 TRUST MANAGEMENT AND FINANCIAL INSTRUMENTS ............................... 37
  3.2 INCORPORATING RISK INTO THE COST ........................................................... 41
  3.3 CREDIT .................................................................................................................. 41
    3.3.1 Money ............................................................................................................. 43
    3.3.2 Trust and Securities ......................................................................................... 45
    3.3.3 Letters of Credit .............................................................................................. 46
    3.3.4 Digital Bearer Bonds ....................................................................................... 47
    3.3.5 Land Owners (Security Deposits) .................................................................. 48
  3.4 THE EFFICACY OF DIGITAL INSTRUMENTS ...................................................... 48

4. TRUST OBVIATION – PROTOCOLS THAT NEGATE THE NEED FOR TRUST ... 50
  4.1 TRUST AND PROTOCOLS ...................................................................................... 50
    4.1.1 Bit Commitment ............................................................................................... 51
    4.1.2 Oblivious Transfer .......................................................................................... 52
    4.1.3 Zero-Knowledge Proof Protocols ..................................................................... 53
    4.1.4 Cut and Choose & Secret Sharing .................................................................... 53
    4.1.5 Fair Coin Flips .................................................................................................. 54
    4.1.6 Contract Signing .............................................................................................. 54
    4.1.7 Other Standard Protocols ............................................................................... 55
    4.1.8 Blind Signatures and Digital Cash ................................................................. 56
  4.2 COIN RIPPING ...................................................................................................... 57

5. TRUST SHIFTING – A PROTOCOL FOR SECURITY DEPOSITS ........................ 59
  5.1 CHARACTERISTICS & SECURITY DEPOSITS ................................................... 60
  5.2 CRYPTOGRAPHIC PROTOCOLS ....................................................................... 61
    5.2.1 Security Deposits and Characteristics using RSA ........................................... 61
    5.2.2 In Depth Protocols ......................................................................................... 62
      5.2.2.1 Protocol Using Hamiltonian Cycles .......................................................... 63
      5.2.2.2 A Proper Complexity Theoretic Problem ............................................... 65
    5.2.3 Information and Cryptographic Economics .................................................... 66

6. ....................................................................................................................................... 69

6. CERTIFICATES – REPRESENTATIONS OF TRUST .............................................. 70
6.1 PUBLIC KEY CRYPTOGRAPHY AND DIGITAL SIGNATURES ........................................... 71
6.2 CERTIFICATES AND X.509/PEM ........................................................................... 72
6.3 CERTIFICATION AND TRUST STRUCTURES ...................................................... 75
  6.3.1 Identity ............................................................................................................. 75
  6.3.2 Characteristics (Attributes) ........................................................................... 77
  6.3.3 Structure – Hierarchical v. Non-Hierarchical ............................................... 82
  6.3.4 Revocation ..................................................................................................... 83
6.4 TRUST MANAGEMENT ......................................................................................... 85
6.5 APPENDIX A: ASN.1 SYNTAX FOR CERTIFICATES AND CRLS ................. 89
6.6 APPENDIX B: AN EXAMPLE ANALYSIS OF THE TRUST MODEL ................ 90

7. TECHNOLOGY POLICY – IMPLICATIONS FOR TRUST MANAGEMENT ......... 92

7.1 POLICY AND RULE MAKING ......................................................................... 94
7.2 CRYPTOGRAPHY ............................................................................................... 95
  7.3.1.1 - Digital Signatures and Contracts ............................................................. 97
  7.3.1.2 - Intellectual Property (IP) and Digital Networks .................................... 98
  7.3.1.3 - Electronic Cash, Banking, Tax Evasion, Money Laundering, and Fraud .. 99
7.4 COMMUNICATIONS ......................................................................................... 101
  7.4.1.1 Anonymity & Privacy ................................................................................. 101
  7.4.1.2 Censorship and Free Speech ..................................................................... 102

8. CONCLUSION ....................................................................................................... 103

9. REFERENCES CITED .......................................................................................... 106
LIST OF FIGURES AND TABLES

FIGURE 6-1 PUBLIC KEY EXCHANGE ................................................................. 71
FIGURE 6-2 SIMPLE HIERARCHY ................................................................. 74
FIGURE 6-3 CA HIERARCHY ......................................................................... 75
FIGURE 6-4 IDENTITY CERTIFICATE ......................................................... 78
FIGURE 6-5 GENERAL CERTIFICATE ......................................................... 79
FIGURE 6-6 POLICY MAKER REQUEST TO ASSERT MATCHING ............... 87

TABLE 2-1 CHARACTERISTICS OF TRUST EVALUATIONS ......................... 24
TABLE 6-1 BENEFITS AND PROBLEMS OF DN/ca NAMING SUBORDINATION .. 77
colophon:

software:

  text processing       Microsoft Word6.0
  bibliographic         Refs7.72

type:

  base text            11/16 pt TimesNewsRomanMS
  chapter titles       16pt Caslon540BT
  headers              14pt and 11pt ZapfHumanist777BT
  protocols            10pt ZapfHumanist777BT
  quotations            11/12 pt TimesNewsRomanMS
  title quotations      11pt SouvenirAllType
  figures and tables   11pt ArialNarrowMS
  example names         11pt CourierMS
BETTER TRUST ALL, AND BE DECEIVED,
AND WEEP THAT TRUST AND THAT DECEIVING,
THAN DOUBT ONE HEART, THAT IF BELIEVED
HAD BLESSED ONE'S LIFE WITH TRUE BELIEVING.
— Frances Anne Kemble. Faith.

As information technologies and networks become more pervasive, actions which were once limited to the confines of the physical world are migrating to the digital world. Such a statement seems overly grandiose, all too common, and in great danger of being considered hype. However, I make the statement to demonstrate a point: while it is easy to envision the extension of the real world into the digital, the task itself will be difficult. This too, would come as no surprise to many. Programming the applications, installing the infrastructure, and developing the standards and protocols — will be difficult. However, this thesis addresses the non-explicit challenges to technologists, entrepreneurs, and users of the Internet that make the above problems even harder.

The challenges of this domain are difficult not only because the solutions are non-trivial, but because the questions are not well understood. For instance, most people take for granted that cash is accepted nearly everywhere, that it grants one a form of anonymity, and that it is controlled by the government. However, the qualities of cash when translated into the digital realm loses some of its real world characteristics. It is something that is both strikingly similar and — for some — frightfully different. A wide range of anonymity similar to that inherent to cash, checks, and debit instruments will be available in

---

1 I define the "real world" as the world in which people interact with objects to effect immediate physical changes in their environment. For instance, a pen allows one to write on paper. Even things like remote controls and ATM machines are intended to cause immediate changes in ones physical surroundings (the television turns off, or one has more cash). The "digital world" is where the immediate effect of an action is the communication, processing, or storing of bits and not to effect an immediate change in the surrounding environment. These definitions are somewhat problematic, but I do not intend that they discretely partition the two worlds — I think the two worlds are merging. Rather, for the sake of general understanding, when I say "real world" I'm referring to buying a cup of coffee with cash. When I refer to the digital world, I am referring to the Internet (cyberspace, the matrix, etc.) where one can cruise the web, send email, or use ecash.
digital form. However, the control, minting, and economic characteristics of these instruments may be very different from what we know today. For instance, electronic cash (ecash) may be wholly outside of the control of any government.

Some things change, some stay the same. It is this strange characteristic of the digital world about which many of the contentious issues of today revolve – be it the right to privacy versus state and corporate interests, or free speech versus the protection of minors.

I focus on one of the assumptions that underlay many technological issues today, and one that most people take for granted: trust. In the real world, trust between people is generally established from personal relationships or legal obligations and it is not given a great deal of thought – or at least not until it is broken. In glory days, this informal relationship worked when a man was as good as his word, and commerce was conducted on the basis of a handshake. However, technology has changed the world into a global village, while increasing the distance between potential business partners. The likelihood of conducting business with a stranger (someone one has less trust in) increases. As such, the mechanism for measuring the amount of trust one should extend and the proper methods for extending it becomes more important. The Internet may provide a future where users can interact in a cryptographic economy in which digital media, smart agents, electronic payments, and pseudonyms will take the place of a firm handshake, or face-to-face business lunch. However, the importance of trust is not replaced – it becomes more explicit.

This thesis addresses the topic of “trust in a cryptographic economy.” By this, I mean how will people establish trust relationships in a market that is created from agents (customers, merchants, computers, and value added services) using information, digital media and strong cryptographic applications to conduct commerce. This thesis argues that trust instruments (tools or services for dealing with trust) will be implemented in a cryptographic economy in a combination of three ways. The first manner in dealing with trust is an extension of how it is addressed in the real world. One can mark up the price of the service so as to recoup losses occurred because of fraud or default; or one can use

Or the level of exposure to risk one takes upon one’s self, or the credit one extends to another.

I have coined the term “cryptographic economy” to represent the economy of the digital world where strong cryptography and competitive agents create a secure marketplace for information and value added services. An instance of this economy is often referred to as an electronic marketplace. Perhaps the most appropriate term would be a “digital agora” but while it captures the richness of this environment, it sounds rather awkward and does not emphasize the importance of security mechanisms to this market.

The topic of people creating and using trust relationships will hereafter be called trust management with deference to the authors of Decentralized Trust Management [BFL96].
existing financial instruments and agreements like credit card networks, and digital signatures (contracts or letters of credit) to mediate commerce between untrusting parties. The second way to deal with trust is to negate the need for it. This can be accomplished in some cases with novel cryptographic protocols. The third is to use cryptographic protocols to shift the direction and amount of trust extended in a transaction. This is similar to the way a security deposit shifts some of the risk of the landlord onto the tenant.

Before investigating how trust will be managed on digital networks, the concept of trust itself is first examined. Chapter 2 introduces the concept of trust and discusses how the concept of trust differs with respect to three disciplines. The first conceptualization is philosophical, the second is an understanding developed in the technical literature of formal models and logics, and the third is trust as expectation in the discipline of decision analysis. I then present my own understanding of trust which is the basis for the rest of the thesis.

Chapter 3 briefly analyzes how the concepts of trust and risk are dealt with in today’s economy. The relation between financial and trust instruments is explored, including credit, money, bearer bonds and security deposits are discussed. This chapter also demonstrates that there is a an incentive to extend and use these instruments in the digital realm.

Chapter 4 reviews a number of cryptographic protocols with respect to their abilities to limit or obviate the need for trust between the principals of a transaction. Protocols for bit commitment, oblivious transfer, fair coin flips, contract signing, and coin ripping are briefly explained.

Chapter 5 extends the cryptographic theme and presents an analysis of a protocol which attempts to shift risk and trust relationships by implementing a digital security deposit. Such a deposit ensures that it is in a user’s best interest not to betray the trust of another.

Chapter 6 investigates the significant technical issue of how to represent trust. The technical representation of trust is examined with respect to certificates, certification structures and policies. Examples of these issues are given, including discussions of PEM/X.509, and PGP. Recommendations are given towards establishing both sophisticated and manageable trust relationships.

Finally, Chapter 7 deals with the technology policy issues that are both affected by and affect the technical aspects spoken of early, including digital signatures, intellectual property, electronic cash, fiscal control and tax evasion. Chapter 8 concludes this thesis.
TRUST

In the province of the mind, what one believes to be true either is true, or becomes true.
— John Lilly.

The term "trust" is increasingly used by those concerned with information security and electronic commerce. The most popular domain for its usage has been research regarding authentication and the infrastructure for public key technology in a networked environment [BLNS86, YKB93, BBK94, ZH95, BBC90]. The issue of how to exchange public keys and their certifications over the Internet has been important to the creators and users of public key applications such as PGP. However, the broader, more traditional usage of the word — beyond the specifications of certification formats for public keys — has increased with the rise of electronic commerce.

Even though the term trust is used, it is rarely defined. Trust is defined by the Oxford English Dictionary as:

1. a. Confidence in or reliance on some quality or attribute of a person or thing, or the truth of a statement.
2. Confident expectation of something; hope.
3. Confidence in the ability and intention of a buyer to pay at a future time for goods supplied without present payment.

Each one of these definitions applies towards an understanding of trust that I shall present in my thesis. The first definition speaks to the common sense understanding of trust. If I trust you, I am relying upon a quality or attribute of something, or the truth of a statement. It also hints at a logical treatment that could apply towards understanding trust. The second definition includes the word "expectation" which reflects the strong mapping between the common sense understanding of trust and concepts from decision analysis. The third definition speaks to the driving force behind interest in trust: commerce. The language of markets, credit, risk, and the law may be successfully extended to the digital realm.
2.1 TRUST AS TRUTH AND BELIEF

In the relatively small but quickly growing\textsuperscript{5} amount of technical literature regarding trust, a few references are made to the significant amount of philosophical literature on the topic of trust and belief. Zurko and Hallam-Baker [ZH95] refer to modern hermeneutics (the study of knowledge descending from Heidegger's philosophies) as an insightful philosophy into the nature of trust: "A key concept in hermeneutics is the assertion that communication and interpretation are central to the concept of truth." By arguing that "much of what we regard as truth is in fact taken as trust" [ZH95] they consider a Hermeneutic philosophy to be a philosophy that parallels the distributed "webs of trust" found on the Internet.

A formalized logical approach can also be taken. Rangan developed an approach for formalizing trust by constructing a theory based on a modal logic in which "first-order predicate logics are enhanced by modal operators such as belief." [Ran88:205] The approach developed a model in which agents maintain a data-base of beliefs regarding the real world. Associated with each agent is a set of states corresponding to the real world, or his belief of the real world:

> An agent's belief arises primarily because of the agent's ignorance about the global state of the distributed system. Thus an agent's state of belief corresponds to the extent to which, based on its local state, the agent can determine what global state it is in.

[Ran88:206]

In a given global state, there are a series of possible-world states which are those worlds that are possible under an agents set of beliefs. An agent believes an assertion if that assertion is true for each of the global states that the agent thinks is possibly the real one. By defining a distributed system in terms that may be related to a Kripke structures (a formal model that relates various propositions, their truth assignments and the possibility of an agent considering one global state to be likely from another) Rangan constructs a formal theory for the evaluation of belief systems of the sort: agent(i) believes agent(j) believes some well formed formula (WFF) in the language of the logic. These WFF may be operated on or transformed by agents using rules of inference and they may be communicated to other agents. Without following Rangan's theoretical treatment and proofs to their conclusion, I would like to describe some of the interesting characteristics of the theory with regards to trust and their application towards my own understanding of trust:

\textsuperscript{5} In reviewing the technical references, it appears much of the seminal work on the topic is from the late 1980s to now! ([Ran88] and [BLN86] being some of the earliest pieces, and [BFL96] being the most recent at the time of this thesis.)
1. A believed proposition need not be true in the real world, but it must be true in all the states corresponding to an agents possible states and their possibility relations.

2. For the global states, \((s,t) \text{"p, is called the possibility relation according to the agent(i)"} [Ran88:206] iff in the global state \(s\), agent(i) can consider the global state \(t\) as possible.

3. A message containing a WFF from agent(j) to agent(i) may trigger a belief acquisition in agent(i)'s data-base if the WFF is consistent with the other messages from agent(j).

4. "In the logic, beliefs received by agent(i) will not be inconsistent with each other." [Ran88:207]. Rangan's interesting example is that of agent(j) telling agent(i) some WFF \(f\). If agent(j) then tells agent(k) \(\neg f\), and then agent(k) tells agent(i) their is no inconsistency in agent(i)'s belief data-base. Rather, he believes that agent(j) believes \(f\), and that agent(j) lied to agent(k) who believes \(\neg f\)."

5. An agent may send beliefs it does not hold, but only beliefs which are consistent with previous messages from an agent will be accepted by others.

Given these properties and other developments Rangan formally develops a theory of trust where the expression of trust is the addition of a WFF that is valid for the axioms of a logic. He is then able to specify various applications by which the chains of trust in a system may be specified formally and proven to be valid by a theorem prover.

[BAN89] provided a logic (referred to as BAN logic) for the analysis of authentication protocols. Often, such protocols are explained by listing a series of messages in the form of

\[
P \rightarrow Q : \text{message}
\]

where message may be a series of clear-text, encrypted, or signed messages that are interchanged between the principles, \(P\) and \(Q\), of the protocol. By enumerating a set of constructs such as belief, once said, sees, and jurisdiction over Burrows, et al. are able to specify a logic with which such authentication protocols can be rigorously tested.

While it does not address the characteristics or nature of trust itself, it does allow one to logically examine whether the assumptions and goals of a protocol with respect to trust are self-consistent.

Gong, Needham, and Yahalom extend BAN logic [GNY90] by providing an approach that exceeds the assumptions required for the original logic:

For example, it does not assume that redundancy is always present in encrypted messages – incorporating instead a new notion of recognizability which captures a recipient's expectation of the contents of messages he receives.

[GNY90:234]
One of the greatest differences is the added distinction one can make for what principles believe in and the information they possess. Such a scheme allows for a much more intuitive understanding of trust and belief since, "this allows us to separately treat the content of a message and the information implied by such a message. It also makes it possible to separate reasoning about the physical world from the reasoning about other principals' beliefs, so that we can consider different levels of trust in the reasoning." [GNY90:234] This difference is reflected in the set of statements for the logic: \( P \) is told, \( P \) possesses, \( P \) once conveyed, and a series of different belief statements regarding the freshness of a formula, the recognizeability of a formula, and the beliefs for the suitability of public and secret keys. As such, one can have a different level of trust for beliefs about the real world, and statements sent by other agents.

This distinction is discussed further in [YKB93:152] where the authors, "distinguish between directly trusting some entity in some respect, and directly trusting an entity in some respect due to some other entity's trust." Given this new distinction, the obvious concern is how does one traverse the network or web of trust (called a trust recommendation path) that develops in an environment in which one trusts an agent, who also can express beliefs about the trustworthiness of others, and the others may do the same? They accomplish this by presenting a trust derivation algorithm which, "generates, from a given set of [trust] expressions, a set \( T \) of all entities in which a corresponding entity, say \( A \), indirectly trusts in respect to \( x \)" [YKB93:156], where \( x \) can be one of the following functions: identification, key generation, escrow, non-interference, clock synchronization, protocol compliance, and providing information about others' trustworthiness (or reputation.)

In Valuations of Trust in Open Networks\(^7\) [BBK94] the analysis of derived trust is further extended for cases in which, "different entities offer different allegedly authentic data..." [BBK94:3] A method of resolving these differing opinions is required. A number of interesting concepts are introduced, one of which is the recording of both positive and negative experiences with other agents. I call this record a history. The concept of direct trust (trust about a direct interaction with another) and recommendation trust (one's level of trust in another as an introducer of strangers) are also defined in the following manner. A direct trust relationship\(^8\) exists if:

\(^6\) The authors also provide an interesting table of functions for which principles are often trusted to do: identification, key generation, escrow, non-interference, clock synchronization, protocol compliance, and providing information about others' trustworthiness (or reputation.)

\(^7\) If the reader is interested in this topic, I encourage them to read [BBK94] which is both interesting and accessible.

\(^8\) The interesting constraint of this analysis is that any negative experience destroys any direct trust relationship.
all experiences with Q with regard to trust class x which P knows about are positive experiences. . . V is the value of the trust relationship which is an estimation of the probability that Q behaves well when being trusted. It is based on the number of positive experiences with Q which P knows about.

[BBK94:5]

and (v) is computed as follows (where p is the number of positive experiences)

\[ v_r(p) = 1 - \alpha^p \]

this value is the probability that Q has a reliability of more than \( \alpha \), founded on the information P possesses about Q.

[BBK94:6]

If r is the actual reliability of an entity, the authors prove that the probability of r being greater than one's "threshold" of \( \alpha \) is \( 1 - \alpha^{p+1} \).

A recommendation trust relationship exists when

if P is willing to accept reports from Q about experiences with third parties with respect to trust class \( x \).

\( v \) in this case is the value of the trust relationship with Q which can exist over both positive \( (p) \) and negative \( (n) \) experiences\(^9\) with the recommended entities and is calculated as follows:

\[ v_r(p,n) = \begin{cases} 1 - \alpha^p & \text{if } p > n \\ 0 & \text{else} \end{cases} \]

As the positive experiences grow with a particular agent, \( v \) will approach 1. If the negative experiences exceed the positive over time, \( v \) will approach 0. Given a non-cyclic network with \( v \) representing the value of the trust relationships (vertices) between the agents (nodes) a derived trust value can be calculated which includes the strength of the recommendation, and how much one trusts the recommender (solid lines are a direct trust and dotted lines are recommended trust.).

\[ \text{Diagram} \]

\[ \text{Diagram} \]

\(^9\) Negative experiences with recommendees do not automatically destroy the relationship with the recommender since he may not have lied, he may have simply been wrong.
The derived values for the relationship between A and C are:

$$V_1 \Diamond V_2 = 1 - (1 - V_2)^{Vi}$$

which is equal to

$$V_1 \Diamond V_2 = 1 - (1 - \alpha^p)^{Vi}$$

A result of this analysis, which I do not show, is that multiple derived values (multiple opinions about, or paths to, one agent) can be combined to form a single, balanced, direct trust. Also, an interesting consideration with regards to the §2.6 is that the valuation is considered in light of economic value.

As mentioned earlier, we assume that the value of each task can be measured in units, e.g. in ECU which are lost when the task is performed incorrectly. Our estimations about the reliability of entities were made relative to tasks consisting of a single unit. If we wish to entrust a task consisting of T units, the trust entity has to fulfill T "atomic" tasks in order to complete the whole task. Bearing in mind, we can estimate the risk when entrusting a task to an entity.

The results of many of the previous calculation can then be combined in order to demonstrate how many economics units one is willing to risk to trust a given entity:

$$l(\alpha, \nu, T) = \int_0^1 \log_\alpha (1 - \nu) \cdot r^{\log_\alpha (1 - \nu)} \cdot T(1 - r^T) dr$$

$$= \frac{T^2}{\log_\alpha (1 - \nu) + T + 1}$$

In their example with $\alpha = 0.99$ and $\nu = .637$, if they are willing to entrust a task worth 100 units ($T$), then they must be willing to risk $l(0.99, 0.637, 100) = 49.5$ units.

Many of the above schemes have a relatively static approach to authentication and trust relationships. However, in a real network, once trust relationships have been established, a further difficulty – especially for distributed authentication – is how does one establish trust relationships or delegate permissions on a temporary and "cascaded" basis. An example given in [Sol88] is that of a business person making travel arrangements which will be paid for by the accounts payable office. The accounts payable office will only deal with one travel agency, which communicates with the user and acts on a user's behalf to make reservations for flights, cars, and hotel reservations. Each reservation will require a prepaid percentage of the total reservation for their service. The accounts payable office will only approve of charges that the user signs for, and the reservation agency will not make any reservations without the accounts payable office having guaranteed the funds. Sollins solution is to analyze a threat model of cascading authentication in which per-
missions may derive and cascade from one source through multiple nodes of a network under certain constraints:

As suggested in the example about travel arrangements it is useful in a distributed environment to be able to allow a remote service to act on one's behalf, but retain some degree of control over the actions that the remote service does on one's behalf. The mechanism proposed here to solve this problem of handing of limited authentication is called a passport. The passport identifies the originator and is digitally signed at each transit point, so that each participating transit point is identifiable.

[Sol88:158]

Sollins then provides a protocol which meets the requirements of unforgeability, accountability, discretionary restriction, modularity, independence, and the combining of identities for the passports.

In An Architecture for Practical Delegation in a Distributed System [GM90] the authors use the word delegation to represent a user's authorization of a system or agent to act with some of his own permissions. Furthermore, their protocol, "goes further than other approaches for delegation in that it also provides termination of a delegation on demand (as when the user logs out) with the assurance that the delegated systems, if subsequently compromised, cannot continue to act on the user's behalf. [GM95:10] The authors liken their architecture to that of Kerberos V4 but use public-key technology so as to avoid the necessity of an authentication server. By constructing a protocol and format for the transport of delegation keys (related to Sollin's passports) which are public/private key pairs that authenticate a delegated system to a server, and which certify the next step of delegation:

This architecture goes further than Kerberos V4 in that it permits a "chain" or cascade of delegations through multiple systems. In our chain of delegations the reference monitor knows that all systems in the chain are authorized delegates. In addition, the reference monitor can determine the identity of those systems for auditing and access control purposes. One additional feature, not known by us to be present in other authentication forwarding schemes, permits the delegation to be explicitly terminated by the user so that the intermediate systems, if compromised subsequent to the termination, cannot make use of the rights attained through delegation.

[GM95:21]

Two recent, related protocols for distributed authentication and cascaded or delegated permissions include [BBS94, SAJ94].

Many of these models use point probabilities in their calculations. However, there is often a great deal of uncertainty in one's beliefs of the real world. A potential solution to this problem is to adopt a "fuzzy" artificial intelligence (AI) approach. May suggested
that AI research on belief systems is relevant to the study of trust. [May96] Particularly
the Dempster-Shafer theory, which May quotes [RN95] as follows:

The Dempster-Shafer theory is designed to deal with the distinction between uncertain-
tainty and ignorance. Rather than computing the probability of a proposition, it
computes the probability that the evidence supports the proposition. This measure of
belief is called a belief function, written Bel(X).

[RN95:462]

An added benefit of this method is that it provides a means of combining evidence to de-
rive new Bel values — in a way similar to derived trust in [BBK94]. However, a concern for
these computational models, and in some aspects the decision analysis presented in §2.3, is the
need for events to be independent:

Further, Dempster-Shafer theory provides rules for combining probabilities and thus
for propagating measures through the system. This fourth point is possibly the most
attractive, but it is also one of the most controversial since the propagation method is
an extension of the multiplication rule for independent events. Since many of the
applications involve events that are surely dependent, that rule is, by classical statisti-
cal criteria, inapplicable. The tendency to assume that events are independent unless
proven otherwise has stimulated a large proportion of the criticism of probability ap-
proaches; as it stands, Dempster-Shafer theory suffers the same ill"

[Nut86:846]

Given the above progression of formal models which become increasingly sophisticated,
my intent is not to replicate the formal methods, but to provide a general understanding
of trust in the most comprehensive manner and to show how that understanding can be
used to represent complex interactions on digital networks — and the interactions of trust
with economic value. As such, I must be philosophical for a moment and define my terms
and discuss my assumptions while avoiding the pitfalls of wondering if a user of a net-
work can trust that not only the network exists, but that the user himself exists!

2.2 A THEORY OF TRUST

In keeping with Rangan's treatment, I posit that there is in fact a real world. However,
each agent can consider potentially contrary beliefs about that real world, each which is
expected to be true with some probability. In an abstraction of the direct trust and rec-
ommended trust of [BBK94], I only consider one form of trust which is the trust one
extends about various assertions. An assertion is a statement which asserts an attribute
of the real world. The abstraction here is that one can place a variable amount of trust on
both first and second hand perceptions and stimuli. Trust is the degree to which an agent
considers an assertion to be valid for the real world. There is an associated risk of the
assertions being wrong. Experience is the creation of a history that contains mappings between various assertions about the real world. For instance, someone may predict (assert) that the sun will rise tomorrow, and when my eyes have told me (assert) that it does, I have gained experience. A belief or assumption is a strong assertion that is innate to an agent's intelligence, or perhaps common to many agents (similar to direct trust in [BBK94].) Assumptions are rarely challenged and are considered to be (1) a seed for the evaluation of all other assertions, (2) a common basis for the creation of histories between agents. For instance, the assertion:

- "I exist" is considered to be a very strong assumption. (~99.999%)

- "I believe what my eyes tell me about the real world" is considered to be a relatively strong assumption. (~99%)

- "I believe what other agents tell me the real world" is not an assumption. (~75%)

For instance, an agent may tell me that I may find $5 under the blue stone. If $5 is found under the blue stone, an experience relative to the assumption that I indeed saw it for my own eyes becomes part of my history – experience is created. In this case:

\[
\text{assertion of $5 under blue stone} \iff \text{assumption of I may believe my eyes that $5 was found under the blue stone}
\]

So as to not to always have to question an agent's first hand knowledge, I define an event to be the eventual result or determination of an assertion based on first hand knowledge or an equally strong assumption. The mapping between two assertions (one often being an assumption) is similar to Rangan's belief acquisitions.

Unlike Rangan's (3) of §2.5, I assume agents may accept new assertions which are contrary to previous assertions. (This seems to place undo weight on the significance of early assertions.) Also, Rangan's discussion of possible worlds is useful, but requires that, "an agent believes p, denoted by Bp, iff p is true in all the global states that the agent

---

10 The term trust is often used in two contrary ways which can be confusing. On one hand, "I don't trust that person," means one has a low expectation of the person's assertions being true, and the consequent risk is higher. (Risk is the inverse of trust.) There is also the concept of, "the less risk there is, the less I need to trust the person." (Trust and risk are related linearly.) However, the second statement is speaking about the need of trust to overcome some amount of risk. Rather than this being the expectation of an assertion being true, one is expressing the existence of risk (or variance in one's expectation) but that one will act upon the assertion regardless. For most of this thesis I shall be using "trust" in the first sense. In a later section entitled, "Trust Obviating" I am referring to protocols which decrease the need for trust – and the riskiness of the transactions. In this sense, the underlying technology coerces someones assertions to be as honest as possible. Hence one expects that person's assertions to be accurate, but not because of the person, but because of the protocol which coerces the persons actions to be accurate.
considers possible." [Ran88:26] Since in real life may people consider a wide range of possible beliefs, including $p$, $\bar{p}$, and $\bar{p}$ probabilistically,\(^{11}\) I assume otherwise.

In place of Rangan's (3)\(^{12}\) in which only assertions consistent with previous assertions in the belief-database are accepted, I consider a more complex trust algorithm akin to the derivation algorithms of [BBK94] which generates the probability with which an agent feels an assertion is likely to pertain to the real world. As an example, an agent may see a ball drop 100 times after being released and have a lot of trust (a high expectation) in the assertion that the ball will drop again if released in the future. Trust algorithms can be considered to be function which describe personal behavior, or a deterministic algorithm of an agent, both of which will have some of the following characteristics:

Table 2-1

<table>
<thead>
<tr>
<th>Characteristics of Trust Evaluations</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1. Closeness – given an experience of the form $A_1 \iff A_2$, if $A_2$ is an assumption, the strength of the mapping between $A_1$ and $A_2$ will be greater. Hence, seeing the $5 dollars under the blue rock is closer (and in this case more likely to be believed) than reading about it. This strong mapping may then be used as a basis for believing other assertions about the world. Also, if no money is found under the blue rock this negative experience is closer than having read about the money not being found under the blue rock.</td>
</tr>
<tr>
<td>C2. Accuracy – the degree to which an assertion matches another. Finding $5 under the blue rock, rather than $4, $3, or no money under the rock leads to a stronger experience.</td>
</tr>
</tbody>
</table>

There are also a number of variables which take into account multiple actions from agents over time.

C3. Sample size – the number of times (or samples) an assertion about the real world is taken (seen). (The amount of experience, similar to the relationship between the number of $p$ and $n$ in [BBK94].)

C4. Variance – the degree to which an assertion varies from aggregated experience. (For instance, an assertion may be "too good to be true.")

And amongst the above variables are the demographic categories with which they are compared to or correlated with:

C5. Expertise – Proclamations by an agent that is known to be a doctor (perhaps he has a digital certificate from the AMA to prove it) is trusted with regards to assertions on medical information, but not with regards to automotive information.

C6. Deferral (Accreditation) – The example above of the AMA asserting that a doctor is a

\(^{11}\) (see next section)

\(^{12}\) In §2.1
good doctor is an example of an agent trusting an assertion about another agent.
C7. Threshold (Group) – One many not trust the individual assertion of Larry, Shep, or Moe; but, if all three assert the same thing, one may have a higher opinion of that assertion. Furthermore, one may examine the above components with respect to a specific individual or demographic group:
C8. Individual History – The history of that particular individual (or threshold group).
C9. Category History – The history of similar individuals (or threshold groups).
Finally, there could be any number of initial conditions and assumptions for the algorithm itself.
C10. An agent is generally (dis)trusting in believing assertions.
C11. An agent does (not) give people the benefit of the doubt initially.

For people, this algorithm is most likely not monotonic, and may non-deterministic (seeming irrational), for instance, a favorite saying of mothers with regards to C7 is, “if everyone jumped off the cliff, would you do it too?” This ambiguity with respect to the rationality and expectations of the agents leads one to consider the realms of risk perception and decision analysis.

2.3 DECISION ANALYSIS

A field other than philosophy and logic which may provide a means for understanding trust in the digital realm is decision analysis. Such a mapping seems particularly appropriate since there is a wide body of literature on preference functions, expected values, and risk assessment – all of which are concepts we are attempting to understand in relation to a networked environment. In such an environment, a number of complex interactions may be possible with the development of electronic cash or other monetary instruments. However, simple cases where one user (agent) is in a risky transaction with another agent of the Internet have existed for years – for.sale boards. Be it a mailing list, conference, or bulletin board people frequently trade or purchase materials without a clue of the identity or trustworthiness of the other person – such a market will be termed a networked market. In such as example, trust is the assessment of one’s risk in a certain situation. This example of two users attempting to decide if or how to conduct transactions in such a market shall be a common example of my analysis of trust.

2.3.1 The Value of Credit Information

A common-place occurrence on the Internet is that of a user wishing to buy a product from another. There is risk for both the seller and buyer in such a scenario depending on

---

13 I have purchased items in this manner twice.
the arrangement chosen. A buyer needs to be concerned about receiving the product in working order in return for the money he spends. The seller in turn, needs to be concerned with the quality of payment for his product: will it be the right amount and on time? The concerns of the buyer and seller often take the form of negotiation regarding whether the product is paid for by check, cash, or credit card; whether the transaction is COD, or prepaid. This negotiation shifts the amount of risk between the parties and the level and direction of trust required in the transaction, and is dependent on the economic properties of the supply and demand elasticities for the product.\(^{14}\) For instance, tenancy places the land-owner at risk since the tenant may ruin the property, but because the owner often has a stronger position in the market (a take or leave it deal) he can force the transference of risk unto the tenant with a security deposit.

Often, the buyers and sellers in such a situation are faced with a decision: to purchase the item, or forgo the purchase. In a more sophisticated case, a user also has an option to purchase information concerning the expected result, this is likened to buying credit or rating information regarding the trust worthiness of the other principal. Decision analysis provides one a way to analyze such a scenario. While it probably would not be a plausible nor efficient exercise for conducting transactions over the Internet,\(^{15}\) it does provide an understanding of the concepts involved. Consider the following example from a buyer's point of view.

### 2.3.1.1 Expectation with No Information

The buyer has been offered 1 megabyte of computer RAM for $30 prepaid. 1MB of RAM is worth approximately $40, the buyer has never done business with the seller before and is not very trusting – he expects the seller will cheat him with a 50% probability. The decision the buyer is then presented with is as follows:

```
Decision
  | PrePay  | Chance
  |         | $10
  |         | $0
```

\(^{14}\)\(^{14}\) Chapter 5 of *Microeconomics* \cite{PR95} entitled “Choice Under Uncertainty” provides an economic treatment of uncertainty, whereas *Applied Systems Analysis* \cite{deN90} provides a much more extensive and practical treatment. For a more theoretical treatment see Fishburn’s texts including \cite{Fis73} and \cite{Fis82}.

\(^{15}\) And there is a significant amount to be gained by establishing stable long standing trust relationships or certification agencies so as to avoid such calculations for the most part.
The expected value for the PrePay decision is \((.5)10 + (.5)(-30) = -10\). The expected value of the "PrePay (and as such no transaction) is 0. Since, \(-10 < 0\) the buyer would not proceed with the transaction. A useful extension to this scenario is the expected value of information. (EVPI) This corresponds to the information about a market, the credit history of a user, or the certification a third party could provide to vouch for the level trustworthiness of another user. Assuming that the third party (referred to as a credit agency) is trust-worthy, what service and increased benefits could be provided?

2.3.1.2 Expectation with Extended Information

deNeufville defines the value of information in decision analysis as, "The increase in expected value to be obtained from a situation due to the information, without regard for the cost of obtaining it." [deN90:330] In this example, assume that the credit agency has aggregate market information that shows that prepaid transactions for RAM are honestly completed 80% of the time.

\[
\begin{array}{c}
\text{PrePay} \\
\text{Chance} \\
\text{~PrePay}
\end{array}
\quad
\begin{array}{c}
P = .8 \\
\$10 \\
P = .2 \\
\$30
\end{array}
\]

The revised expected value for the decision is \((.8)10 + (.2)(-30) = 2\). As such, over a significant number of transactions, on average this information provided the buyer with a benefit of $2 - some of which can be collected by the credit agency.

2.3.1.3 Expectation with Perfect Information

The credit agency would be remiss if it was not able to provide specific information about the seller. In such a case, the buyer could attain information about the character of that seller, or it could procure the results of a test in which the credit agency would either "approve" or "disapprove" the transaction on a basis of its own models (perhaps it has a "better", more sophisticated trust algorithm and with the specific information it has on agents it is able to make recommendations regarding a transaction.)

In this case, the credit agency's service is to provide specific information which the user can than apply towards his own preferences, or the agency can give a simple recommendation for conducting the transaction.\(^{16}\) (Or in a similar example, the acceptance of a

---

\(^{16}\) The question of whether the information provided by the credit agency is a valid test is discussed in the §2.5 Appendix A, regardless I must caution that the example of Expectation with Extended Information is an example of a perfect information problem, but the nature of the information is much more general (the whole market), as such, I found it useful to break up the
credit card – every store can not process all the trust information regarding every transaction, hence they defer such decisions to credit card agencies.) Since this recommendation is an assertion (even if it is an assertion about another's assertion) it too is subject to the exercise of trust – or in other words has a probability of being an accurate assertion. The credit agency may be able to assert that it's predictions are accurate 85% at the time, or perhaps one has enough experience with the credit agency to come to this conclusion on one's own. So as to not get too complex, the buyer will continue to trust the assertions of the credit agency, and as such will not worry that the agency is lying about it 85% accuracy rate. In this situation, the user could calculate the expected value of perfect information and assume the credit agency is always accurate. In such a case, the new calculations would be correspond to the following:

First, every test result, $T_{rk}$, from the perfect test will tell us exactly what will happen subsequently, and its associated outcome, $O_{rk}$, will have probability one in the revised decision tree following the test result. [deN90:337]

In our case, we would conduct the calculation taking the branch of each decision with the best outcome. Since we have perfect information and know exactly when to conduct a transaction, the decision tree and expect value is simple: $(.8)10 + (.2)0 = 8.$

```
Decision
   PrePay ——— Chance
      P=.8 $10
     P=.2  $0
   -PrePay ——— $0
```

The expected value of perfect information is then our new result less the old: $8 - 2 = 6.$

### 2.3.1.4 Expectation with Sample Information

Calculations for the expected value of sample information are more complex and require one to consider the fact that predictions are incorrect 15% of the time. Such errors will decrease our benefit because of the good business we lost, and the bad risks that we needlessly took. The calculations are shown for completeness:

Cheating = “C”; Playing Fair = “F”; CheatingPredited = “CP”; FairPredicted = “FP”;

\[
\begin{align*}
P(C) &= .2 & P(CP/C) &= P(FP/F) = .85 \\
P(F) &= .8 & P(CP/F) &= P(FP/C) = .15
\end{align*}
\]

two cases to demonstrate that credit agencies could provide a wide range of information in depth, breadth, and detail regardless of whether the information was theoretically perfect or imperfect.
The probabilities for the actual predcitions are as follows.

\[ P(CP) = P(CP|C) \times P(C) + P(CP|F) \times P(F) \]

\[ (.85)(.2) + (.15)(.8) = .29 \]

\[ P(FP) = 1 - P(CP) = .71 \]

The revised prior probabilities are calculated using the probabilities of the predictions and Baye's Rule.

\[ P(C|CP) = P(C) \times \left( \frac{P(CP|C)}{P(CP)} \right) = (.2)[(.85)/(.29)] = .5862 \]

\[ P(F|CP) = 1 - P(C|CP) = .4138 \]

\[ P(F|FP) = P(F) \times \left( \frac{P(FP|F)}{P(FP)} \right) = (.8)[(.85)/(.71)] = .9577 \]

\[ P(C|FP) = 1 - P(F|FP) = .0422 \]

Now, choosing the best decision for each of the given test results:

\[ EV(\text{Prepaying} | CP) = (.5862) \times (10) + (.4138) \times (-30) = -6.5552 \]

\[ EV(\neg \text{Prepaying} | CP) = 0 \]

\[ EV(\text{Prepaying} | FP) = (.9577) \times (10) + (.0422) \times (-30) = 8.3110 \]

\[ EV(\neg \text{Prepaying} | FP) = 0 \]

Then, “We can calculate the expected value after the test as the sum of the probability of each test result times the value of the best decision after the test result.” [deN90:341]

\[ EV^* = (.29)(0) + (.71)(8.311) = 5.9008 \]

Hence, the value of the sample information in this case was \((5.9008 - 2) = 3.9\).

2.3.2 Trust Algorithms as an Expression of Preference

We have already mentioned the trust algorithm as an expression of the phenomena in which users gauge the probabilities of various events. We defined it as such because while a computer agent could be programmed to calculate the level of trust it will extend towards another (and the risk it will take upon itself), humans generally do not make such explicit calculations. Rather they indicate preferences about various risky situations without the benefit of knowing all of the probabilities. However, even when people are
presented with cases in which they can explicitly calculate an optimum outcome, people behave in an unexpected manner in order to accept or avoid additional risk.

Decision analysis attempts to measure preferences by assuming that the probabilities are relatively straightforward and it is upon these simple probabilities that people express different preferences. While expressed preferences differ for people over the same probabilities, I would argue that the probabilities for the same situation may be perceived very differently; however, it is useful to consider trust in the terms of preference functions for the sake of completeness.

One of the main goals of the study of preference is how to evaluate non-linear valuations of costs and benefits. This is accomplished by expressing these valuations in terms of a value or utility functions. A value function "\( V(x) \) is a means of ranking the order of relative preference between sets of consequences. It assigns a number to every \( X \) such that for any two sets, \( X_1 \) and \( X_2 \), one is preferred to the other only if its value is greater than the other's." [deN90:360] Value functions are dependent on three basic axioms: 1) completeness - for every pair of possible consequences a user will prefer one to the other or be indifferent 2) transitivity - if \( X_1 > X_2 \) and \( X_2 > X_3 \) then \( X_1 > X_3 \) and 3) monotonicity - more of a good thing is better. Utility functions differ from value functions in that the utility function's units have meaning relative to each other. The result being "provided always that the axiomatic assumptions are valid, it is possible to evaluate choices analytically, even when people have nonlinear preferences." [deN90:365] The axioms necessary for the utility function are the same as that of the value functions plus three additional ones 1) existence - probabilities exist and can be quantified (not always the case in every day life) 2) monotonicity for probability - the higher the probability of a good thing the better, and finally 3) substitution or independence axiom - "a person's preferences for an item should vary linearly with the probability of its occurrence." [deN90:366] The result being that a person's preference may be non-linear with respect to the benefit (when conducting a transaction for $10,000 rather than $1 a user is much more likely to be risk averse), but not in probability (the preference of getting $1 with

---

17 Some basic mathematical axioms such as existence, or definitions of the ranges of the probabilities s.t. probabilities must be between [0,1], etc. are required as well but have "little significance to the analyst" [deN90:361]. See the Fishburn references [Fis82] and [Fis73] for a theoretical treatment.

18 The terminology of this field is be non-standard (less so than complexity theory though!). Often these utility functions are known as von Neumann-Morgenstern utility functions and different discussions of the axioms may tend to focus on one axiom more than another. Fisher presents four axioms for which the von Neumann-Morgenstern utility functions hold which are generally the same concepts discussed elsewhere: complete ordering, continuity, importance of ultimate prizes and their probabilities, and the most important/controversial - strong independence.
probability = .5, should vary linearly with the preference of getting $1 with the probability = .6.)

The question then is, is trust an expression of a utility function? Utility functions are often measured with a lottery which is similar to our example of a transaction over the Internet. A lottery is a set of possible outcomes each with a separate probability of occurrence. Utility is then measured by offering a series of lotteries to the user, each converging closer to a point where the user is indifferent between the two lotteries. A series of these exercises can approximate a person's "utility function." It is unlikely that such calculations would be of use when measuring the preferences of users on the Internet other than to describe the general behaviors of an aggregate number of these users. Also, some of the axioms above, particularly the existence and substitution axiom, may not hold for some transactions, but deNeufville states:

As a practical matter, recent research we have conducted indicates that the substitution axiom holds, at least as a first-order approximation, except when the probability of some greater consequence is either very small or close to certain.

[deN90:368]

Regardless of the contentious questions of this discipline, we have come to an understanding of trust which is reflected in the following three definitions.

*trust* - the expectation of an assertion being true.19

*trust algorithm* - an algorithm that determines/explains the creation of the expectations.

*trust-utility function* - the way in which agents will exhibit non-linear preferences with regards to risk (i.e. amount of trust.)

### 2.4 TRUST AS COMMERCE

The third definition of trust was given as follows, "Confidence in the ability and intention of a buyer to pay at a future time for goods supplied without present payment." This definition allows one to consider aspects of trust not given in the straightforward treatment of the previous sections. For while it may seem intuitive to consider trust in light of decision analysis, the expected value and probabilities in such an analysis are considered to be phenomena of the real world and not interactions with competitive agents. For instance, consider the case where agent(B) -- who plans to cheat -- offers agent(A) $20 for a 1M of RAM. Agent(A) may be suspicious and not accept the offer based on his expectations (level of trust) of agent(B). Knowing this beforehand, perhaps agent(B) would offer $100 for 1M of RAM. If this were a very simple expected value calculation, in which the

19 Within the confines of the philosophical issue of "in as far as we can know what is true."
probability of the $20 and $100 deals were the same, a cheating agent could inflate the outcome so as to turn the decision to his favor. However, one of the variables considered in the treatment of the trust algorithm was the consideration for outcomes which seemed "too good to be true." This section will deal with such topics more specifically and shall refer to concepts from micro-economics and game theory.

In [FD95] Hal Finney and Wei Dai discuss a concept related to trust, that of reputation. Reputation is the amount of trust an agent has created for himself through interactions with other agents.\(^\text{20}\) Hence, if one's assertions consistently meet the expectations of other agents, they will have higher expectations of later assertions being valid. Reputation is valuable for three main reasons. A user may prefer to conduct transactions with users he trusts. The costs of transactions between trusting users may be smaller because third party reputation services need not be consulted. Finally, if the conditions are right, one can betray one's reputation for a very large gain.

The exact economic nature of reputation and trust is not often addressed with regards to transactions over information networks aside from discussions on the cypherpunks list. In [FD95] Dai wrote:

In a reputation based market, each entity's reputation has three values. First is the present value of expected future profits, given the reputation (let's call it the operating value). Note that the entity's reputation allows him to make positive economic profits, because it makes him a price-maker to some extent. Second is the profit he could make if he threw away his reputation by cheating all of his customers (throw-away value). Third is the expected cost of recreating an equivalent reputation if he threw away his current one (replacement cost).

In more traditional economic terms, reputation could be viewed as an asset: "something that provides a monetary flow to its owner. For example, an apartment can be rented, proving a flow of rental income to the owner of the building." [PR95:157] It probably cannot be considered a product in that concepts of supply, demand, marginal cost and other costs associated with production do not generally hold. For instance, consider how trust can be created:

- Trust is created through the development of experience with other agents. Hence, it is a relation rather than a product. For instance, if agent(B) successfully completes a transaction with agent(A), agent(B)'s reputation is still a product of the "arbitrary" trust algorithm agent(A) employs. (Agent(A) may be distrustful no matter how many satisfactory transactions occur.)

\(^{20}\) Reputation already plays a key part of many Internet communities including Internet Relay Chat (IRC), Backgammon Servers, Chess Servers, Netrek, and MUDS.
• The only relevant cost in the creation of trust seems to be the opportunity cost of betraying that cost. (Any costs pertaining to the transaction itself (i.e. the cost of being on the network) would be accounted for in the cost of the transactions.
• There is an bound on both how much (100%) and how little (0%) trust can be generated.
• Agents can transfer trust by certifying another agent.
• The creation or destruction of trust is not a zero sum game. The net sum of trust may increase or decrease.

However, perhaps these two general rules could be applied in decision making regarding reputation creation as asset creation:

1. An agent should maximize profit over its planning horizon, where profit is defined in the economic sense as revenue less costs, including opportunity cost. The opportunity cost of reputation is the excess revenue that could be generated from the exchange of the reputation (and its revenue over the planning horizon) for immediate revenue (by cheating) over ones planning horizon.

2. The decision as to whether to invest in building reputation is subject to the NPV Criterion which states: Invest if the present value of the expected future cash flows from an investment is larger than the cost of the investment” [PR95:532] or if the following equation is positive for cost $C$, discount rate $R$, and time horizon $n$.

$$NPV = -C + \frac{\pi_1}{(1 + R)} + \frac{\pi_2}{(1 + R)^2} + \ldots + \frac{\pi_n}{(1 + R)^n}$$

The important consideration here is that $C$ is a function of an agents reputation algorithm and the trust algorithms of agents with which he will interact.

Clearly, the defining and characterization of trust and reputation in such a scenario soon becomes very complex, however, in having brought trust to the field of economics there are a number of sub-disciplines with which trust can be considered. The first is thought regarding markets with asymmetric information.21 The most common example is that of the used car market where the buyer has very little knowledge regarding the quality of an object being purchased.[PR95:594] Such a market is characterized as failing because of asymmetric information (or the lack of trust). The example of the used car often fails because the market is perceived as being one of low quality cars, which exacerbates the removal of high quality cars from the market, which in turn exacerbates the perception of their being a disproportionate amount of “lemons.”

Also, there is a fair amount of economic literature pertaining to product quality and asymmetric information particularly in the field of insurance and credit. Another field

---

21 For work on information economics see [Bab93, EMN89]
where this is an issue is for a temporary good or service. An example of which is a highway motel or restaurant that is not likely to have repeat customers and cannot build a personal reputation. The solution may have an interesting application in the network world and consists of creating a reputation – market brand – through standardization. For instance, all McDonald's have relatively the same color schemes, foods and prices and attract customers that may have never eaten in that particular restaurant. [PR95:598] Hence, brand in the form of logos, seals, or labels on the WEB may be of great importance. Other common economic concepts are that of market signaling, particularly guarantees and warranties. [LMN94]

The second field of economics which is of interest is competitive strategy and game theory. In [WD95] Dai mentions a reputation algorithm (the counterpart of the trust algorithm) which determines the optimum conditions for increasing utility, be it through creating a strong reputation, or exchanging it for some other value. Dai posits that a good reputation algorithm (1) need be efficient (I assume this ranges from optimal efficiency to at least a "competitive" efficiency), (2) not too costly to evaluate, (3) and relatively stable in an evolutionary system. Such characteristics apply towards trust algorithm as well. In fact, trust algorithms and reputation algorithms can be thought of as competitors in a networked market where information and one's algorithms determine one's success over time.

Already we have seen that agents are employing both their trust and reputation algorithms so as to make the best choice against potential competitors. In [FD95] Finney refers to the Prisoner's Dilemma (PD) game as an example of a simulation of agents concerned with reputation. In the PD game two prisoners are each accused of having committed a crime. They are in separate jail cells, and cannot communicate with each other but are familiar with the rules of the game and the consequences of their own and the other's actions. Each has been asked to confess. They both know if neither confesses (they have "cooperated" to their mutual benefit), the prosecutor's case is difficult and each prisoner goes to jail for only one year. If both confess, they will both receive a five year sentence. However, if one confesses (defects), and the other does not, the one who confessed will go to jail for only one year still, and the one who did not will go to jail for ten years. The dilemma is then does one risk a ten year jail term by not confessing, or does one confess and only have a one year jail term if they other guy cooperated (his ten year jail term is tough luck) or a five year term for the both of you if he confesses as well. The dominant strategy in this game is to confess (defect) since confessing betters ones position regardless of the others' actions:

if the other guy defects, you get five years instead of ten.

if the other guy cooperates, you get one year instead of ten.
Such games can be played multiple times, over which the agents playing have a fixed amount of memory with which to hold a grudge or to preen their reputations. In 1984 Axelrod conducted a tournament [Axe84, Axe87] in which algorithms programmed by humans competed against each other in the iterated PD game. Interestingly, such games also lend themselves to the employment of “genetic algorithms” [Hol75] in which competitive algorithms evolve by promulgating the “fit” strategies through the lifetime of the game by the reproduction of winners, the crossing of two different winning strategies, or through random mutation. For instance, three intuitive strategies for a two agent, one-round memory game (in which the opponents previous move was considered to be that of cooperation (C) or defection (D)) are shown: [MMC95:5]

<table>
<thead>
<tr>
<th>Strategy</th>
<th>CC</th>
<th>DC</th>
<th>CD</th>
<th>DD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Always Cooperate</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6. Tit for Tat</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16. Always Defect</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

The representative string of digits is akin to a biological's creatures “chromosome.” This rather intriguing research has been extended to analyses of various real-life market places. In [MMC95] the authors use genetic algorithms to simulate the creation of brand by marketing managers in simulated regional coffee markets and show “that in the limited tests we can feasibly conduct these agents outperform the historical actions of brand managers in this regional market.” [MMC95:3]

Consequently, the fascinating realms of competitive strategy and game-theory [Eis95, MMS90, MMC95, KSS95, LMM095, PT95]; emergent behavior/institutions [Kel95, Vri94]; electronic [Bar94, DF95, Var95] and information [Dys95, EMN89, Wit82] markets and complexity [Wal92] are relevant to the study of trust.
2.5 APPENDIX A: CREDIT RATING AS A TEST\textsuperscript{22}

In my examples I have used the term for information gained from another agent or service in the electronic market as a "test," a term used in the domain of decision analysis as explained earlier in the chapter. Examples of such tests in decision analysis are usually very applied, one takes an x-ray, or one runs a series of diagnostics on a machine, each of which has some sort of associated cost. Most of this information is collected and used in a period after the decision to actually use it. (A doctor decides she wants an x-ray, and she then takes it.) The motivating example with regards to credit rating is where a third party actually takes specific information from the past. I strongly expect this information is equally valid as a "test" as deNeufville states, "A test simply refers to any specific process of collecting information." [deN90:330] This information will still have a cost associated with collecting and processing it and will increase the information available to the decision maker and as such, will have some value. Consider a simple example of a dice being rolled at a casino table. A single die is being repeatedly tossed and one will wish to know if the die is biased. A logical desire is to test the die, perhaps throw it a significant number of times before agreeing to bet on it. An equally valid approach would be to consult a trusted friend who has been recording the coin tosses before one arrived. This is in essence what a credit agency does. As long as the tosses are independently distributed (IID) over time, looking to the past is as equally valid as looking to the future.

Information gathered from the past does have an interesting characteristic. Often, a history only goes back so far, the friend at the table may only remember the last thousand tosses. However, one could also conceive of a die that is very fragile, and will only last a thousand more tosses from when one decided to test it. As such, the repeated consultation of a credit agency over an over in quick succession may not provide any more information than the first consultation (though agencies certainly could sell information services on an incremental basis). However, providing the exact same inputs to an algorithm to test its validity do not increase one’s information either, hence I trust the definition of "test" in the paragraph above (any information) with respect to information gathered and analyzed in the past.

\textsuperscript{22} This analysis is based on a conversation held with Brett Leida of MIT on 4/8/96.
TRUST INSTRUMENTS

When I consider life, 't is all a cheat.
Yet fool'd with hope, men favour the deceit;
Trust on, and think to-morrow will repay.
To-morrow's falser than the former day; ...

As stated, the thesis of this analysis has been that trust in an electronic market can be dealt with in one or more of three ways. The most obvious way is to extend the basic principles of trust and risk management from the "real" market to the digital world. For instance, instruments such as letters of credit facilitate low-trust international transactions. Or, in cases where the rate of default is predictable, the cost of default can be included in the cost of the service. Additionally, methods like credit rating and scoring will exist in one form or another to lessen the likelihood of default and bridge the gap between untrusting principals of digital transactions. This chapter focuses upon real world trust and financial instruments and their relationship to electronic markets.

3.1 TRUST MANAGEMENT AND FINANCIAL INSTRUMENTS

The terms of this subsection's title, "Trust Management Instruments" and "Financial Instruments," are used to represent the broad range of tools one can use to help exchange value in a market place. Each tool (instrument) has a quality or attribute that makes it more suitable for aiding certain transactions than others — when anonymity is required, cash is an instrument of choice. In the real world, a whole range of financial instruments exist to satisfy the needs of market participants. Each varies with respect to operational qualities such as anonymity, immediacy, and cost. They also differ with respect to the strength and direction of trust that is inherent to the use of that instrument. Some instruments may require a great deal of trust between the participants (but may facilitate very fast transactions), others are specifically intended to allow transactions to occur in a

[23] For an example of how this is being done see [ABA95a].
low trust environment. The low trust environment is similar to the motivating problem of how to buy and sell things over the Internet as described in Chapter One.

Unfortunately, the term "instrument" may be misleading because it connotes a sense of physical substance, as if the real world object necessarily embodies the functionality of the instrument (for instance a piece of paper or token). This is not necessarily true, and is becoming less true as finance becomes further digitized. The piece of paper or token has little *intrinsic* value. Rather it is the representation of a capability or service. Hence, it is useful to think of an instrument as *both* the underlying service and its representation (physical or digital). With the above discussion in mind, I shall henceforth make the following distinctions for the digital world:

_instrument_ — a *service* that is provided to facilitate the exchange of value and its representation or *certificate*; similar to "3. that with or by which something is effected; means; agency" [Web89:737]

_service* — "13. ... the performance of any duties or work for another; helpful or professional activity." *[Web89:1304]_

_certificate* — the representation of a *service*; this may be in the form of a token, legal agreement, security, digitally signed assertion, etc. Similar to "1. a document serving as evidence or as written testimony, as of status, qualifications, privileges, the truth of something, etc." *[Web89:242]_

Hence, a dollar bill is an instrument, it is the representation of a service provided by the government that allows users to exchange value.

The provision of a service usually has an associated cost. Costs of cash include the printing and distribution of the currency. Indirectly, if the service is to maintain a healthy economy so that value may be easily exchanged, there is a cost of keeping that economy stable. However, the "cost" of maintaining a healthy economy is hardly a cost at all since it provides a means of mitigating information asymmetries and other market failures. A healthy economy supports increased trust in the underlying currency, this allows distrustful parties to easily exchange value (since they can pay each other in the currency) which fuels the economy onward. The result is that the increase in social welfare outweighs the associated costs of producing the currency. However, as many countries are aware of, if the reputation of the underlying service is poor, people simply won't use it.

Hence, physical bills, stock and bond certificates and legal contracts are the representations of a service, even if that service is not related to the immediate transaction. (The US government is not providing a direct service to drug dealers dealing in cash, it is an indirect — but useful — service.) These instruments are not commodities with intrinsic value.
Rather, they are a representation of a complex web of trust relationships and services — just as digital certificates are representations of trust in the digital world.\footnote{See Chapter Six.}

With the above definitions in mind, it is useful to consider the differences between digital and real world instruments today. Real world instruments have a large real world service component and are the end product of a rich institutional and technological history. For instance, credit cards evolved from the growth of store credit to their customers. In the 1940's general purpose cards (which could be used in multiple places) like the Diner's Club or American Express Card became popular for obvious reasons. In the 1950's banks became interested and issued cards to local merchants who did not have credit cards, or found it more effective to allow the bank to manage the accounts and billing/management costs. Eventually networks of cooperating banks (led by Bank of America) emerged which led to the market as it stands today. The technical standards of this instrument are very much associated with the structure of the credit card industry: the processing and fee structure between the card issuing banks, merchant acquiring banks, settling banks, and the transaction processing network. [LR95]

Digital instruments have some interesting attributes, the first of which is the entities that use digital instruments are just as likely to be computerized infrastructure or agents as humans. At the market level they are very dynamic – only those services which are competitive will profit. They aren't hindered by antiquated technologies nor business models. However, they do not have the resources (both financial and institutional) nor the brand of traditional real world instruments. Regardless of technical superiority, the success of independent payment mechanisms, such as DigiCash, in the face of large credit card companies is far from conclusive.

Digital instruments as discussed in Chapters Four and Five are very young yet. The number of services provided are few, and the technical representations are still being standardized. One of the few technical pieces describing instruments beyond direct payment is Endorsements, Licensing, and Insurance for Distributed System Services [LMN94] which distinguishes three types of certification instruments:

\begin{itemize}
\item \textit{license} – a credential that indicates a service provider is legally authorized to provide a service \{which\} \ldots has been found to meet certain minimal qualifications required by the law \ldots [This is a form of a credential.]
\item \textit{endorsement} – provides assurance that a service provider meets more rigorous requirements determined by the endorser \ldots [Another form of a credential.]
\item \textit{liability insurance policy or surety bond} – provides a client with a means to recover damages in the event of a loss that is the fault of the service provider. [Where] liability insurance policy represents an agreement between two parties, and a surety bond
\end{itemize}
represents an agreement between three parties: the surety, the obligee, and the principal.

[LMN94:170]

Licenses and endorsements are both forms of general certificates described in Chapter Six. Liability insurance policies and surety bonds are related to instruments discussed in this chapter, Chapters Four and specifically Chapter Five in that shift the assumed risk between the principals.

However, as exciting as the possibilities are, the number of instruments available to Internet users is small, as Tim May stated, "... the 'ontology' of digital money, the instruments and forms it can take, are impoverished compared to the real world," and he challenged readers for the cryptographic equivalents of:

- ~ options
- ~ warrants
- ~ bearer bonds
- ~ promissory notes
- ~ zero coupon bonds
- ~ checks
- ~ receipts
- ~ lockboxes
- ~ coupons
- ~ time deposits
- ~ money orders
- ~ escrow
- ~ IOUs

Furthermore, May states that the structure of such instruments are similar to the realm of computers:

A look at any book on money and finance shows a rich "microworld" of "things" and "procedures" (classes and methods attached to classes). The classes have subclasses, and the methods have various behaviors and "expectations" attached (more than just simple class behavior, more of an AI or agent flavor, in my view).

[May95]

Rather than categorizing such structures as those familiar to the artificial intelligence discipline, I would tend to characterize the instruments in terms of protocol development. In Chapters Four and Five the exposition demonstrates that complex protocols are built upon a hierarchical of other protocols and concepts. For instance, a security deposit is dependent on electronic cash and zero knowledge proof protocols, which are dependent on blind signatures, which are dependent on encryption and signature algorithms.

Finally, before proceeding on to examples of such instruments, I must qualify my aggregation of trust instruments with financial instruments. I have already stated that financial instruments provide services, and a significant service is the provision of trust (introduction, reputation, certification, etc.) Financial instruments are a means of transferring or creating value. Trust services, as discussed elsewhere in this thesis, increase the efficiency or likelihood of the transference of value by acting as a "market making" or at least "market honing" force that brings otherwise recalcitrant buyers and sellers
together. Many hope that such digital "intermediary" services will increase the efficiency of a market, decrease the costs of search in a market, and act as a "market maker" where a market would otherwise fail – leading to "friction free capitalism." [BLW96]

Furthermore, it is my belief that in the digital world bits will be bits. For instance, one string of bits may certify that a user should have access to a service. This string of bits could be exchanged for bits that represent an equivalent value in ecash. Just as stocks, bonds, and certificates have a market value, so will digital instruments. As we discussed in Chapter One, reputation itself has value and can be both purchased (by enduring the cost of staying honest), sold (by betraying trust) and transferred (credit agencies). Due to the nature of digital technology and an ubiquitous network, the liquidity of value as represented in various instruments will be very high. However, certain digital instruments will still be valued more than others (or more cost effective) for the appropriate transaction. While one may be able to simulate one instrument using another, the added costs of such transformation may be counter-productive.

As such, an understanding of the concepts regarding value and how they relate to trust instruments shall have a direct bearing on how trust develops in an electronic market.

### 3.2 INCORPORATING RISK INTO THE COST

If one has a predictable model of customer untrustworthiness, a simple way to handle the lack of trust is to simply incorporate the cost of the defaults into the charges levied for those services. For instance, if 5% of credit card users default, the companies can increase the fees associated with having a credit card.²⁵ Of course, credit card companies also compete on the basis of fees (particularly the annual fees)²⁶ hence it befits them to keep the default rate as low as possible using various selection methods.

### 3.3 CREDIT

One of the most ubiquitous and profitable trust brokers of the real world are banks which extend both personal and corporate credit through charge accounts or loans. These services provide the "lubrication" for much of the activity of our economy. With respect to personal credit, every merchant cannot know the trustworthiness of every customer.

---

²⁵ However, this is not to say credit card companies like for everyone to pay their full balance every month since they make most of their profits on the interest rates charged against the outstanding balance.

²⁶ Annual fees are driven to $0 while interest fees are non-competitive. Lawrence Ausubel, of the University of Maryland has suggested that this is because customer's shop on the basis of annual fees, and assume they will always pay their bills on time and need not worry about interest fees. (Passell, Peter. NYT, August 17, 1995, p. D2.)
Banks and credit card companies alleviate this problem by exposing themselves to risk – for a price – while allowing most transactions to proceed unhindered. However, banks and credit agencies also wish to minimize their risk\textsuperscript{27} with respect to the price so as to profit\textsuperscript{28} the greatest. In order to accomplish this they have developed sophisticated systems of measuring the trustworthiness of potential customers. This system is known as credit scoring and is the subject of many texts and theoretical treatments. Credit scoring is described as the “scientific approach to determining which applicants are granted credit,” [TCE92:v], and has existed for many years – but only became serious when scoring tables become widely used in the 1970s.\textsuperscript{29} A whole profession exists for credit management and many reference books [Bon93, Ban98] and analyses of the credit market itself [Cle86] exist to help reduce the risk and increase the profits inherent to such operations.

Banks use a number of factors to determine the ability of a loanee to repay a loan. While credit card analysts may tend to look at previous history (hence reputation is key) banks making business loans also look at a wider range of variables. Some of the factors include:

1. *character* – the intention of the borrower to repay a loan.
2. *capacity* – the ability of the borrower to repay.
3. *capital* – value of the firm and the present value of future earnings. (sales, assets, liabilities, profits, net-worth.)
4. *collateral* – assets that a borrower will pledge as a security.
5. *condition* – the economic environment of the borrower.
6. *coverage* – amount of resources available to discharge liabilities normally.

I expect that the model of business loans will proceed unscathed into the digital era. However, those that make personal loans and credit will be extremely interested in the expressed preferences and demographic data that will be available to WEB merchants and marketers.

\textsuperscript{27} This individual loan risk management is separate from the “portfolio effect” which entails, “considering the relationship between the new loan in question and other existing loans or potential loan opportunities. By lending to both cyclical and counter cyclical industries a bank would attain an optimal combination of loans whose total credit risk would be less than the sum of the credit risk of individual loans” [Onu79:21] in a manner similar to investment portfolios.

\textsuperscript{28} See [TBE89:3] for *Credit Scoring and Credit Control From Four Points of View* by E.M. Lewis for a historical background to credit scoring.
3.3.1 Money

The economic instrument most familiar to people throughout the world is money. I have already mentioned that intermediary trust services allow one to exchange services in a market that would have otherwise failed. Money accomplishes the same thing except that there is an extra step of indirection – one is not trusting an immediate third party, but the stability of the currency. Just as one tried to find a certification path\(^{30}\) in privacy enhancing technologies, one attempts to find a path in a market through which one may trust that the transaction can occur in a fair and valid manner. (Also, one wishes this path to be cost effective as well.)

This history of money\(^{31}\) is fascinating, and the future capabilities of digital money are very exciting. An example of some of the economic research being done with respect to modeling money is *Money as a Medium of Exchange in an Economy with Artificially Intelligent Agents* [MMS90]. In this paper the authors extended a model of Kiyotaki and Wright\(^{32}\) that describe economies “in which particularly commodities emerge as media of exchange .... or in which a good from which no agent derives utility emerges as fiat money” using agents employing Holland's classifier systems.\(^{33}\) However, with the truly daunting amount of literature on the nature of money, and a quickly growing series of treatises, analysis\(^{34}\) and articles on ecash\(^{35}\) I cannot address all aspects of digital money in depth. However, I will briefly touch upon those aspects with respect to trust.

As in other places in this thesis, a definition would be useful. In a 1992 Forbes article, Peter Huber defined money in the following way:

> Money, as I wrote here two years ago (FORBES, May 14, 1990), is just another network, our oldest medium of systematic communication. And new communications technologies are fast surpassing the old. The paperless office, unlike the paperless office, is at hand.

[Hub92]

---

\(^{30}\) See Chapter Six.

\(^{31}\) See *Money – History, Present & Future - and E-Money* http://www.ex.ac.uk/~RDavies/arian/money.html for an excellent site and further resources on the topic.

\(^{32}\) Kiyotai, N. and R.Wright 1989, On money as a medium of exchange, Jour of Political Economy 97, 927-954.

\(^{33}\) Holland's classifier system is a form of a evolutionary modeling system. Similar work is referred to in §2.4.

\(^{34}\) For instance, see *The Economics of Digital Commerce: An analysis of Digital Cash, Electronic Fund Transfers, and eCash* [Pan95] in which Panurach describes the characteristics of various digital instruments with respect to anonymity, liquidity, velocity, and money-supply.

\(^{35}\) *International Design* even had an article on ecash entitled “What is Money” by Karrie Jacobs. ID March/April 96.
Cathcart wrote an excellent text on money, credit, and the economics of monetary systems [Cat89] in which he enumerated the operational qualities of money:

1. the ability to be readily recognized and agreed upon as to exchange value
2. the ability to be stored securely and inexpensively
3. the ability to be transferred securely and at low cost
4. the ability to be combined in small or large values so as to be useable conveniently in a wide range of transactions
5. stability in purchasing power
6. wide acceptability

[Cat89:226]

Electronic money systems will most likely have no difficulties with respect to points {1,5, and 6}, points {2,3} will be the advantage of electronic money systems over current systems. While {4} has been difficult technically with respect to splitting anonymous, digital cash it does not provide a problem for stored-value smart cards.

Cathcart also provides a classification of money, by which trust in the value of money is based on the acceptance and conditions of the money supply:

1. commodity money has intrinsic value (gold, cigarettes, etc.)
2. traditional money enjoys confidence and acceptance primarily because it has become institutionalized in the economy
3. fiat money is legally mandated.
4. debt money is a debt instrument or deposit balance redeemable on demand in another asset of value—usually commodity money or fiat money—and is backed by less than 100 percent by reserves. [This leads to an increase in credit and money—like trust.]

[Cat89:230]

Commodity money is the most trusted because of its inherent value, hence if it is not used as money, it can be used for its innate utility. (As decoration for royalty in the case of gold or as a fix for an addiction in prisoners who smoke.) The stability of fiat money is dependent on the economy of the government backing the money or the capability of the government to enforce its acceptance by force of arms. In an examination of Yeltsin's efforts to prop up the Russian ruble Huber wrote of the importance of trust:

But new governments of young nations, especially nations with turbulent histories, can't make money, either. Nobody quite trusts them, and without trust the paper lovingly engraved at the government mint is valueless.

[Hub92]

An example of institutional money is the use of stones by the culture on the Isle of Yap in the Pacific where a person's wealth is dependent on the number and size of rocks he possesses, such a form of money can be easily destabilized with the introduction of new supplies or a distrust of the institution. Debt money is trusted in as far as one expects to
the issuers to repay their depts. As we discussed in Chapter One, (4) points out that trust and money supplies are not necessarily zero sum games, the use of some instruments, and the increase in the faith of a money or its backing institution can lead to an overall greater money supply and trust in the economy – and consequently the capability to easily exchange value at a low cost.

3.3.2 Trust and Securities

To further support the argument that trust and financial instruments are tightly coupled, consider the nature of trust, and then the markets for securities. Earlier, I quoted the OED for the definition of trust:

1. a. Confidence in or reliance on some quality or attribute of a person or thing, or the truth of a statement. Const. in (+of, on, upon, to, unto).
2. Confident expectation of something; hope.
3. Confidence in the ability and intention of a buyer to pay at a future time for goods supplied without present payment: = credit sb. 9 a. Chiefly in phrases

Futures, options, stocks, and bond markets are all creatures of trust. In a futures market one obligates oneself to sell or purchase a commodity at a set price sometime in the future. With an option, one acquires the right to sell or purchase a commodity at a certain price. Each is an expectation of the future and an attempt to profit by or hedge one’s risks against uncertainty. (Both assertions, expect values, and risks are discussed in Chapter One.) When one plays the stock market one is again making assertions about the expected performance of a company or the market itself. A bond is a loan to a company, government or other institution and the market hinges upon the buyers confidence in or reliance upon the ability of the issuer to meet the interest payments and to redeem the full value of the bond upon its maturity. Each market, and particularly the bond market, has certification and reputation agents that provide information services (as is obvious from the innumerable number of indexes, portfolios and rating services). For instance, the reputation of a bond is extremely important and are rated according to the riskiness of the loan. Rating services such as Standard and Poor’s and Moody’s ratings dramatically influence the attractiveness and the rates of bonds offered. (Lower rated bonds must offer higher rates to compete against higher rated bonds.)

Primitive markets are forming on the Internet which may come to resemble the more sophisticated markets of Wall Street. The Security Exchange Commission (SEC) recently gave permission to Spring Street Brewing Co., to continue offering information services
as an introducer of interested buyers to sellers of its stocks.\footnote{\textnormal{[Tay96]WEB stock reporting and index services are becoming widely available and popular to on-line investors, and a nickel and dimes futures market for the presidential election is being conducted by the University of Iowa College of Business Administration!}} However, older institutions are not standing idly by, as seen in the SEC’s own offering of its EDGAR company and mutual filings database to the Internet with the SEC-Live service. Given this, the Internet will increase the efficiency of normal market institutions (replacing or extending the capabilities of current indexes and services) and it will give birth to new and hybrid services such as distributed ecash based trading. An oft cited example of such a market is the Electronic Cash Market web servers and mailing lists\footnote{\textnormal{See the Electronic Cash Market Mailing List at http://www.ai.mit.edu/people/lethin/ecm.html.}} on which people buy and sell various demo (and now real) ecash instruments.

### 3.3.3 Letters of Credit

Letters of credit are perhaps the most striking example of an instrument that is suitable for electronic markets.\footnote{\textnormal{See LETTERS OF CREDIT http://www.cambtrade.com/Letters.html}} They are commonly used in international commerce when a buyer does not trust the seller nor the foreign (legal) institutions involved in the transaction. The same concern is felt by the supplier. Note, that it is not merely a lack of trust in each other that may hinder an international transaction, but also an inability of each party to rely upon an infrastructure (collection agencies and legal systems) to collect money even when one party cheats. What can bridge the gap of trust to allow the transaction to occur? \textit{Letters of credit} structure payments through trusted intermediaries and credible commitments so that each party is confident of payment. For instance, a supplier of sprockets in the US wishes to sell his merchandise to a buyer in Japan. Each party and his representative bank make certain commitments for payment. The buyer’s (Japan) bank makes promises to pay the supplier’s (US) bank if the sprockets are delivered according to the contract regardless if the actual buyer pays his bank. Likewise, the supplier is not paid until he has supplied the purchaser in compliance with the contract. While each country has its own sets of law and regulations regarding banking and collecting debts – which is why the international transactions are difficult in the first place – terms for defining and documenting letters of credit and the resulting transaction are fairly uniform and are defined in Uniform Customs and Practices for Documentary
Credits. Any problems within each jurisdiction (for instance, if the buyer doesn’t pay for the sprockets) can be resolved within that jurisdiction (the Japanese banks sues the buyers according to domestic law) but the supplier still receives his payment.

A obvious digital counterpart to the letter of credit is not financially orientated, but it is the exclusive focus of Chapter Six: certificates. Certificates enable two users who may be a world apart to mutually exchange keys by relying upon a hierarchical system of intermediary trust services. So just as I may not be able to trust a user’s bank in Japan, I do not trust his key server, but I do trust the US server, which trusts the UN server, which trusts the Japan server.

3.3.4 Digital Bearer Bonds

Another trust instrument which enables strong anonymity like cash is the digital bearer certificates. Robert Hettinga has argued that digital bearer certificates may return the method of securities exchange to its relatively anonymous state when a bond could be transferred between parties. [Het95] Prior to the 1970’s bonds were anonymous bearer instruments, every bond certificate had a number of detachable coupons which could be sent in to the issuer for redemption. This meant that the bond could be exchanged anonymously and out of sight of various government agencies such as the IRS. However, after legislation requiring that such transactions be reported and a 1983 SEC ruling, many bond holders do not even receive a certificate, all payments and transactions are conducted (and reported) electronically and are called book-entry bonds and are easily traceable.

Hettinga argued that while the low cost and hierarchical structure of the communications networks on which trading services occurred made it easy for government to regulate these securities, the even cheaper and distributed nature of Internet style communications will make the regulation of digital bearer bonds all but impossible:

So, with a digital bearer bond, you would have in effect a bundle of digital certificates. One would be for the principal and would be good for the repayment of that principal on the date the bond was called or the redemption date, however the bond offering is written. The other certificates would represent coupons, one for each interest period for the life of the bond.

These digital certificates, in combination [with] increasingly geodesic networks enabled by exponentially falling microprocessor prices and strong cryptography,

---

39 See U.C.C. - ARTICLE 5 - LETTERS OF CREDIT

40 Dimitri Vulis opined that “The fact that bearer bonds were outlawed suggests that if and when new ways are invented to conduct financial transactions that are conductive to tax evasion (e.g., using anonymous electronic payments), they too may become outlawed.” [Vul95]
theoretically allow secure, point-to-point trading of any security of any amount with instantaneous clearing and cash settlement.

[Het95]

3.3.5 Land Owners (Security Deposits)

Another mechanism, which is the significant focus of Chapter Five, for bridging the gap between buyers (renters) and sellers (land owners) of leases for apartments is the security deposit. In such a situation, the security deposit can be thought of as a coercion of the renter’s behavior. In a microeconomic framework, this market is characterized by asymmetric information: the land owner does not know if the renter will be able to pay for any excess damage to the property. In this case, a security deposit acts as a signaling mechanism that tells a land owner that the renter is a trustworthy person, and has made a credible commitment to demonstrate the fact—much like warranties, insurance, and brand enables parties to signal in other types of markets.

3.4 THE EFFICACY OF DIGITAL INSTRUMENTS

There is currently no conclusive evidence that many of the instruments mentioned in this thesis shall enjoy wide spread support because of ease of use, and efficiency. Rather, this is the expectation that is driving the research and development of electronic payment systems. Some of the mechanisms proposed thus far include Netbill, the OMI Payment Switch, CyberCash, DigiCash, First Virtual's Green Commerce Protocol, Netcheque/Cash, and MasterCard's and VISA's Secure Electronic Transaction protocol. The market for digital instruments is still terribly immature. Finding the proper business model or even the right pricing strategy is something some of these companies are currently trying to resolve.

While there is no conclusive evidence for the success of these services—many are just leaving the demo stage of development—I feel there are strong arguments for their success. Consider an Internet bank that I use for some of my own banking purposes. In answer to the question, "How do you make money?" the Security First Network Bank (SFNB) FAQ responds:

"We make money because our business model is far more efficient than traditional banking models. We have a "footprint" that spans the entire U.S. through the Internet. Yet all our Internet operations are located in Atlanta along with our banking

---

41 Some of the proposals that have fallen by the way-side or have been incorporated into other proposals include iKIP, Secure Transaction Technology, Secure Electronic Payment Protocol, Netscape's Secure Courier, the Simple Network Payment Protocol, the Hewlett Packard Payment Method, and Anonymous Internet Mercantile Protocol.
office in Pineville, Kentucky. A traditional bank would need to have fully staffed branches all over the country to achieve the same reach. As a result, our operating costs are far lower than a traditional bank and we can pass the savings onto our customers.

Subject to regulatory approval, we plan to offer brokerage, insurance, loans, and other financial services. Although we intend to generate fee revenue for these services, we anticipate the fees will still be lower than what is competitively available to you. Because our costs are lower, everyone benefits.

Much of this thesis has focusses on the competitive nature of commerce in a cryptographic economy. The success of digital instruments will consequently be dependent on improvements that they can offer over other instruments in terms of quality of service, efficiency and security. However, since most real world services are moving to employ digital networks as part of their underlying infrastructure, I assert that purely digital services will be at least as competitive. If the infrastructure proves to be more efficient in digital form, the user interface should be doubly so. For instance, a traditional bank may offer ATM or tele-banking services. SFNB uses the same banking infrastructure, and in addition to the costs saving at the infrastructure and user interface level, the user has the added capability to check bank statements on the WEB, conduct electronic payments and transfers, schedule automatic payments, and dynamically generate financial reports.

\[SFNB FAQ\]42

\[42\] See http://www.sfnb.com/caq.html .
4.1 TRUST AND PROTOCOLS

This thesis proposes that trust in an electronic market can be dealt with in one or more of three ways: 1) recoup losses from the risk of misplaced trust in the transaction fees, or pay for reputation services before a transaction so as to limit risk in a transaction, 2) develop protocols which eliminate the need for trust, and 3) shift trust. This chapter deals with the second group of protocols which limit the need of trust between parties for even sophisticated relationships. This class of protocols can be further subdivided into three levels. Pfleeger defines these levels as:

Arbitrated protocols, in which a trusted third party participates in each transaction to ensure that both sides act fairly.

Adjudicated protocols, in which a third party judges — after the fact — whether both parties acted fairly and if not, which party had not, and

Self-enforcing protocols, in which an attempt to cheat become immediately obvious to the other party and the protocol is safely terminated.

[Pf89:132]

Many of the schemes I discuss elsewhere in this thesis are in fact one of the first two types of protocols and it useful to repeat Pfleeger's analysis of each. Arbitrated protocols, he argued, have several disadvantages:

The two sides may not be able to find a neutral third party that both sides trust. Suspicious users are rightfully suspicious of an unknown arbiter in a network.

Maintaining the availability of an arbiter is a cost of the users or the network; that cost may be high.

Arbitration causes a time delay in communication, because a third party must receive, act on, and then forward every transaction.

If the arbitration service is heavily used, it may become a bottleneck in the network as many users try to access a single arbiter.
Secrecy become vulnerable, since the arbiter has access to much sensitive information.

[Fn89:131]

The difficulty with adjudicated protocols is that adjudication only comes after damage or cheating has already occurred. Many electronic payment or trust schemes are a combination of self-enforcing protocols, and protocols which rely upon financially arbitrated or adjudicated schemes. For instance, two untrusting principals may rely indirectly upon their trust in ecash to exchange value over the Internet. These protocols are not strictly self-enforcing because they rely upon the trust of a bank to redeem electronic tokens. Nor are they strictly arbitrated or adjudicated because a bank – the trusted third party – may not even realize that its services are used for arbitration or adjudication. With respect to personal privacy, a bank should not be aware of the details of any transaction beyond the fact that it validly creates, exchanges, or processes forms of payment in an efficient manner. Hence, the class of protocols which are self-enforcing but rely on an underlying financial system that may be an arbitrated or adjudicated system is defined as a indirectly arbitrated or indirectly adjudicated (collectively abbreviated as IAA) in the sense that they are often not acting actively to arbitrate or adjudicate the protocol which is employing their service. Banks and governments do well in the real world by providing a basis for others’ transactions, and there is no reason why they would fail to be equally successful in the digital world.

Regardless, this chapter focuses on self-enforcing and IAA protocols by reviewing some of the basic protocols and then discussing some of the current research that employs the protocols for more sophisticated transactions. 43

4.1.1 Bit Commitment

An obvious problem for networked transactions is how does one make a prediction before an event without revealing the prediction prematurely. Bit-commitment44 is the technique of committing oneself to a bit without revealing it. For instance, a stockbroker may wish to show that he knows whether a stock will fall (0) or rise (1) so a customer will contract with him. However, the stockbroker does not wish to disclose information prematurely. A simple solution to this problem is for the skeptical customer to present the

---

43 Two general references for many of the concepts discussed are Schneir’s Applied Cryptography [Sch94] and Denning’s Cryptography and Data Security [Den83].
44 This examples is taken from [Sch94:72].
stockbroker with a random number $R$, the stockbroker then encrypts his prediction

$(b\in\{0,1\})$ with $R$ in a secret key $K$ and sends it to the customer:

$$E_k(R,b) \rightarrow \text{Client}$$

When the stockbroker wishes to prove his prediction was accurate, he sends the key to the customer, who can then decrypt the message and determine if the random number is the number he originally sent. If so, the commitment of bit $b$ is valid.

### 4.1.2 Oblivious Transfer

One of the most significant protocols with regards to trust is that of *oblivious transfer*. This protocol has the – at first – rather strange property

1. Bob has a 50% chance of receiving a secret
2. Bob knows when he received the secret, Alice will not know if Bob received the secret.

What possibly could be the motivation for such a protocol? Actually, Blum discussed three applications of oblivious transfer in [Blue81], including fair coin flips, contract signing and digitally certified mail. Blum’s examples generally involved finding roots mod $n$ (for oblivious transfer) or factoring $n$ (for coin flipping) with a 50% probability given that the originator of the problem gives a hint that allows it to be solved 50% of the time. We present a common example for coin flipping that involves less mathematical insight – though [Blu81] and [Den82:116] both present the original scheme. The following is a common example of oblivious transfer based on Pfleeger’s [Pfl89:146] example:

1. Pete chooses two pairs of public encryption keys (four keys in all).
2. Nancy chooses a conventional encryption key $K$
3. Pete sends both public keys $(i,j)$ to Nancy, keeping the private keys for himself.
4. Nancy picks one of Pete’s public keys at random, call it $(k)$ and encrypts her private key in one of Pete’s public keys and sends it to Pete.
5. Pete does not know which public key of his public keys Nancy encrypted her private key in. However, the protocol requires that he decrypt it with one of his keys. The one he chooses, call it $(j)$, results in Nancy’s $K$, or (if he picks the wrong public key) a meaningless string of bits $MS$.
6. Pete encrypts the statement “Nancy wins – Pete” in the result of 5 ($K$ or $MS$) and sends this and the key pairs he chose $(j)$ to Nancy.
7. Nancy decrypts Pete’s message with $K$. If both Nancy and Pete choose the same key pair $(j=k)$ then Nancy has a copy of Pete’s message and wins, if $(j\neq k)$ then Nancy has nothing and Pete wins.
8. After the winner is known, Pete reveals his private keys so that Nancy can determine if Pete played fairly by making sure he sent the proper public keys.
4.1.3 Zero-Knowledge Proof Protocols

A zero-knowledge proof (ZKP) is a way to prove one knows something without releasing what one knows. For example they are explained in [QG89] by the following story:

Peggy, the prover, wishes to prove that she knows a secret magic word that allows her to pass between the wall CD. In order to prove to Victor, the verifier, that she knows the secret she enters the cave at A, and heads to either C or D. Then Victor enters and stands at B and shouts for Peggy to return from the right or left side. Peggy had a 50% chance of guessing which side Victor would choose. If this protocol were to be repeated multiple times, it would soon be obvious that Peggy does indeed know the secret.

Real protocols for zero-knowledge proofs rely upon the difficulty of solving hard computational problems and the ability to prove one problem is equivalent to another without revealing the original problem. In this example I use the hamiltonian cycles of graph theory [Blu86], the protocol works as follows:

1. Peggy, the prover, generates a graph of many random points in such a way that she can trace a single, continuous, and circular line through each of the points exactly once. This is her original hard problem H(1). She gives this to Victor the verifier.
2. Peggy moves the points about and changes the points labels so as to get a new graph H(2), but that is still easily solved with her knowledge of the solution for H(1). Only Peggy, who know the path of H(1) could solve H(2) in a reasonable amount of time. Also, only Peggy could prove that H(1) and H(2) are related or isomorphic.
3. Peggy gives H(2) to Victor.
4. Victor asks Peggy to prove H(2) is either isomorphic to H(1) or to give the solution to H(2). Neither of these queries result in Victor figuring out H(1) and they both can easily be responded to by Peggy.
5. This protocol is repeated multiple times; each repeat is a further assurance that Peggy knows the solution to H(1).

4.1.4 Cut and Choose & Secret Sharing

The protocol above employs a method known as cut and choose. This involves multiple plays of a protocol with 2 outcomes. At each play, Peggy would have a 50% chance of guessing the right outcome. The probability that Peggy is cheating by bluffing (guessing) is \( (0.5)^\text{number of plays} \). After 20 rounds, Peggy has only a one in a million \((1,048,576)\) chance of fooling Victor.
Secret sharing is the splitting of a secret among different parties. The secret may be recombined only when all of the parties cooperate to reconstruct the secret. Threshold schemes allow for some portion of the greater number of participants to be able to reconstruct the secret. For instance, a (3,5) scheme allows 3 out of the 5 people who have a piece of the secret to reconstruct it. A simple scheme is for Alice to XOR her secret $S$ with a random number $R$. She then gives the result to one friend, and $R$ to another. They may recombined the secret by XOR their pieces together such that $(S \oplus R) \oplus R = S$

4.1.5 Fair Coin Flips

In 1982, Blum published a more intuitive scheme for a fair coin flip in the aptly named Coin Flipping by Telephone: A Protocol for Solving Impossible Problems [Blu82]. Often two parties would like to generate a random bit — or string of bits — of information that each feels the other party did not coerce. For instance, if Alice and Bob wish to flip a coin over the telephone it would not do to have Bob call “heads,” and then have Alice tell him, “sorry, it was tails.” The digital form of a coin flip would be for Alice and Bob to each choose a random number and Alice sends her’s to Bob who XORs the two together. This isn’t fair because Bob can always manipulate his bit.

The scheme that does work is for Alice to commit to (or predict) the setting of a bit by Bob. Once she commits, Bob tells Alice his fancy, be it a “1” or “0”. Alice can then reveal her commitment and if she predicted Bob’s action properly she wins, if not, Bob wins.

4.1.6 Contract Signing

There is even a scheme for contract signing between two mutually distrustful agents on a network. Pfleeger described a protocol in real life that mirrors how one would accomplish this in cyberspace. Assume Charles and Diane live very far apart but both wish to fairly sign a contract they are mutually satisfied with. How would they go about signing the contract through the post? Pfleeger’s suggestion is for Charles to sign the first letter of his name to the contract and mail it to Diane, who would do the same and return it to Charles. Charles would then sign the second letter of his name and repeat the protocol until both had signed their complete names. The issue at hand is once Charles signed the first “C,” surely no judge would hold him to the contract, so he may do this without worry and as an act of good faith. Diane will see this and sign her first letter “D.” As the protocol continues, and assuming they both have equally long names, they both become mutually committed to the contract with ever increasing probability. “Since they are uncertain at exactly what point they become bound, both Charles and Diane must fear they are bound at each point.” [Pfl89:148] The techni-
cal protocols involve the same "baby-step" approach. A generalization of the technical protocols is for Alice and Bob to break their signatures into 100 little pieces and encrypt each piece in a DES key, then send them to the other. The protocol unfolds as Alice and Bob send each other one key at a time, and the signature slowly builds.\(^4\)

### 4.1.7 Other Standard Protocols

The following is a general description of many of the rather surprising protocols which can be implemented over networks. The protocols are painted with a rather broad brush and there are technical weaknesses or considerations one must account for most all of the protocols. However, the list does represent characteristics of self-enforcing protocols, or at least IAA protocols. For instance, there is a certified mail protocol that allows Alice to require a signed receipt from Bob if he reads the message. Other schemes include:

- **threshold schemes** \((m,n)\) – allow for \((m)\) people out of a total \((n)\) to reconstruct an escrowed key or digitally sign a message. The capabilities of the protocols can be quite sophisticated. For instance, one can require specific thresholds from specific groups, such as \((3,5)\) from group A, and \((2,5)\) from group B. Schemes for negative votes in which, "any qualified minority can prohibit the intended action" [Beu:491] exist as well.

- **secure elections** – retain the properties of normal elections in which voters must be authorized and can only vote once, votes are anonymous, and the process itself has confirmable integrity. This scheme generally relies upon the existence of a Central Tabulating Facility to tally the votes.

- **secure multiparty computation** – allow a group of people to mutually compute a function without revealing their individual inputs to the function.

- **anonymous message broadcast** – (or the Dining Cryptographer's Problem) is a ring of agents that can broadcast anonymous messages to the other users of the ring. The basis of the protocol is the flipping of coins and announcing to the group whether one's neighbors' coin had the same side come up as one's own coin. Information is created (messages are broadcast) by lying. If one lies about the "sameness" or "difference" of a neighbors' flip, this will create a disparity in the parity of the announcements by all the agents on the ring. This disparity communicates one bit of information, but the other cryptographers will not be able to determine who is creating the discrepancy.

- **mental poker** – is when two people are able to play a game of poker through the mail, much in a way similar to the oblivious transfer protocol between Pete and

\(^4\) Real protocols for contract signing are quite varied, and more complex than my description. For instance, see [EGL85].
Nancy. However, instead of encrypting only one bit of data the game generally
flows as follows. Pete shuffles the deck by encrypting each card in his own private
key. He then sends the results to Nancy. Nancy chooses five of the cards and re-
turns them to Pete, he decrypts them, and that is his hand. Nancy chooses five
more cards and encrypts them in her private key\textsuperscript{46} and returns them to Pete. Pete
decrypts the cards and returns them to Nancy. Now, they are only encrypted in
Nancy's key, which she can reverse and she now has her hand of cards!

Electronic cash schemes are considered to by a hybrid scheme of self-enforcing protocols
and IAA protocols and are examined in the following sections.

\textbf{4.1.8 Blind Signatures and Digital Cash}

Digital cash uses the cut and choose method in conjunction with blind signatures. Blind
signatures, developed by David Chaum [Cha87], \textit{blind} an object so that someone else
may sign it without knowing the real content. This is accomplished in RSA by multiplying
the original message by a random number, having the person sign it, and then dividing
the number out again.

Digital cash uses the blinding technique to give a user anonymity by making it difficult
for the bank to see the money order it is signing and consequently track it through the
network. However, the bank does not wish to sign a statement that it can not see nor ver-
ify! (Perhaps the money order asks for the redemption of an exaggerated value.) However,
if the cut-and-choose protocol is employed, it is highly probable that the bank is signing a
valid blinded money order if all the other money orders are revealed and found to be
valid.

A digital cash scheme that does not use the cut-and-choose method would begin when
Alice wishes to generate digital cash. She begins by creating a statement, "Please pay
$1000 to the holder of this certificate", as well as some random uniqueness strings and
identity strings. (This prevents \textit{double-spending} and cheating in cases where someone
tries to submit the same certificate twice.) Alice then gives this to the bank to digitally
sign, it does so, and takes a $1000 from Alice's account for itself. If it ever encounters
the money order again, it knows it has a valid obligation to pay it. However, if the bank
saw these uniqueness strings at the beginning of the protocol it could easily track where
and by whom the money was spent!

If the protocol uses the cut and choose protocol and blind signatures, anonymous ecash
that one can not double spend is possible. It works as follows, Alice generates 100

\textsuperscript{46} This is referred to as blinding.
anonymous money orders, blinds all 100 of them, and submits them to the bank. Within each money order, Alice also creates identity string pairs. To do this Alice first takes a string that unambiguously identifies herself to the bank and splits it into two strings using a secret sharing scheme. Alice does this 100 times, creating 100 pairs of the left and right side of the identity string. (Of course, one side without the other is useless as is discussed above in §4.1.4.).

The bank then challenges Alice to hand over 99 of the blinds, unblinds 99 of the statements, verifies the statements and the identity strings, and signs the last blinded money order, and returns it to Alice, who then may spend it where she wishes.

Later, when Alice is conducting a transaction with a merchant she must submit the money order to the merchant and randomly reveal one half of her identity strings. (The merchant challenges her with a 100 random bit string by which Alice reveals either the left or right side of each pair of her identity strings.) The merchant then shows the money order and the identity strings to the bank. The bank can look at its records and see if that money order was already submitted (double spent) by comparing the uniqueness string to those already spent. If the money order is a duplicate, the bank will refuse to honor the money order. Furthermore, if the identity strings are exactly the same as those submitted before, the bank will know that the merchant is at fault. If the identity strings are different, then with a high probability the bank now has the left and right side of at least one of the identity strings with which it may recombine and discover the cheater.

4.2 COIN RIPPING

Given the basic flavor of self-enforcing and limited protocols which require little to no trust between the parties, we can now concern ourselves with an interesting protocol which applies directly to the motivating scenario of two untrusting agents attempting to conduct transactions over the Internet. In Ripping Coins for a Fair Exchange [Jak95] Markus Jakobsson presents an example in which a person wishes to send a taxi driver to pick up some goods. The taxi driver does not wish to make a trip without payment, and the person does not wish to pay the taxi who may take the money and never return. His solution to the problem is described as:

Let us say that the taxi driver’s payment would be $100. Paying him half the amount before he leaves, and the rest upon delivery would not solve the problem. However, if you tear the $100 bill in two parts and give the driver one half before he leaves, and the other upon delivery, this would solve the problem. Neither you nor the driver will be able to use half a bill, since this will be without value, but when he returns and you give him the second half, he can tape the parts together and use the bill. If you are concerned that the taxi driver might not return just on spite, thus making you lose $100 although he will not get it, you can both tear bills and give each other halves in
order to keep each other honest. When he returns, you will exchange the second halves, and then perform a standard payment.

[Jak95:220]

The scheme has characteristics like those of electronic cash. The added operational requirements for the ripped coin are:

- a ripped coin is unredeemable.
- a ripped coin is still verifiable as coming from a well formed coin.
- the receiver can withhold the coin to penalize the payer for cheating.
- the receiver can verify that the two parts fit together properly.
- if a payer spends or rip-spends more coins than he has withdrawn he will be detected and identified.

Two technical requirements are that coins that can be ripped should not have larger technical representations than normal coins, and the bank will never know if a coin will be or has been ripped, its history is transparent to the bank for fair transactions.

The protocol works as follows. A coin is of the form \((x,s)\) where \(x\) is a random number known only to the drawer and \(s\) is the signature of the bank on a blinded \(x\). A coin is ripped when a transcript of the form \((x',c,x')\) is created from a interaction between the payer and merchant, where \(c\) is a challenge that the merchant made to the payer and \(x'\) and \(x\) are the results of that challenged applied to \(x\) and \(s\) with a secret function \(f\). The merchant then sends the \((x',c,x')\) to the bank. When the payer wishes to give the merchant the other half of the coin, he provides enough information to produce the transcript \((x,c,x)\) which is redeemable at the bank. It is important to note that \((x',c,x')\) is not redeemable at the bank, but it does prevent the payer from trying to produce other transcripts \((x',c,x')_n\) with other people because the bank will only redeem the coin for the first instance of \((x',c,x')\). Also, if any other transcripts \(- (x',c,x')_n -\) are created and submitted to a merchant, the bank will discover the double spending and the identity of payer will be revealed.

Interestingly, this protocol resembles the commitment that results from the security deposit protocol of the next chapter. However, instead of the protocol enforcing a non-action (prevent re-distribution) it coerces an action (the completion of a multiple-step transaction.)

---

47 The requirements of typical ecash schemes are: only the bank can create coins, blind signatures allow for user anonymity, receivers can verify a well formed coin, and over-spending is detectable by the bank.
TRUST SHIFTING

So, when a raging fever burns,
We shift from side to side by turns;
And 't is a poor relief we gain
To change the place, but keep the pain.
— Isaac Watts. Hymn 146.

In some instances it is possible to lessen the trust required between the principals of transactions. In others cases, significantly decreasing the amount of trust required is not possible; however, shifting the direction of the trust relationships inherent to the transactions may be useful. The explicit purpose of this chapter is to demonstrate a protocol with the capability to shift the direction of trust relationships. For instance, an intriguing problem of networked environments is how can one sell a user a digital object and trust them not to give that object to others? Embedding that digital object in a tamper resistant smart card obviates the need for trust between the two people (but each must trust the card). Short of that, the person selling the digital object (the merchant) is at the mercy of the customer not to redistribute the object. This chapter approaches this problem by developing a cryptographic protocol which changes the nature of the trust relationships involved. This is accomplished by developing a security deposit that creates a disincentive for the customer not to abuse the trust of the merchant.48 It is akin to the method of a land owner requiring a security deposit from his tenants.

Consider the case of a WEB merchant Bob who provides an information service. Alice wishes to pay for a password to access the service. The question I address is how does one prevent Alice from sharing her password or from reselling it to many people once she legitimately purchased it from Bob?

A rather pernicious way to prevent a person from sharing their password is for the WEB service to publicly display the user’s credit card number encrypted with the user’s access password. The publicly displayed credit card number hidden (blinded) in the user’s password is safe as long as the user does not reveal the password to anyone else. Hence,

48 Bruce Wilcox has suggested escrowing value with a third party who will adjudicate protocols for fair compliance. My protocol does not require the direct involvement of a trusted third party but on the stability of an underlying ecash protocol (an iaa protocol of §4.1).
this blinded credit card number acts as a security deposit to the degree that it provides an incentive for the user to not to abuse a specific policy regarding redistribution. However, it is important to note, that in the case of abuse, the WEB service does not recover its "loss" of providing service to two people for the price of one.

The approach of creating disincentives which move the risk and the nature of the trust relationship from the WEB service (who now is much less likely to be cheated) to the user (who if he cheats, may be betrayed himself when the friend that he gave the password to makes a grab for the credit card number) is the general method and is described in greater detail in the sections below. The most general variant of the protocol is described using RSA public-key as the primary mechanism. However, while the RSA scheme does reflect the operation of the protocol it is not the best scheme on which to base the protocol and later sections discuss technical aspects related to the security and complexity theoretic aspects of the protocol.

5.1 CHARACTERISTICS & SECURITY DEPOSITS

Since I wish to consider cases broader than just instances of passwords being improperly shared, I define the term characteristic as an authenticatable attribute of a user in a networked environment. For instance, Alice may be characterized as having a legitimate right to access Bob's WEB service, or as having the legal ability to drive in Maryland. Characteristics can be represented technically as a password, or attributes of an extended certificate as discussed in Chapter Six. With a valid characteristic one should be able to prove possession of a characteristic without revealing it.

In this chapter, I present a generalized method of creating a security deposit that:

1. Proves an agent has a characteristic.
2. Does not allow another agent to derive the characteristic from the proof.
3. Provides the agent with the characteristic a strong disincentive to share it with others.
4. Ensures that if an agent does share the characteristic, she is economically penalized and discovered by the issuer of the characteristic.

Those familiar with cryptographic protocols, will quickly realize that conditions one and two are the results of zero-knowledge proof protocols. By combining zero-knowledge proofs with blind signatures and ecash, I provide a protocol that satisfies all four conditions.
5.2 CRYPTOGRAPHIC PROTOCOLS

This chapter relies upon a variety cryptographic concepts. An overview of many of the concepts is provided in Chapter Four, otherwise I refer readers unfamiliar with the concepts to the original works as cited.49

5.2.1 Security Deposits and Characteristics using RSA

Consider the following protocol in which a characteristic is the private key of a public/private key pair:

1. Alice is given a private/public key pair by Bob for her interactions with Bob. Alice can then prove she knows the private key by signing challenges in the private key.
2. Alice states, "In order to prove I won't give anyone my characteristic (private key) I will encrypt a money order for a $1000 from my account with my public key." She signs this statement, and gives the statement and the encrypted money order to Bob.
3. Bob knows Alice's private key, he can confirm the validity of the security deposit and should post it publicly as a deterrent to Alice sharing her characteristic/private key.
4. If Alice gives her private key (characteristic) away, she would have to hope that person did not make a run for the $1000! Hence, within a crypto-economy the security-deposit provides a disincentive for cheating.

However, this is not general nor secure! Perhaps the characteristic is only valued at $1, and while Alice wants to conduct this transaction with Bob, she doesn't trust him! (Bob could take the money himself.)

A different algorithm would be to employ the cut-and-choose protocol yet again:

1. Alice generates 100 private/public key pairs and encrypts a security-deposit statement, "In order to prove I won't give anyone my characteristic (private key) I will encrypt a money order for a $1000 from my account with my public key." She then signs the statement and encrypts the money order in each of her 100 public keys. She gives all of this to Bob.
2. Bob challenges her for 99 of the key pairs and confirms all the statements are valid. If they are valid, the left over key pair can be used to authenticate Alice and act as her characteristic. The encrypted money-order is the security deposit.
3. Since the encrypted money order acts as a security-deposit, Alice will not wish to give it away.

However, while no one but Alice knows her private key the generation of these key pairs may contribute significant overhead to the protocol. A subtler problem in using RSA in this protocol stems from the general rule that one should not sign arbitrary messages

49 Or to texts on these topics including [Sch94, Den82, Pfl89].
using RSA. (This is sometimes called the chosen-ciphertext attack.) For instance, to prove that Alice knows a characteristic (her private key) she will be challenged to sign certain statements or random strings. This way anyone with her public key can verify that Alice is indeed using the proper private key by checking the signature. However, since in RSA one can blind a message before having them signed, people could submit the public-key encrypted money-order blinded in a random number as a challenge to Alice for authentication. Alice would sign the seemingly random challenge, but she is actually decrypting the money order! The common solution for authentication given this problem is for Alice to sign a hash of the challenge rather than the challenge itself and this prevents her from being tricked into decrypting things she previously encrypted but that are now blinded. However, because of the importance of using this protocol to repeatedly authenticate users a stronger approach may be taken by employing zero-knowledge proofs as the blind on the security deposit.

Reiterating the characteristics of before, a good scheme should adhere to the four principles I wrote of earlier:

1. Proves an agent has a characteristic.
2. No other agent can derive the characteristic from the proof.
3. The agent with the characteristic has a strong disincentive to share it with others.
4. If an agent does share the characteristic, she is economically penalized and discovered by the issuer of the characteristic.

The exposition of the protocol employing the RSA scheme demonstrates the operation of the security deposit scheme and generally meets the above requirements. However, readers interested in the use of zero-knowledge proofs that strengthens the protocol with regards to requirement (2) may wish to read the next section. (For instance, returning to the cave example of §4.3, Alice would encrypt the money order in the secret password that lets her pass through the wall.) Others may skip to §5.2.3 on Information Economics.

5.2.2 In Depth Protocols

We have already specified the four requirements of a good security deposit scheme, this section addresses the technical requirements for a protocol to be successful in meeting the four requirements above. The following sections refer to a number of general concepts regarding zero-knowledge proofs and complexity theory. Useful references on these topics include an introduction to algorithmics [Har87] and the “bible” of algorithmics [CLR90]. Specific to complexity theory, Garey and Johnson’s [GJ79] provides a good guide to the theory of intractability – including a historical treatment to the development
of the field and the nomenclature – and an interesting appendix of annotated NP-complete problems. [vanL90] also proved to be an invaluable reference, particularly Johnson's chapter [Joh90] on "A Catalog of Complexity Classes." See [Gol90]\textsuperscript{50} for a very complete treatment of ZKPs.

\textbf{5.2.2.1 Protocol Using Hamiltonian Cycles}

We have already mentioned some of the disadvantages in using RSA as described. The first was the overhead in generating 100 or more public-private key pairs, the second was the arbitrary signing or chosen ciphertext attacks. Another consideration is that the RSA implementation does not meet the stronger characteristics of a zero-knowledge proof (ZKP) system in which:

1. the prover can demonstrate conclusive knowledge of the proof only if he knows it
2. the verifier learns nothing about the proof other than the prover knows the proof and that
3. no other knowledge can be gained from the prover which the verifier could not obtain for himself.

The RSA implementation fails the third condition because of (at least) the arbitrary signing attack, which grants the verifier a certain amount of knowledge – the signed challenge – that he could not have generated himself with any ease. Generally, the question of what constitutes knowledge and what is zero-knowledge is problematic. But, "when discussing zero-knowledge proofs, we avoid the first question (which is quite complex), and treat the second question directly. Namely, \textit{without} presenting a definition of knowledge, we present a generic case in which it is certainly justified to say that no knowledge is gained."

[Goldreich then presents a problem from graph theory in which:

\ldots we can say that Bob gained knowledge from the interaction if his computational ability, concerning the publicly known graph, \textit{has increased} (i.e. if after the interaction he can easily compute something that he could not have efficiently computed before the interaction). On the other hand, if whatever Bob can efficiently compute about the graph after interacting with Alice, he can also efficiently compute by himself (from the graph) then we say that Bob gained no knowledge from the interaction.

[Goldreich then presents a problem from graph theory in which:

A similar, informal statement regarding graph isomorphism in [GMW90:14] is that the graph isomorphic protocol is, "Intuitively \ldots 0-knowledge since whatever the verifier receives is 'useless', because he can generate random isomorphic copies of the input

\textsuperscript{50} Which is currently unfinished, but available on the WEB at http://theory.lcs.mit.edu/~oded/frag.html
graphs by himself.” However, if one were to consider the RSA scheme in which challenges to the prover were not strings chosen by the verifier, but rather a random string generated from a coin flipping protocol, the information returned to the verifier would be equally as useless unless there is a cipher-text only attack on RSA.

Regardless, extending the security deposit protocol to work with zero-knowledge proofs would certainly be worthwhile, and could be done (incorrectly) in the following manner:

1. Alice generates a hundred hard problems (100 graphs G(x)) for a zero-knowledge proof. For instance, Alice can easily and quickly generate each hamiltonian cycle problem by constructing a random line that connects a series of dots that are "laid down" as the line is drawn in a plane.

2. The graphical description (a random bitstring) of the lines in each of the hundred hard problems is then used to blind 100 promissory notes.

3. The hundred blinded promissory notes and the corresponding graphs are submitted to a bank for signing. The bank asks for the blinds for 99 of the problems and checks the validity of the promissory notes and the validity of the blind as a solution to each of the graphs. If all 99 check, the bank signs the unchosen promissory note which is appended to the single hard problem (G) and signs this combined statement. The bank keeps this certificate in a public repository.

4. This is Alice's characteristic. She can now prove that she knows the answer to the hard problem G without giving the solution away, not even the bank knows the answer! If Alice were to give her characteristic away, she would potentially be exposing the promissory note and incur a loss!

However, this scheme has a problem in that Alice could generate multiple Hamiltonian paths for a single graph. In such a case, Alice could give each solution to a different friend, who would be capable of proving they knew a solution to G, without being able to grab Alice's security deposit, which is encrypted in the particular solution that Alice kept for herself.

As such, a different type of problem with a unique solution is required. The graph isomorphism problem is considered. Two graphs $G_1=(V_1,E_1)$ and $G_2=(V_2,E_2)$ are isomorphic if there exists a bijective function $f : V_1 \rightarrow V_2$ such that any two vertices $(u,v) \in E_1$ iff $(f(u),f(v)) \in E_2$. The problem of proving that two graphs are isomorphic, or non-isomorphic, can be used in ZKP protocols [GMW90:13] and can be employed in the security deposit protocol. Alice would keep the mappings $(f : V_1 \rightarrow V_2)$ between the two original graphs secret and would be her characteristic with which the money order is encrypted in. However, both graph isomorphism and hamiltonian cycles problems are not very practical for this protocol. Methods for generating hard hamiltonian cycle and graph

---

51 Bijective is both injective (or one-to-one) and surjective (or onto).
isomorphic problems are not well understood. In *Computers and Intractability* Garey writes that the hamiltonian cycle problem:

Remains NP-complete (1) if G is planar, cubic, 3-connected, and has no face with fewer than 5 edges, (2) if G is bipartite, (3) if G is in the square of a graph and (4) if a Hamiltonian path for G is given as part of the instance. Solvable in polynomial time if G has no vertex with degree exceeding 2 or if G is an edged graph.

References Omitted [Gar79:199]

Hence, while it may be appropriately hard – and Alice will always wish to create hard problems so people cannot cheat and claim her security deposit – they can be created with multiple solutions which would be unacceptable to the party that is accepting the security deposit.

Graph isomorphism on the other hand is described as:

The problem remains open even if $G_1$ and $G_2$ are restricted to regular graphs, bipartite graphs, line graphs, comparability graphs, chordal graphs, or undirected path graphs.... Solvable in polynomial time for planar graphs and inverted graphs. The problem is in NP $\cap$ Co-NP for “arc transitive” cubic graphs.

References Omitted [Gar79:285]

Hence, producing hard isomorphic graph problems may even be more difficult and though while the mapping function is bijective, this does not equate with uniqueness and as such this problem is not appropriate either.

5.2.2.2 A Proper Complexity Theoretic Problem

However, after considering these examples, we can now enunciate some complexity theoretic requirements for the problem that is at the center of the security deposit protocol:

1. the problem should be difficult (intractable) – this should extend for the average case (or even the best case) instance of the problem and not just the worst case.
2. the solution must be unique
3. it must be transferable to a ZKP protocol.

Fortunately, there is such a class of problems.

ONE To satisfy the first requirement we would like to have a problem which is hard enough to break so that people will not be able to figure out Alice’s proof and claim her security deposit. The complexity class of problems (NP – P), all the nondeterministic polynomial time problems less the polynomial time problems, would provide such a class. As such, NP-complete problems would be appropriate, but NP-hard problems that are as intractable as NP-complete but not in NP would not be since their proofs cannot be checked in polynomial time.
Johnson in [VanLee] defines a class of problems that operate on *unambiguous Turing machines* (UTM)\(^{52}\) as:

FUP (UP)\(^{53}\) is the class of all partial functions computable by polynomial-time UTM.

\[\text{Joh90:111}\]

Hence, given (1), the appropriate problem would come from the (FUP – FP) set of problems, an example of which is an augmented DISCRETE LOG in which given a prime \(p\), a primitive root \(a\) modulo \(p\), and an integer \(b\), \(0 < b < p\) and proofs that \(p\) is prime and \(a\) is a primitive root, find the unique integer \(c\) s.t. \(a^c \equiv b \mod p\). In such an example \(c\) is one’s certificate. These problems are not known to be NP-complete, but this is not a requirement.

In [GMW90] the authors demonstrated that all problems in NP have zero-knowledge proofs. Demonstrations of DISCRETE LOG ZKPs are provided in [CEG86] and [CEGP87]. Examples of these protocols and Koyama’s RSA ZKP\(^{54}\) protocol can be found in [Sch90:401-403]. With many possible problems to choose from, it is not unreasonable to assume that the security deposit protocol could be implemented with the proper problem with the difficulty of the same order of magnitude of implementing the underlying ecash protocol.

Hence, the appropriate complexity theoretic problem provides a proof (sometimes call a certificate) that is used to blind a money order or batch of digital cash, and the proof remains a secret unless its user purposefully divulges it to another.

### 5.2.3 Information and Cryptographic Economics

With the general protocol described, one can now turn to the environment in which the transactions occur and examine the various scenarios in which people may attempt to cheat the system. This section is titled "Information and Cryptographic Economics" because it relies upon the importance of reputation, trust, value, arbitrage and disincentive in a free market between many anonymous agents. To show how these various concepts relate to the protocol I now examine the general economic scenario in which the protocol would be employed.

---

52 A UTM is “a NDTM [Non-deterministic Turning machine] that for each possible input string, has at most one accepting computation.”[Joh90:111]

53 Where, the “F” denotes search problems, and the lack of the F denotes a decision problem.

54 At last, we can use RSA after all!
Consider the case where Alice wishes to buy access to Bob's homepage. The cost of the access is small, maybe $5 for life, but Bob does not wish to give it away for free, nor does he want people to cheat. Also, consider the fact that Alice does not wish anyone to know that she has access to Bob's page and creates a pseudonym of Anne. Various levels of anonymity can be employed in this environment but for the purposes of this example, only Alice will be anonymous. Knowing that Bob requires $5 for access to his page, and a $100 security deposit Alice goes to the bank for $5 money order and a $100 security deposit using the protocol above. At the end of the protocol the bank generated a security deposit that includes, (1) a statement that "the possessor of the zero-knowledge problem H(1) (the solution of which is the characteristic) has committed herself to a $1000 loss if she reveals it to anyone else", (2) the problem H(1), (3) and a signature on the concatenation of (1) and (2). It then posts this statement on a public board and returns it to Alice.

Alice then "transfers" the $5, the characteristic, and the security deposit to herself as 'Anne' and Anne contacts Bob. Bob checks the signature on the security deposit, cashes in the $5 and agrees to use the zero-knowledge proof as the proof that Anne knows the characteristic. (He then uses this as an authentication service for his home page.)

Now, what are the problems with the scenario?

What if Alice cashes in her security deposit and continues to use the characteristic?

In this example I have been using a bank as a security deposit service. One of the functions of the bank will to be maintain the public listing. Perhaps, Bob can be asked to be notified if the security deposit for a certificate is ever cashed (he will then discontinue the service), or maybe he himself will have to check the listing every once in a while. This method requires an on-line third party (the bank) as does the underlying ecash scheme.

How does Alice cancel her characteristic?

One cancels one's security deposit by merely claiming the security deposit. A consequent of which is that one then looses access to the service.
What if Alice gives her characteristic to others?

In such a case, the others would have a strong incentive to go get the security deposit! An appropriate security deposit should be required for the service provided. If Alice gives Claire and David the characteristic, and Claire grabs the security deposit, Bob will be notified that the characteristic is bad. Furthermore, if David or Alice attempted to get the security deposit after Claire, the protocol would result in a "double-spending" and Alice would be revealed as the identity that distributed the characteristic – her reputation from this would suffer greatly.

How much should the security deposit be for?

I was not able to find much literature on the setting of security deposit rates, hence I do not address this topic in great detail. However, consider if Alice were to commit herself to a promissory note of $10 (and pay the $5 fee for access), she then would only have to sell the characteristic to four people at $4 to make a profit of $1. The amount of the security deposit should exceed the amount of money a crook could make by selling the characteristic at a lower price. The importance of reputation in this scheme is of immense importance. Perhaps, as time passes and an agent accumulates trust, the amount required for a security deposit of that trust-worthy agent would decrease. Also, consider since people know they are buying illegal copies of a characteristic, they know other untrust-worthy people are probably doing the same as well! And if one of those people go for the security deposit, then everyone’s $4 black market characteristic is wasted! Hence, a black market for criminals would be infeasible unless reputation and trust networks develop within the black market.

How rich does Alice have to be to buy a piece of bubble gum and shoestrings?

If Alice is a legitimate user, and has many characteristics, she may have committed to thousands of dollars in security deposits! If she is an honest user, that money is safe, and she may reclaim it in time, however this could require her to tie up a large part of her capital. One potential solution has already been mentioned: trust. One can forsee the formation of credit agencies that will vouch for the niceness of a new agent in the world if that agent does not have all the required money. Or one established pseudonym could anonymously but validly vouch for another. The fraudulent rate of the agents would determine the feasibility of this scenario. Also, most characteristics could have expiration dates embedded in the original certificate that the bank signs. This would enable the security deposit capital to be fluid and dynamic as the needs of the agents change. It could also be possible to pay the security deposit paid back over time as the information devalues, or the trust in the agent increases.
There is also the possibility of allowing the same security deposit to be used for multiple relationships. For instance, the fact that Anne can solve problems isomorphic to H(1) would not only be used by Bob, but by Eric and Frank. I expect this scenario is similar to determining the right price of a security deposit, but this is as yet unclear. The pricing of these type of services is an area for future research. However it is important to remember that not only is the loss of money a deterrent, but so is the "double spending" detection of cheaters.

Is there really no way for Alice to give her characteristic to another?

Another area of further research is the existence of secondary markets for characteristics and arbitrage. I have already mentioned the possibility of black markets forming based on their own networks of trust. How would one accomplish such a thing – with security deposits of course! For instance, maybe Anne could sell her characteristic to Claire, if Claire committed herself with a security deposit to never give the characteristic to someone else. In such cases, arbitrage would move the original market price of the object and the security deposit to a value more in keeping with the actual demand!

In this chapter I have presented a protocol for employing non-transferable characteristics in a cryptographic-economy by using digital security deposits. While the protocol itself may be interesting, it leads to a truly fascinating view of an economy where the value of information and characteristics are not determined by a neo-classical methodologies, but by a very complex and potentially emergent system of exchanges between anonymous agents exchanging value over a network.
Of greatest practical concern to those interested in the topic of trust in an electronic environment is how does one effectively represent trust relationships and manage them in the distributed and rather chaotic environment of the Internet? The usual way of representing trust relationships in a networked environment is through the use of "certificates" which historically have asserted that an identity (such as an email address) is linked to a specific public key. For instance, my certificate is signed by a number of individuals who vouch for the fact that the public key is linked to my email address reagle@mit.edu.

Trust management is a fascinating topic because the technology (which is interesting itself) is entangled in numerous policy issues. Creating certificates are relatively easy. Creating a certification structure (or public-key infrastructure (PKI)) that is extensible, easily managed, robust, and efficient is difficult. The difficulty is reflected in the question, what should a "good" public-key infrastructure do? Should it map identity to public keys? Should it handle attributes beyond identity? Should every person in cyberspace be issued a unique identity (UID) to be used exclusively for every transaction? Who would issue that ID? Furthermore, it is difficult to technically represent trust relationships that are both expressive enough to represent realistic relationships, and simple enough to safely construct the proper (secure) expression.

This chapter reviews the representation of trust with certificates by first presenting some of the basic concepts of public key technology and digital signatures; the basic X.509/PEM scheme; and then by examining many of the current proposals and issues regarding public key and trust management. However, the reader should keep in mind

---

55 The usage of the term "certificate" is different from that used elsewhere in this thesis when it is used to describe conclusive evidence of a solution to NP-hard problems.

56 "Public-key certificates such as X.509 certificates determine the skeletal structure of trust within a distributed public-key system." [RSA93a:7]

57 In the PGP scheme this is not exactly true since the user may change the identity field after signatures are placed on the key.
that many of these issues are being actively debated\textsuperscript{58} as this thesis is being written with new ideas, papers and proposals arriving daily to the Web or in the email.

6.1 PUBLIC KEY CRYPTOGRAPHY AND DIGITAL SINGNATURES

A difficulty of any cryptographic system is how to distribute the keys used for encryption and decryption.\textsuperscript{59} In symmetric systems (with a single shared key) one has to trust a secure mechanism, perhaps a human courier, to exchange keys. Such mechanisms tend to be unreliable and expensive. With public key cryptography one can send a secure message to anyone in the world using their public key, and only the person with the matching private key may decrypt the message. For instance, in Figure 6-1 Alice can get Bob's public key from a public key server, or from Bob himself and send any arbitrary message $M$ to Bob encrypted in Bob's public key. The encrypted message is represented as $E_b(M)$, and only Bob will be able to decrypt it with his private key. Bob could send a secret message, $E_{A}(M')$, to Alice using the same technique. The other interesting feature of public key cryptography is the ability to sign a message. Since only Bob knows his private key, if he performs a transformation on a message (or signs it) with his private key, then anyone in the world can use the matching public key to verify if the transformation is a valid signature.

However, public key technology does not fully solve the key management problem. For example, what if Bob writes beautiful poetry and posts it to the Internet and Alice falls in love with Bob and wishes to write him a confidential love letter? She must get

\textsuperscript{58} A number of recent works in progress, reports, symposia, and the like have been circulating amongst the security community on the topic of PKI and certificates. These topics are receiving a great amount of attention on mailing lists such as cypherpunks, and the Simple Public Key Infrastructure (SPKI) Working Group. Others lists that discussed the topic include perm, www-buyinfo, e-payments, and ietf-payments.

\textsuperscript{59} For a brief description of trust and cryptography with respect to an applied topic (email) see [BBC90].
his public key. This is not trivial. There could be a rather jealous and lonely person by
the name of Clark (the cheater) who wants to masquerade as Bob. The central question
is how does Alice know that the public key she receives from a person claiming to be
Bob is really the key of the poet Bob, and not the impersonator Clark? This problem
is that of identifying the other principle. Once a person is properly identified, then they
are authenticated on the basis of the identification. An advantage of public key cryptog-
raphy is that one does not need to share one’s private key with a trusted third party or
courier for encryption or authentication. However, for key distribution it is useful to use
a trusted third party for the purpose of identification. One can use a trusted agent to
certify that a given public key really maps to a given identity. The solution to our poetry
example is to depend on a trusted third party, perhaps the moderator of the poetry con-
fERENCE, to direct Alice to the real Bob’s public key and email address. The moderator
in effect certifies that the public key and email address correspond to the email address of
the authenticate Bob. How does Alice know that the moderator is indeed who he ap-
ppears to be? At some point Alice will have received a public key that she trusts,
perhaps a friend gives her his public key in person. If this friend signed the moderator’s
public key, she can trust that the moderator is authentic.

6.2 CERTIFICATES AND X.509/PEM

Certificates make “assertions” in a networked world. While most digitally signed state-
ments could be thought of as an assertion, certificates can be thought of as assertions
with relatively long life times and that come from a third party. For instance, Alice
may show a police officer the license (certificate) the DMV issued her. Röscheisen char-
acterizes a certificate as a “third party-opinion about an attribute of some entity,” and that
each assertion on a certificate (if there is more than one) has four parts:

- an identifier for the certifying authority (issuerID).
- an identifier for the subject of the assertion (subjectID).
- the attribute about which one is making an assertion.
- the value of the attribute.\(^\text{61}\)

If an attribute is used to determine if a request for a particular action or permission, this
process is called authorization. [ZB95:4] Attributes that are used for authorization can be
specified in terms of positives (approved access), negatives (denied access), or both
(binurated attributes).

\(^{60}\) For much of this discussion I combine the discussion of X.509 with PEM – though there are
differences which I discuss.

\(^{61}\) These four elements are based on the four expressed in [Rös95:1].
Certificates are used by the X.500\textsuperscript{62} [CCITT88a] standard for the management of a large directory of users on a distributed network. X.500 could provide a database by which everyone's email address could be accessed by anyone else on the network. A problem of such a service is that the directory can not be managed statically; changes in email addresses or telephone numbers have to be updated. To do this the owner of a directory entry would have to first authenticate themselves to the directory server so malicious changes could not be initiated by others. Also, the networked environment may be complex, with multiple geographical and operational domains – users in Britain and California would not use the same directory server. Consider the case where a user in Britain wishes to communicate with the user in California. This problem is similar to a poetry problem since merely possessing a key and asserting an identity is not enough. Rather, one needs a trusted party who's public key one knows one has a valid copy of to sign or certify other peoples' public keys. As stated, two users might not have an immediate, mutually trusted friend to rely upon. In such a case, the users in California and Britain each should have an immediate (local) trusted party whom they can trust. Then each friend (perhaps the US and British governments) would have a trusted friend at the UN. It is through this hierarchical approach that authentication in a distributed environment occurs in the X.509/PEM standard.

X.509: The Directory—Authentication Framework [CCITT88b] is the standard that defines the means by which certificates are properly formed and employed in the authentication of a user to a directory server. It is often referred to as an identity certificate scheme because the certificate is a signed statement that maps an identity to a public key.

The entity which provides this mapping is the certification authority (CA). Within X.509, every lower level CA (the immediately trusted agency) is also responsible for ensuring each principal (user) within its domain has a unique distinguished name (DN) network wide. The distinguished name structure is specified within X.501 and is of the format:

\[
\text{C=USA, L=Boston, O=MIT, OU=RPCP, CN=Reagle}
\]

Where C=Country; L=Location; O=Organization, OU=Organization Unit; CN=CommonName

Privacy Enhanced Mail (PEM) was an attempt to create a standard for secure Internet mail. While it failed to be adopted in any practical sense, it provides a basis for much of the discussion regarding PKI today. The certification hierarchy proposed in RFC1422, is a

\textsuperscript{62} X.500 is entitled Recommendation X.500: The Directory—Overview of Concepts, Models and Services, and is the first section of 8 of ISO 9594 – X.500 is often used to denote the whole class of standards of ISO 9594-8.
subset of that allowed under X.509. The significant difference is that while under X.509 any CA can grant a certificate to any user, PEM requires that a CA may only sign a user's certificate if that user is a member of the CA’s domain — this is called subordination and is accomplished by prefixing the user's DN with the CA’s name. The enforcement of unique CA names is a function of a higher level certification mechanism — the IRPA which is discussed later.

If any two users within the same domain wish to exchange certificates they must merely query their mutual CA to obtain the proper certificates (as do the two figures in Figure 6-2 Simple Hierarchy along the solid line.) When two users are not of the same domain, the hierarchy must be traveled so that the proper certification path (CP), or trusted path, is formed (represented by the dotted line).

The PEM scheme also defines roles for a root (called the Internet Policy Registration Authority (IRPA), or Top Level Registration Authority (TLCA)) and Policy Certification Authorities (PCA):

The IRPA establishes global policies, described in this document, which apply to all certification effected under this hierarchy. Beneath the IRPA root are Policy Certification Authorities (PCAs), each of which establishes and publishes (in the form of an informational RFC) its policies for registration of users or organizations. Each PCA is certified by the IRPA. (It is desirable that there be a relatively small number of PCAs, each with a substantially different policy, to facilitate user familiarity with the set of PCA policies. However there is no explicit requirement that the set of PCAs be limited in this fashion.)

[RF1422]

As such, the IRPA certifies PCAs according to the following policy classes:

- (identity), (scope), (security and privacy), (certification policy), (management),
  (naming conventions), (business issues), and (other).

A PCA then certifies CAs, and the CAs certify their domain of users (often an organization) for which it maintains a unique DN name space. It is important to note that a CA or PCA may have multiple logical roles:

Note that in this architecture, only PCAs may be certified by the IRPA, and every CA’s certification path can be traced to a PCA, through zero or more CAs. If a CA is certified by more than one PCA, each certificate issued by a PCA for the CA must contain a distinct public component.
Figure 6-3

CA Heirarchy represents the case in which Alice, a MIT student, wishes to talk to Bob at OpenMarket.™ If Alice trusts the IRPA, then the certificate would be of the form:

\[
[[\text{IRPA, PCA.COM}]], [[\text{PCA.COM, CA.OMI}]], [[\text{CA.OMI, Bob}]]
\]

Where each enclosure \([x,y]\) is a signed certification by \(x\) of \(y\). When Alice considers this certification path, she must also consider the policies of the CAs. In our example, perhaps a job offer is being made and Alice feels that the policies of the commercial PCA are robust enough to trust that she is being offered a job by a real company and not a rogue. However, if Alice gets a piece of email saying she won $10,000 and she must only pay $10 to claim her prize from a person who is certified by a persona CA (which issues certificates under an anonymous policy) she should think twice about sending her $10.

It is often the case that one may wish to revoke a certification that one has placed in the public space (perhaps one loses one’s private key). In this case revocation certificates can be distributed. Or, before accepting a certificate, servers may wish to check a Certification Revocation List (CRL) to make sure the certificate is still valid. Revocation is further discussed in §6.3.4.

6.3 CERTIFICATION AND TRUST STRUCTURES

Given the basic framework of technically representing trust, I now focus on some of the specific technical and policy issues of trust and certification structures.

6.3.1 Identity

Earlier I stated that the most common form of the certificate is the identity certificate which links an identity to a public key, this key is then used for authentication. A problem of the networked world is that many forms of identity tend to be unstable. A certificate that asserts an email address is linked to a public key is invalid when the email address changes. The question of what is the proper tag to “hang” ones identity off of is problematic, but two things are clear. One, if a tag means something (it has semantic meaning), then it is not likely to be stable. Email, employment, and location all change.
Recent discussions between members of the PKI groups have mentioned the use of URL's. While URL's tend to be as unstable as email addresses, research regarding URX's\(^6^3\) could provide leverage on the naming in certificates problem. Second, one need not require any semantic meaning to exist in the tag, rather all one needs is a unique tag that is stable. For instance, Bob could have a certificate that is bound to a random number, rID, and with this certificate Bob conducts commerce with merchants. However, this does not mean this the rID forces anonymity upon Bob. He should be able to have name servers certify that his name maps to rID. But, if his name changes, or if there are problems with many people have the same name, all other attributes need not be changed as long as the rID is unique. Even if Bob does wish to be anonymous, he does not lose his capability to have persistent relationships with other users. If a merchant is willing to certify that he has done SX amount of valid commerce with the rID, then one has created a pseudonym. Pseudonyms are anonymous identities with a persistent reputation. For instance, even though Bob is anonymous, the merchant knows he has conducted valid commerce with rID a number of times and may be more willing to do business with him in the future. Since rID has no basis in the real world, it is very stable and can be kept as long as it proves to be useful.

An added difficulty in key management protocols is complicated by the naming schemes on which they rely. Some of the above issues are represented in the chart compiled from Trc94:

\(^6^3\) The class of problems that include Uniform Resource Locators (URL), Names (URN), Characteristics (URC), and Information (URI). More information can be found at http://webreference.com/urx.html.
Table 6-1
Benefits and Problems of DN/CA Naming Subordination

<table>
<thead>
<tr>
<th>Benefits of X.509 DN/CA Naming Subordination</th>
<th>Problems of the X.509 DN/CA Naming Subordination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Automatically enforces unique name assignment</td>
<td>The CA is responsible for data that it has no</td>
</tr>
<tr>
<td></td>
<td>control of – a CA has to put the data in a Direct-</td>
</tr>
<tr>
<td></td>
<td>tory System Agent that is operated by another</td>
</tr>
<tr>
<td></td>
<td>organization</td>
</tr>
<tr>
<td>Protects a user from a rogue CA (which might</td>
<td>Any modification of CA data are subject to</td>
</tr>
<tr>
<td>generate the keys and corresponding certificate</td>
<td>significant time delays and effective is not</td>
</tr>
<tr>
<td>to impersonate one user)</td>
<td>possible</td>
</tr>
<tr>
<td>Simplifies the public key verification (unwrapping</td>
<td>Moreover, one who has been given the right to</td>
</tr>
<tr>
<td>the certification path)</td>
<td>manage a certain part of the name space,</td>
</tr>
<tr>
<td></td>
<td>automatically gains control over the corres-</td>
</tr>
<tr>
<td></td>
<td>ponding part of the certification hierarchy.</td>
</tr>
<tr>
<td>Implies which CA is authorized to certify a user</td>
<td>Finally, the existence of only one TLCA does not</td>
</tr>
<tr>
<td></td>
<td>reflect a real life situation, e.g. which coun-</td>
</tr>
<tr>
<td></td>
<td>try in the world will be authorized to operate a</td>
</tr>
<tr>
<td></td>
<td>TLCA for the whole world?</td>
</tr>
</tbody>
</table>

Table created from bullet points of [Tru94:74]

Pretty Good Privacy (PGP) is the most widely known privacy application with which users have had a chance to manage trust relationships and keys. The identity of a PGP principal is generally an email address. However, there is no restriction on what the identity tag must be, it could be an anonymous remailer return address, or even a useless string. This anonymity and the lack of a strict CA hierarchy and naming scheme has led to its wide use adoption over PEM. PGP’s trust management is discussed further in §6.3.3..

6.3.2 Characteristics (Attributes)

A user requesting access to a file on a system (or perhaps to change an X.501 entry) would employ X.509 as follows. The user identifies himself with a certificate, on the basis of the certificate the service authenticates the user, and if the authentication is valid it checks an access control list (ACL, a database of permissions) to see if the identified and authenticated user has legal access to the file or requested action.
Given the fact that one can carry about a signed statement by a server – which that server can later validate – a useful feature of a certificate would be for Alice to carry around any number of assertions beyond her identity within her certificate. For instance, instead of having to identify and authenticate herself to a service, within her certificate she could carry a statement signed by the service which states “I should give Alice permission to X when requested.” The service can easily check that it had actually signed this statement, and that Alice is authentic. It need not consult an ACL. Such a unforgeable token that is used to gain access to an object is called a capability in traditional computer security literature [Pl89:218].

Since I wish to consider general statements beyond access permissions, for instance, “Alice has a GPA of 3.5.” In the previous chapter I defined such assertions as characteristics, within the domain of this topic, they are often referred to as attributes as well. A certificate that includes characteristics beyond identify are called general certificates. While few people would argue with their usefulness, there are disagreements over how they should be technically implemented.

64 In an email to the SPKI list, Ron Rivest recently proposed a way of viewing certificates, those which map a public key to a “real world” entity (like an identity), and those which only correspond to relationship in “cyber space” (like permissions). The interesting result is that not all public keys need to have real world principals, and that permissions generally are granted to public keys and not real world principals.
There are a number of benefits of a general certificate over that of the X.509/PEM scheme. A simple one is that the policy regarding the certification is embedded in the statement itself. A general certificate can state, "Alice can drive in MA", a X.509/PEM scheme might require that Alice be checked against an ACL, and then the policy under which she was certified might be checked. Carl Ellison parodies this with a quote from the Hitchhiker's Guide to the Galaxy when a character is told of the earth's imminent disaster, "There's no point in acting all surprised about it. All the planning charts and demolition orders have been on display in your local planning department in Alpha Centauri for fifty of your Earth years..." [El1996:3] Furthermore, if one wishes to have different certificates with different meanings a new certificate and public key need to be created under each different PCA policy. This leads to structural complexity and unneeded work [El1996:3].

One of the most contentious issues is how to technically represent general certificates. New schemes and extensions to existing standards have been proposed. One means of extending the capabilities of X.509 are specified in PKCS#6 developed by RSA Laboratories. PKCS#6 [RSA93a] allows one to extend the attributes included in the X.509 certificate, some of which are collectively signed by the certificate issuer. The specification lists four benefits of basing the standard on X.509 rather than replacing it:

1. changes to X.509 certificate syntax are easily followed.
2. an X.509 certificate can be extracted for compatibility with other standards.

---

65 See Generalized Certificates at: http://www.clark.net/pub/cme/html/cert.html
66 Though Trek states, "However, it should be emphasized that it is a certificate that embodies a CA and not a key... Therefore, if a CA has not put restrictions on a public key (its length, for example), a user may have the same key in multiple certificates, signed by different CAs." [Trc94:78]
3. both the X.509 certificate and the extended certificate can be verified with a single public-key operation, since they are signed together by the same certificate issuer.

4. there is little redundancy—having an extended certificate in place of an X.509 certificate would require that the extended certificate contain much of the information already in the X.509 certificates, such as issuer and subject names.

[RSA93a:4]

However, it is clear that the utility of such an extended certificate is limited, for while it may be useful to include some basic information (such as email address, telephone and fax numbers or postal information, as enumerated in [RSA93b]), it cannot represent a richer set of statements where multiple CAs assert attributes beyond trivial directory information. For instance financial, medical, legal, and other institutional information should be represented, and a single CA cannot validly sign all of this information.

An interesting argument for a general certificate is that of Röscheisen's "General Certificates" [Rös95]. These certificates express a "third party-opinion about an attribute of some entity" where an entity is represented by an immutable minimal principle handle (EpersID, which is equivalent to the rID discussed in the previous section). The syntax of the EpersID is similar to the DN/CA name subordination used in X.509, but rather than a DN, a CA assigns a UID (which can/should be a random number) of the form CA:UID. The certificate and assertions are independent (immutable) of real world changes. This semantically null string also provides the principal privacy during interactions.

Descriptive properties are listed using a dot notation. For instance, Verisign's assertion of subject 72355's phone number might look like the following:

[vs:72355].phone=415-7110 [Rös95:2]

Interestingly, Röscheisen's general certificate is very similar to the PKCS#6 extended certificate. One difference is that Röscheisen incorporates the fields of the X.509 directly into his new type GeneralCertificateInfo, while PKCS#6 includes the certificate X.509 field as a distinct subset. The PKCS#6 method allows for compatibility since the canonical X.509 fields can be easily extracted. The second difference is the fact that

subject Name is replaced with subjectID EpersID

The greatest benefit of general certificates is the ability of certificates to be recursively included as an attribute of another certificate. Also, the definition of the CompositeGeneralCertificate type allows one to easily aggregate a collection of certificates into a "meta" certificate. With a composite certificate a user has control over
which certificates are displayed to whom during transactions. As such, one need not expose employment information every time one shops at the corner store.\textsuperscript{67}

The relevant question given the distinct advantages of a generalized certificate is how will such a concept come to be implemented for the immediate future? This question is the topic of substantial debate in the PKIX,\textsuperscript{68} and SPKI\textsuperscript{69} groups and their respective mailing lists. PKIX is an IETF working group that is chartered to extend X.509 to support a public key infrastructure for the Internet. SPKI is not yet a working group, but its members hope to develop a standard that provides alternatives to the X.509 model of certificates. Many participants of these groups have a stake in developing systems, applications, and standards for public key technologies. However, while the discussions are interesting they are typical of such standards processes – difficult to mediate and sometimes prone to religious wars (ASN.1 is a favorite topic). For instance, the competitive credit card payment standards STT and SEPP (which now have been unified into SET) were partly products of an argument between Visa and Mastercard over X.509’s concept of identity.\textsuperscript{70} (SET uses X.509 but with only a unique serial number and no principal identity information included.)

Much of the debate regarding generalized certificates versus X.509 involves the X.509v3 standard which extends the capabilities of X.509 similar to PKCS#6. An amendment to X.509v3 also provides for extensions to the certificate that a) provides additional information about the key, the key ID, and restrictions on the key usage or certification policy, b) provides alternative names, and c) extends the constraints one can place on CRLs and CPs rule matching. [CCITT95] These debates have been focused on whether such modifications are bending an archaic standard to fit modern needs, or abandoning a valid and worthwhile standard for untested ideas.

Even though it is "unpublished" X.509v3 is the standard which its supporters tend to argue will be able to meet the needs of a generalized certificate\textsuperscript{71} Others argue that X.509 should be retired. As such, the result of these standards process is unclear (though I tend to think some form of X.509 will be come quite common). However, to include the flavor of such discussions, I include a chart of Laurence Lundblade (Appendix B) that demon-

\textsuperscript{67} It is possible that sophisticated access controls could be created for rich certificates to the point that a user could encrypt attribute fields, or the certificate issuer could even encrypt fields and hide the semantic content from the user.

\textsuperscript{68} http://www.ietf.cri.reston.va.us/html.charters/pkix-charter.html

\textsuperscript{69} http://www.hks.net/cpunks/codepunks-0/0109.html

\textsuperscript{70} This topic is further analyzed in ELECTRONIC COMMERCE PROTOCOLS AND COMPETITIVE STRATEGIES: Credit Card Transactions over the Internet [LR95] at http://rpcp.mit.edu/~reagle/2Misc/web/career/stuff/compete/Toc.html.

\textsuperscript{71} A X.509v5 is in the wings.
strates some of the issues and standards involved in deploying privacy enhancing infrastructures.

6.3.3 Structure – Hierarchical v. Non-Hierarchical

One of the pivotal issues in representing trust is the model by which the trust relationships are structured. We have already seen the hierarchical model of X.509/PEM. In contrast, PGP offers a “grass roots” or decentralized model and its trust structure is referred to as a web of trust. Such a system allows the users to be responsible for creating and certifying each others keys and eventually a “web” of trust relationships (employer, to friend, to student) forms by which one hopes to find a certification path. Rather than relying upon certification authorities, one relies upon trusted introducers. For instance, I might trust my father to the degree that if he signs a key saying “key K really belongs to person P” I will accept that as a fact. It is important to keep in mind that PGP allows one to sign a key in one of two ways as discussed in the documentation:

1) Does the key actually belong to whom it appears to belong?
   In other words, has it been certified with a trusted signature?

2) Does it belong to someone you can trust to certify other keys? [PGP90]

The first point is a relatively simple mapping between an identity and a public key. The second is more interesting and is an assessment of a person’s trustworthiness and competence in key management. This distinction reflects the points made in Chapter One, point five. Characteristics of Trust Evaluations with respect to “expertise”. However, unlike those aspects of trust, the PGP scheme is simple. One can rate the trustworthiness of the person as an introducer at unknown, untrusted, marginally trusted, or completely trusted levels. One can then define a small trust algorithm which generates a validity score for any key added to one’s key ring. For instance one may accept an introducer if his key has two marginally trust signatures, or one completely trusted signature.

As stated before, many feel that the capabilities of being able to form immediate relationships that are not dependent on complex CA scheme led to the widespread use of PGP in favor of PEM. However, in the future a hybrid scheme may be a likely alternative to either strictly hierarchical or distributed structures. For instance, a student of MIT could have his PGP key signed by an MIT key server (which might authenticate him as a valid MIT user by checking a registration database with a SSN number). The key server would then sign and store the key so that users at MIT can grab the public key when needed. Perhaps Harvard also has a similar key server, and the administrators of the MIT and Harvard systems know each other and are willing to sign the opposite server’s key with their own
server's signature. In such a scenario, a pseudo hierarchical system of key servers could evolve so as to assure that public keys in the .edu domain belong to real students. However, at this point in time this is not how PGP key servers actually work because what I proposed by having the administrators sign the other servers keys keeps the MIT key server in "charge" of the keys for MIT students. PGP servers are never in charge of a public key, rather they are public bulletin boards where users may post their keys, and the servers then form a network to propagate keys (i.e. exchange key rings) over time. Regardless, if a newer structure becomes popular, a hybrid system may allow one to combine a hierarchical certificate (for instance a payment certificate from VISA) and non-hierarchical certificates (for personal communication) in one generalized certificate. For an example of such a system, see Scaling the Web of Trust: Combining Kerberos and PGP to Provide Large Scale Authentication [SA95].

6.3.4 Revocation

Trust is not static, and not all assertions last indefinitely. This is a significant difficulty for certification schemes. For instance, Alice may be a student at Harvard and has a public key certificate signed by Harvard so that people may properly identify and authenticate her. When Alice graduates and gets accepted into MIT, she may no longer wish to be associated with Harvard and she or Harvard may wish to revoke the certificate that asserts her status as a student there. With PGP, it is fairly common for people to either lose their private keys, or wish to change keys to better the security (one's key may be compromised or one may wish to increase the key length). In both situations, it is necessary to revoke the old key so that no messages are mistakenly sent using the old keys (which will be unreadable if the old key is lost, or read by others if it is compromised). The necessity of revoking certificates (both identity and general certificates) adds an added level of complexity to key management. In the X.509/PEM model authenticating services post and check certification revocation lists (CRLs) for keys that still may be floating about, but which are no longer valid.

In keeping with PGP's distributed nature, keys are revoked by creating a key revocation certificate (KRC) that is then distributed and tells other users' PGP programs not to use that key again. Unfortunately – and fortunately, otherwise we would have spoofed KRCs – one needs the private key to create a KRC. If one can generate KRCs one needs PGP key servers to propagate KRCs so that eventually a key will have been revoked throughout a network of servers. It is not generally useful for a key server administrator to delete a key

---

72 Micali, using NIST PKI estimates assumptions, states that, "For instance, if each user verifies 100 digital signatures per day on average, then the total PKI yearly costs are $10,848 Millions, of which $10,234 Millions are due to CRL transmissions." [Mic96:3]
because other sites with that old key will propagate the old key to that site again. Admin-
isters may manually disable a key for their site if a user convinces them in the absence of a valid KRC that the old key really is lost. This disallows the old key from being fetched by users from that specific site, but the disabling only applies to that specific site and old copies of a key may persist elsewhere on the network.

An obvious solution to the above problem is to include expiration dates on certificates (or the attributes therein). However, the setting of a revocation date into even the near future rarely coincides with the event that necessities that the key/attribute be revoked. For instance, if a key is compromised 1.5 years from now and it expires in 2 years, one is confronted with the same problem, but one that only lasts 6 months. Research on revocation often involves increasing the granularity by which one can control a certificate’s lifetime. One method of accomplishing this in [LABW91:6]. In their example a set of key servers (caches) on a network may hold a certificate which asserts key #574897 is linked to user Burrows. If this assertion no longer holds, the typical method requires that every server be notified that this certificate is no longer valid. Their proposal is to define limited lifetimes for every certificate that is refreshed from the source whenever it reaches its expiration date but remains uncompromised. A requirement of their scheme is that it be “fail-secure”-- if a fault occurs a system will deny a valid access request rather than grant an invalid request.

A balance that once must keep with this scheme and others is the relation between the number of refreshes (increased granularity) and the network bandwidth or overhead costs. Also, “it’s hard to make a source of information that is both highly secure and highly available.” [LABW91:6] The solution to the latter problem is to use two sources (CAS) to refresh keys, one which is highly secure (K) and has a long lifetime, and another which is highly available and has a shorter lifetime (O). Using these two sources the trusted computing base (TCB, the amount of hardware or software one must trust to remain secure) is able to have the following useful characteristics.

- The TCB for granting access is dependent only on the security of the highly se-
cure CA K. (This is the result of a rather involved argument in [LABW91:12] and their joint authority theorem (P12) in which a chain of “assertions” and “quotations” are made between K and O.)
- The TCB for not granting access is dependent on K and O since either can pre-
vent a certificate from being refreshed.

Silvio Micali in Efficient Certification Revocation [Mic96] has also proposed a scheme in which CAS send “refresh” lists to certificate distributors (directories). A significant problem of current CRL schemes is that when a user queries a directory for the liveliness of a certificate, the directory must return the whole CRL of the appropriate CA. Two benefits of his Certificate Revocation System (CRS) are:
completeness – if a user queries for the status of a certificate the CRS provides a “complete and satisfactory” answer. With a CRL, if a user queries about a certificate, \( n \), which does not belong to any not-yet-expired list, the directory cannot prove that the certificate has not yet expired (it can only argue that it doesn’t have a certificate revocation for it.)

efficiency – the savings in communication costs (sending refreshed lists rather than whole CRL’s between the directory and user) is expected to be on the order of 900 times cheaper than CRLs.\(^{73}\) A computational advantage is that digital signatures are not required by the CA for recertification or revocation.

Within a CRL, rather than having a single preset expiration date, a certificate can be repeatedly refreshed or extended up to some finite limit. One could use a CRS to create a 30 day certificate with each increment applying to one day, The system works as follows:\(^{74}\)

For a certificate that is valid for one month, two quantities \( N_0 \) (for “NO”) and for \( Y \) (for “YES”) are included to represent the status of the revocation. The CA used two values secret values \( N_{32} \) and \( Y' \) to generate \( N_0 \) and \( Y \) accordingly:

\[
\begin{align*}
H(Y') &= Y \\
H(N_{32}) &= N_{31} \\
H(N_{31}) &= N_{30} \\
&\vdots \\
H(N_1) &= N_0
\end{align*}
\]

where \( H \) is a one-way-function

When a directory is queried with respect to the revocation of a certificate, the status is dependent on the certification policy:

1. If it produces a \( Y' \) such that \( H(Y') = Y \), then (“YES”) this certificate is revoked.
2. This certificate should only be considered as valid on day \( i \) of the month if it is accompanied by an \( N_i \) such that \( H^{<i}(N_i) = N_0 \) (or “NO” the certificate is not revoked.)

The CA can "recertify" the certificate for one more day by handing out \( N_i \) to the distributors of the certificate (the owner and directories). The CA can revoke the cert by passing out \( Y' \).

### 6.4 TRUST MANAGEMENT

In Chapter Two I described a wide ranging concept of trust in which one may hold a whole series of views about the world and the trustworthiness of different agents with

---

\(^{73}\) Comparisons are made with the NIST PKI specifications. [Bur95, Naz95]

\(^{74}\) This specific explanation is partly based on an email from Ron Rivest to the cypherpunks list.
respect to different things. (See Table 2-1) However, the certification schemes discussed
until now only allow for a narrow range of assertions – namely identity. However, recent
research into trust management\textsuperscript{75} may allow for the development of systems which sur-
pass current schemes and move towards those envisioned in the discussion of generalized
certificates. Blaze, Feigenbaum, and Lacy [BFL96] have addressed the topic of managing
the potentially complex generalized certificate model in Decentralized Trust Manage-
ment. This scheme allows for generalized certificates and decouples the expression and
the checking of a policy from the enforcement mechanism.

The AT&T system is known as the PolicyMaker and allows for rich trust expressions.\textsuperscript{76}
For instance, it supports threshold schemes where a policy could require 3 out of 5 bank
officer signatures for loans in excess of $1,000,000. The syntax is relatively straightforward. Assertions are of the form:

\textbf{Source ASSERTS AuthorityStruct WHERE Filter}

Where Source is the agent making the assertions; source may be a local policy or another
certificate. (This is similar to my own development in Chapter Two with respect to as-
sertions and assumptions. In this system, a local policy is an assumption an need not be
signed by any other authority for it is strongly believed regardless.) The AuthorityStruct
is the object of the assertion, and Filter is the predicate or way in which a specific policy
is expressed.

A user may query the PolicyMaker system with the following:

\textbf{key}_1, \textbf{key}_2,...,\textbf{key}_n \textbf{REQUESTS} ActionString

Where the \textbf{key}_n are the identities applying for the action, and the ActionString is the re-
quested action. In the bank example, two bank officers may request a loan of $1,000,000
more. The ActionString is matched against a filter (Figure 6-6) which is a predicate for
the policy and can be represented in a language such as SafeAWK or SafeTCL. For in-
stance, any number of immediate conditions on a policy (for instance, \( x > 1 \times 10^6 \)) can be
expressed as a filter. Less immediate restrictions regarding the number of trust deferrals
can also be called by filters and by annotations (a type of filter).

The PolicyMaker may have a certificate from the Vice President of Pperations that states
a requirement for three bank officers to approve loans greater than $1,000,000. Also,
perhaps the Vice President’s key was signed by the CEO’s signature, which is accepted as
valid as an assumption (local policy). Hence, not only is the ActionString required to be

\textsuperscript{75} A term that I am borrowing from [BFL96].

\textsuperscript{76} However trust in this system is monotonic (similar to [Ran88]) and does not allow negative permissions.
valid, but the deferral of the trust from the CEO to the VP must also be acceptable (and can be specified in the ActionString). Such a query could be represented by:

**Figure 6-6**
PolicyMaker REQUEST to ASSERT matching

<table>
<thead>
<tr>
<th>CLIENT:</th>
<th>REQUESTS</th>
<th>ActionString</th>
</tr>
</thead>
<tbody>
<tr>
<td>key, key₂, ..., keyₙ</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

At each stage of the processing, the series of filters, and deferrals are checked. This can also be aided by annotations which add additional restrictions or information not present in the original request. Once the original query is passed through a chain of filters and picks up annotations, it is passed through the chain once again to make sure a filter at the beginning of a chain can approve of all of the annotations added through the first processing.

The authors argue that such a system allows for a unified mechanism, flexibility, locality of control and separation of mechanism from policy. As such, it is important to remember that the PolicyMaker does not understand the meaning of ActionStrings, they are specific to an application. The PolicyMaker’s only concern is that it is able to parse such strings against a filter. Also, all signature verification are done outside of the PolicyMaker, and can be done before the query is passed to the PolicyMaker. (If a signature fails, then one need not waste time processing it further.)

The success of a system like PolicyMaker is dependent on how well it meets the goal of providing an expressive language while remaining easy enough so one can easily specify common trust relationships. Additionally, there are some technical questions such as:

- how to find the best path through a path of policies
- how to deal with contrary permissions and policy paths (for instance, one path may allow an access, another may not)
- and how to deal with malicious policies that are submitted with the intent to confuse or slow PolicyMaker parsing
With respect to the broader issue of richness and ease of use it is instructive to consider real life experience. People have many complex intuitions regarding degrees of trust. Yet, in the day-to-day experiences of living most people proceed by relying upon an enormous number of assumptions and default rules. For instance, I generally assume the traffic lights in Cambridge have a three second yellow light. Rather than evaluating a complex function, I may rely on that default rule. Of course, when such a rule fails me, I may get a ticket. While default rules may provide a solution to complexity, they must be made carefully with regards to secure applications. Hence, I expect, if an expressive – and complex/complicated – method of representing trust is developed, simpler default rules which have been repeatedly tested and accredited by some agency shall be used for the majority of all trust transactions on the Internet.
6.5 APPENDIX A: ASN.1 SYNTAX FOR CERTIFICATES AND CRLS

The X.509 certificate format is defined by the following ASN.1 syntax:

Certificate ::= SIGNED SEQUENCE{
    version [0] Version DEFAULT v1988, 
    serialNumber CertificateSerialNumber, 
    signature AlgorithmIdentifier, 
    issuer Name, 
    validity Validity, 
    subject Name, 
    subjectPublicKeyInfo SubjectPublicKeyInfo
}

Version ::= INTEGER {v1988(0)}

CertificateSerialNumber ::= INTEGER

Validity ::= SEQUENCE{
    notBefore UTCTime,
    notAfter UTCTime
}

SubjectPublicKeyInfo ::= SEQUENCE{
    algorithm AlgorithmIdentifier, 
    subjectPublicKey BIT STRING
}

AlgorithmIdentifier ::= SEQUENCE{
    algorithm OBJECT IDENTIFIER, 
    parameters ANY DEFINED BY algorithm OPTIONAL}
6.6 APPENDIX B: AN EXAMPLE ANALYSIS OF THE TRUST MODEL

Reprinted with permission of the author Laurence Lundblade.

Any fully implemented system will have to choose some form of a trust model. Some possibilities are:

* web of trust
* strict hierarchy
* web of hierarchies or some other hybrid

The important thing here is that there are many trust models that are valid and useful and it may be useful for other components of the system to be neutral to the trust model as is clearly the case with MOSS.

The Key Distribution System

A lot of components may go into this (protocols, client/server architectures, local key stores) and it is probably the most complicated part of any system. Some options are:

* distribution of keys manually via e-mail
* automatic non-interactive lookup of keys from a server
* interactive browsing of a key store for keys
* revocation lists or none
* online certificate verification via a secure channel
* certificate caching

Probably the best thing to say, is that there's a lot of work to do here.

The Certificate format

It seems possible to pick a certificate format independent of the other issues. Doing so would allow us to leverage components like as one does with other data objects like MIME. There probably only two major contenders:

* X.509 v3
  + broadly supported by standard bodies
  + supported by several industries (e.g., banking)
  + very rich and flexible
  + ASN.1
  - ASN.1 (tough for a student to get an ASN.1 compiler)
  - complicated
* PGP keys
  + widely deployed
  + simple to write code for
  - difficult to lookup (linear search on key id required)
  - too simple to support many trust models and distribution systems

Note that both use the RSA algorithms, so they are interchangeable at some very basic level.

The Transport of Content format

This is the format of the actual message that is sent from one user to the next. I'm going to discard anything that doesn't handle MIME because I don't think they are important any more. Raph described a lot of this so I'll just mention a few considerations explicitly about transport formats.

* PGP/MIME

From a data structure format this is a compact binary format. It seems reasonable to implement, is documented and requires no special tools. There is a performance problem with key look ups for signed message because a linear search is required unless the key or other data is always included with the message.

* S/MIME (PKCS + MIME)

Uses PKCS format with some MIME formatting. The main problems here are the multipart/alternative format for signatures and the ASN.1 requirement. An ASN.1 compiler is required to implement this. PKCS has actually been around for a while and has been used for a number of cryptographic systems.

* MOSS

MOSS is perhaps the easiest to implement and the most flexible since it is an ASCII text protocol like other Internet protocols and because it explicitly supports several trust models.

I think the most important observation is that PGP/MIME and MOSS share the security multipart structure from RFC 1847. It is also possible to use the security multipart format with PKCS #7 and thus S/MIME could be changed to support it. If this happened we'd have something in common for all formats and it would make life much easier for all e-mail client authors. An added bonus is that RFC-1847 support allows an e-mail client to support encryption and signing of full MIME entities with an external program that can be configured like MIME content viewers are with something like the mailcap facility. It can be something as simple as a UNIX pipe to a command like pgp.
Government is a trust, and the officers of the
government are trustees; and both the trust and the
trustees are created for the benefit of the people.
— Henry Clay. *Speech, 1813*

Can complex trust instruments be implemented in a digital market? Previous chapters
addressed the technical and economic aspects of this question. This chapter focuses on
the broader – but no less important – policy issues. Elsewhere, this thesis discussed
“policy” repeatedly with respect to trust and certification structures. Certification poli-
cies explain the conditions under which one vouches for another. Policy in the broader
sense of the word – the conditions and guidelines under which an institution legislates,
regulates, or acts – is also important: policy is often tightly coupled with, or biased by,
the technology it applies to.

The digital world presents a number of challenges to typical policy processes. First,
technology often changes faster than policy. Second, networking technologies are capable
of affecting the policy process itself – people can use the network to communicate and
organize. Third, networking technologies, while similar in many ways to other communi-
cation technologies, could exceed the effects of any previous technology in the depth and
breadth of their impact on society. The ability to develop sophisticated markets that em-
ploy a variety of trust and financial instruments – as well as provide communication,
entertainment, and civil functionality – is dependent on the underlying technology, which
can be shaped by government policy.

Currently, a debate is taking place through-out the world about what roles governments
should have with respect to this technology. The debate is partly the result of the phe-
nomena I define as *precedent dependency*.\(^{77}\) Regulatory structures become dependent on
technological and political precedents (accidents) rather than general principles. For in-
stance, in the US a number of general principles (rights) were non-exclusively
enumerated in the amended Constitution. An oft cited right is that of free speech. How-
ever, this rather simple principle has evolved into a complex policy structure wherein the

\(^{77}\) The concept of precedent dependency is related to “path-dependence” as discussed by Kavas-
salis [Kav95].
right of free speech is different with respect to the communications media it employs: person-to-person, broadcast, common carrier, or print. Digital network technologies, which make those distinctions moot, confuse policy makers.\textsuperscript{78} The unfortunate result of this dependency is regulations that are no longer relevant to the current environment or technology.

An example of precedent dependency is the controversial US wire tap legislation.\textsuperscript{79} Wire taps, which obligate a communications carrier to assist in the monitoring of a communication, are generally approved as a limited exception to the right of privacy. Law enforcement agencies have since become dependent on this mechanism, and – to their alarm – this capability may be threatened by new digital technologies. Consequently, law enforcement agencies promoted new legislation that required telecommunications carriers to build automatic wire tapping\textsuperscript{80} capabilities into their networks. Hence, a judicial exception to the right of privacy and the technological “accident” of being able to put a clip on a wire has become an unusual technological requirement of our communications infrastructure.

I am not so cynical (or foolish considering the history of the Internet) to think that all government involvement is folly. Yet, the issues at hand are truly difficult and policy makers should perhaps think before they leap. Froomkin described this situation fairly:

As U.S. lawyers we are most accustomed to thinking about the problems of data creation, dissemination, and access in certain delimited categories such as the First Amendment, intellectual property rules, the torts of invasion of privacy and defamation, and perhaps in the ambit of a few narrowly defined statutes such as the Privacy Act or the Fair Credit Reporting Act. The categories are valuable, but are collective inadequate to the regulatory and social challenges posed by the information production, collection, and processing booms now under way.

[Fro96:3\textsuperscript{81}]

This chapter reviews examples of trust related services (or their underlying technologies) that are traditionally supplied or regulated by governments, but that may be significantly changed by new technologies.

\textsuperscript{78} See Pool's \textit{Technologies of Freedom} [Pol83].
\textsuperscript{79} See http://www.cpsr.org/doi/wiretap.html
\textsuperscript{80} The wire taps would be very convenient, the law enforcement agency could request that all pertinent calls be routed to their own facilities for monitoring.
\textsuperscript{81} Unfortunately, my pagination will be different from that of the forthcoming publication in a Symposium volume of the University of Pittsburgh Journal of Law and Commerce.
7.1 POLICY AND RULE MAKING

Before proceeding into the specific technologies and issues that are so contentious, I would first like to extend the discussion regarding the nature of policy and its relationship to technology.

In the real world, one's expectations about that world is partly a social construct. As I discussed earlier, people make many assumptions in order to conduct business, or to just get through the day. One could argue that social institutions are motivated by the dissatisfaction of members of a chaotic (untrusting) society and regulatory institutions develop to limit this dissatisfaction and increase stability. A world without traffic lights would be intolerable, but laws exist so that when one comes to a green light, one expects one can safely proceed through the intersection. Society has encoded a number of useful default expectations about the world and attempts to enforce them.

Hence, governments mint money, regulate markets, and legislate. The digital realm is quickly becoming an area that is ripe for the emergence of trust intermediaries, brokers, and third party services. Consequently, governments around the globe are becoming increasingly interested in extending their regulatory powers into this new realm. However, are governments the proper institution to fill this void? This question is a complex one because underlying it are a number of equally difficult questions:

- From a political theory point of view, what abilities are granted and restrictions placed upon governments? (A very old question and one confused by precedent dependency.)
- Do those abilities and restrictions upon governments change when they enter into this new realm? Examples of free speech and wire tapping have already been mentioned.
- Are aspects of the digital realm so unlike the real world that many government services are no longer needed? (For example, need they be the sole “creator” of currency in the future?)
- How does one deal with the international aspects of regulating networks?

Post examines similar questions in *Anarchy, State, and the Internet: An Essay on Law-Making in Cyberspace* [Pos95a] in which he addresses who governs cyberspace by relying upon Robert Ellickson’s framework of behavioral controls, and the role various entities play in regulating an environment. He also argues – as I did early – for the capability of policy to dramatically affect technology:

Networks – electronic or otherwise – are particular kinds of “organizations” that are not merely capable of promulgating substantive rules of conduct; their very essence

---

82 The significant civil issue is to whom are the default rules useful for? Democratic societies should strive to make the legislation fair and useful to all.
is define by such rules – in this case, the "network protocols". Accordingly, the person or entity in a position to dictate the content of these network protocols is, in the first instance at least, a primary "rule maker" in regard to behavior on the network.

[Pos95a, par. 12]

The Internet is not immune from government coercion. This chapter includes many examples of governments attempting to control protocols to effect policy.

7.2 CRYPTOGRAPHY

Some degree of trust will exist between the participants of any electronic communications. For example, I have purchased computer RAM over the Internet without requiring anything beyond simple email. However, encryption technologies – as should be evident from this thesis – are essential to the development of sophisticated and efficient trust and financial instruments. I, and the seller of the RAM, were each at significant risk which would be lessened by encryption technologies. Unfortunately, governments across the world have hampered the development of widely useable encryption in a number of ways. Countries like France and Russia have made the general use of cryptography illegal!\(^{83}\)

The US government has attempted to hamper cryptographic deployment in a number of ways:\(^{84}\)

- by attempting to restrict cryptographic research (it has in the past "requested" that the dissemination and publication of research results be postponed).\(^ {85}\)
- by regulating the export of encryption technologies as a munitions under the US's International Traffic in Arms Regulations (ITAR). Hence, only very weak technologies can be exported or sold on the world market. Many have ar-

---

\(^{83}\) See Bert-Jaap Koops’ extensive Crypto-Law Survey at http://cwis.kub.nl/~frw/people/koops/lawsurvy.htm

\(^{84}\) Immense amounts of information on all of these points and more can be found in the EFF "Privacy, Security, Crypto, Surveillance" archive at http://www.eff.org/pub/Privacy/ .

\(^{85}\) "In 1984, the President had issued a National Security Decision Directive, NSDD-145, which gave the intelligence agencies broad authority to peruse computer databases, for so-called 'sensitive but unclassified information.' A subsequent memorandum from John Poindexter, expanded this authority still further to include ‘all computer and communications security for the Federal Government and private industry.’ As the government's authority to control access to computerized information for the purpose of protecting national security expanded, the free of flow of information diminished. Stories of agents visiting private information vendors and public libraries soon followed. At the same time, a wide range of other activities by the government further threatened to restrict access to information." Marc Rotenberg’s prepared testimony on The Computer Security Act Of 1987 (P.L. 100-235) and The Memorandum Of Understanding Between The National Institute Of Standards Technology (NIST) And The National Security Agency (NSA) Before The Subcommittee on Legislation and National Security, Committee on Government Operations U.s. House of Representatives. MAy 4, 1989. See http://www.epic.org/crypto/csa/rotenberg_testimony.txt
gued this hampers US competitiveness in the world market for communications and transaction technologies.²⁶

- by offering a number of substitute technologies that are weaker or limited such as EES (a standard for escrow keys for government access) or CLIPPER (a chip for voice communications that is part of the EES). Another example of offering an alternative technology is the DSS scheme which unlike RSA can only be use for authentication and not confidentiality.
- by proceeding in a often underhanded manner:
  - Documents have shown that advanced telephony has not yet been a problem for law enforcement agencies.²⁷
  - In cases where US companies attempt to move their cryptographic development for international products off shore, they receive warnings that they will be subject to other regulatory pressures if the company continues with its off-shore strategy.²⁸
  - The government prosecuted Philip Zimmermann for the distribution of PGP when the proliferation of such software, or mathematical knowledge in general is impossible to stop.²⁹
  - The CLIPPER initiative, which agencies such as the FBI professed to voluntary system, was actually intended to be mandatory.³⁰

Also, attempts to create weak (or restricted) encryption policies at the international level are being promoted by many nations.

This topic is a fascinating one, and information regarding technologies, services, faults, risks, and policy are in the news on a daily basis. The story of Daniel Bernstein’s publication of cryptographic source code resulting from academic research was recently in the news.³¹ Judge Marilyn Patel in the Northern District of California denied the request to

²⁶ See EFF "Privacy - Crypto - ITAR Export Restrictions" Archive at: http://www.eff.org/pub/Privacy/Key_escrow/ITAR_export/.
²⁷ However, despite efforts to obtain documentation from the field in support of Bureau claims of a "crisis facing law enforcement," the response from FBI Field Offices was that they experienced no difficulty in conducting electronic surveillance. See http://www.eff.org/pub/Privacy/Clipper/wiretap_fbi.memo.
²⁸ From a reputable but anonymous source. Also, see footnote 85.
²⁹ See EFF "Legal Cases - PGP & Phil Zimmermann" Archive at: http://www.eff.org/pub/Privacy/ITAR_export/PGP_Zimmerman_Viacrypt_AustinCW_case/.
³⁰ Newly-released government documents show that key federal agencies concluded more than two years ago that the "Clipper Chip" encryption initiative will only succeed if alternative security techniques are outlawed.... The conclusions contained in the documents appear to conflict with frequent Administration claims that use of Clipper technology will remain "voluntary." at http://www.eff.org/pub/Privacy/Clipper/Clipper FOIA/epic_clipper_secrets.announce.
dismiss the case which has implications on cryptographic export controls and communications:

Judge Patel's acknowledgment that source code enjoys Constitutional protection has implications that reach far beyond cases involving the export of cryptography. The decision holds importance to the future of secure electronic commerce and lays the groundwork needed to expand First Amendment protection to electronic communication.

[EFF Press Release at http://www.eff.org/]

It is unclear how this issue will be resolved, but it is clear that what is at stake is of both a global and personal significance. This technology is important to the development of a global information economy and to the rights of the individual participants.

## 7.3 COMMERCE

Encryption technologies can enable a number of instruments or tools that are strongly related to trust. However, even if things like electronic cash, security deposits, digital signatures, contracts, intermediary agents and notaries are technically possible, the legal standing of these instruments will have an immense impact on their acceptance. Many instruments such as digital signatures will be used by real world companies for real world commerce. Hence, a legal understanding will have to be reached before these instruments are used for even a small portion of the many transactions that occur across the globe. Of course, the legal and regulatory system is often far behind the cutting edge of technology, but it is sometimes in step or even ahead of daily business practice – at least in terms of the topics being addressed and not necessarily the quality of the decisions.

### 7.3.1.1 - Digital Signatures and Contracts

Digital signatures\(^{92}\) are perhaps the most widely legislated upon topics in this section and they enable the encoding and authentication of contracts, purchase orders, and the like in digital form. Currently, ten states (including Utah\(^{93}\), Washington, and Wyoming) are considering digital signature legislation. California passed the California Digital Signature on October 4, 1995. The Legal Counsel's Notes [Cal95] for the legislation are relatively straightforward:

This bill would provide that, in any written communication with a public entity, a signature may be affixed using a digital signature and that in those communications,

\(^{92}\) Legislation and reports on digital signature may be found at http://web.aimnet.com/~software/industry_issues/digsig.htm

\(^{93}\) See the excellent Utah Digital Signature Legislation Base at http://www.state.ut.us/ccji/digsig/default.htm
the use of a digital signature would have the same force and effect as the use of a manual signature if it complies with the bill's requirements...

And the definition of the signature is informative.

(1) It is unique to the person using it.

(2) It is capable of verification.

(3) It is under the sole control of the person using it.

(4) It is linked to data in such a manner that if the data are changed, the digital signature is invalidated.

(5) It conforms to regulations adopted by the Secretary of State. . .

(b) The use or acceptance of a digital signature shall be at the option of the parties. Nothing in this section shall require a public entity to use or permit the use of a digital signature.

(c) Digital signatures employed pursuant to Section 71066 of the Public Resources Code are exempted from this section.

(d) "Digital signature" means an electronic identifier, created by computer, intended by the party using it to have the same force and effect as the use of a manual signature.

To guide policy makers the American Bar Association, Section of Science and Technology has written Digital Signature Guidelines to help inform legal processes with regards to this topic.94 The acceptance of digital signatures shall be one of the first indicators of the ability of legal and corporate cultures to adapt to the digital world.

7.3.1.2 - Intellectual Property (IP) and Digital Networks

Ways to limit the potential abuse of intellectual property is a very contentious issue. First, there is the question of the proper balance between the rights of the reader/user and owner of the IP. Second – and this should not be surprising – those that favor the large producers of information wish to vigorously protect their IP by specifying technical requirements for the networking infrastructure. One possible scenario is that users loose much of their fair use and physical copy (personal property) rights and everything from their client to the network infrastructure is design to enforce this restriction. The other extreme is an environment in which copyright owners have absolutely no control over the distribution (nor reimbursement for the creation) of content. A number of proposals have been offered, the first of which was the Clinton Administration's Intellectual Property Working Group "Green Paper" (July 1994) which:

proposed to modify the concept of copies in a digital network. It recommended three things: to expand the Copyright Act so as to include the transmission of a copy-

---

94 The guideline is no longer freely available to the public. However, information on the guidelines can be found at: http://www.intermarket.com/ec1/.
righted work as infringement, to exempt digital works from the First Sale Doctrine, and to disallow the creation of devices with a primary purposes of defeating a technology meant to protect IP.

[Rea95]

While this proposal was derided by the broader Internet community for favoring the interests of large media companies, the proposal has allegedly regained momentum and has influenced possible legislation on the topic. Another alternative is the complete rewrite of existing intellectual property laws, the most significant of which proposes that software be protected with regard to its behavior rather than its expression (copyright) or novelty (patents). [SDKR94] The other proposal has been to rely on existing legislation and case law that—if read properly—is general enough to apply to digital networks. [Rea95]

7.3.1.3 - Electronic Cash, Banking, Tax Evasion, Money Laundering, and Fraud

The US Federal government began minting money because of the incompatibility of independent currencies from different banks or states, making interstate commerce extremely difficult.95 With electronic cash systems, this incompatibility may no longer be a problem. If the incompatibility of digital instruments were to be a problem, it may not be one best solved by a central entity issuing fiat currency. Standards bodies could attempt to mediate interoperability or third party intermediaries would no doubt extend services for the exchange of electronic currencies.

Speculation about the impact of electronic currencies outside of the control of governments is imaginative and wide ranging.96 Australian and Swiss banks have been pressured from the international community (particularly the US) to remove anonymous accounts and services.97 Bonded offshore services that employ PGP are easily reachable

---

95 "In colonial times, it was very common to find varieties of currency in use from such countries as Great Britain, Portugal, Spain, France, and Germany. In addition to local currency issued by the colonies, wampum, livestock and local crops, such as tobacco were also used to conduct business and other daily transactions of commerce. This variety caused confusion and slowed the process of trade and economic growth. The critical need for a standard monetary system became an issue of great concern to the framers of the U.S. Constitution. "By 1776, Thomas Jefferson had become the emerging Nation's strongest advocate for a unified system of coinage to lay the foundation for expanded trade and economic growth in the 13 states. To this end, Jefferson proposed the decimal coinage system that we use today. Moreover, over the next 15 years Jefferson would become a leading advocate for founding a national mint on American soil, both for practical reasons and as a symbolic reminder of the Nation's recently-won sovereignty."


96 [Fro95, Pos95a, Pos95b, Pan95] all address these topics.

on the Internet. Companies are using the Internet to exchange IPO information, and wish to use the WEB for actual trading. [Dav95] A major worry beyond terrorists using ecash, is that a significant portion of the tax base would begin using electronic finance instruments and stop paying taxes:

The Clinton administration's reluctance to ease up on export controls for encryption software stems in part from pressure from U.S. law enforcement agencies, and the owner of a New York-based software company sees heavy lobbying behind the government's desire to regulate content on the Internet: "I think the Internal Revenue Service and the FBI are watching this one very carefully. They wouldn't mind seeing the government set a precedent for deciding what can and cannot go on the Internet."

The IRS fears that easy access to cheap and sophisticated encryption software will make income- and sales-tax evasion too easy, and the FBI worries about criminal and terrorist plots hatched in cyberspace, but some observers say government control tactics are too little, too late. A Hudson Institute economist says, "Electronic money gets really interesting when you realize how impossible it is to put national walls around it, mandate the use of national currencies, or require that transactions go through banks... The country will have no practical choice but to rely more than ever on voluntary tax compliance. That means tax rates will have to be kept as low as possible on people and on businesses."

[Inv96]

However, in Dollars & Borders: U.S. Government Attempts to Restrict Capital Flows, 1960–1980, Hawley demonstrates that even before digital cash was a threat to governments:

State intervention to form, structure, and regulate markets, unintentionally but inevitable produces financial and market innovations circumventing state barriers, as individual units of capital pursue profits and unimpeded growth. Each instance of U.S. capital controls, along with other national restrictions on the free movement of international capital, only intensified the pace of the internationalization of finance, thereby further limiting the ability of the U.S. and other states to conduct monetary system and fiscal policy as they had previously.

[Haw87:170]

New digital financial and trust instruments may not prove to be a crushing blow to governments around the world, but perhaps they are indicators of the importance of the underlying network infrastructure. Digital instruments are icons of the ever increasing velocity, liquidity and lack of control of even "normal" currencies, instruments, and ideas that use digital networks services.

---

98 Quoting from an email addressed to the author, "We can execute your financial transactions, move cash from bank to bank, from brokerage accounts to bank, attorney, escrow account, etc. You instruct us by PGP as to what you need accomplished."
7.4 COMMUNICATIONS

The characteristics of real world communications today are generally well established and a product of a complex history. For instance, communications has developed along three technological and regulatory tracks: print, broadcast, and telephony.\footnote{See Pool's \textit{Technologies of Freedom} [Pol83].} People tend to be familiar with the degree of trustworthiness inherent to each media – network television news are considered to be more trustworthy than people handing out pamphlets on the street corner. People have also become acclimated to the level of freedom exercised in each media. Digital networks change this. All content from a traditional media will have a digital form. Pamphlets, news papers, editorials, music, talk shows, television programming, movies, and general voice and video communications will all take place over the same medium. Many of the contentious issues which were resolved because of political expediency and technological accident in the past are now open questions again. Some of those issues such as anonymity, freedom of speech and censorship are discussed below.

7.4.1.1 Anonymity & Privacy

While encryption technologies are useful because they allow one to create digital signatures, this does not preclude one from using the technology to conduct anonymous conversations or transactions. In \textit{Pooling Intellectual Capital: Anonymity, Pseudonymity, and Contingent Identity in Cyberspace}, [Pos95b] Post defined many of the relevant terms:

an "anonymous" communication is one in which the message itself contains (and hence the recipient of the message receives) no information regarding the identity of the sender.

a “pseudonymous” message is one that contains information (of varying reliability, to be sure) about the identity of the originator – the cognizable entity responsible for transmitting the message – without simultaneously providing information about the actual, biological, individuals responsible for transmission of the message.

“Traceability” measures the cost of obtaining information about the identity of the sender in addition to the information that is “readily apparent” – i.e., obtainable at (virtually) no cost – from the message itself.

From previous discussion, one can see that governments wish to keep the cost of traceability low by taking advantage of the previous generation of technology's characteristics, preventing more anonymous technologies from becoming widespread, and by mandating traceability requirements for new technologies. Even so, one can guarantee a fairly high level of anonymity on the Internet today by using anonymous
remainlers—though it is not a trivial exercise.\textsuperscript{100} Such remainlers provide a level of indirection that prevents the recipient from seeing whom the original message came from. By chaining remainlers and encrypting each segment of the chain in that remainler’s public key, no individual remainler know both the recipient and the sender. The functionality of remainlers need not be limited to email, they can be used to redirect news, and web queries. In fact, the use of web proxies (which provide a single point of access to the Internet to a group of users—for instance, all the AOL users) hides the identity of the individual user.\textsuperscript{101}

The issue of anonymity is significant beyond the issues of pornographers, drug dealers, terrorists, fascists, and spies. Commercial privacy will be a legitimate and pressing concern for the customer in the electronic market. Froomkin discusses Posner’s belief that anonymity with respect to the concealment of disreputable facts creates inefficiencies in a market—shifting the cost or risk to the other parties. In other words, anonymity increases the inefficiency of a market since it creates information asymmetries—valid transactions might not occur for the lack of trust, or extra effort must be expended in building or determining reputations. However, Froomkin correctly adds that, “Posner’s formulation, however, has been criticized for neglecting the strategic aspects of the individual’s desire to control the release of personal information that is not disreputable.” [Fro96:6]

\textbf{7.4.1.2 Censorship and Free Speech}

The issues of censorship and free speech are only indirectly related to issues of trust but are fascinating topics to follow—such as the ACLU v. Reno case.\textsuperscript{102} The relation with trust (and that of many of the topics of this chapter) is that the issue of censorship and free speech impacts the policy, which affects the technology, that enable sophisticated trust relationships. Unfortunately, there is no guarantee that the Internet will act as a democratic force, and that the suppression of political rights, and the rigid control of encryption technologies will prevent the formation of a sophisticated—but less interesting and less innovative—market with “big brother inside.”\textsuperscript{103} Readers interested in the topics of censorship, free speech, and privacy should see Froomkin’s \textit{Flood Control on the Information Ocean: Living With Anonymity, Digital Cash, and Distributed Databases}. [Fro95a]

\textsuperscript{100} One must also remember that the Internet was never built with the purpose of authenticating messages. Users can can often directly telnet to port 25 (the port that the mailer daemon sits on in unix) of a host and forge an email.

\textsuperscript{101} \url{http://www-sloan.mit.edu/15.967/group07/} for more information on web privacy and anonymity.

\textsuperscript{102} See the ACLU home page at \url{http://www.aclu.org/}.

\textsuperscript{103} Quoting a .sig of Timothy May, “Boycott big brother inside.”
CONCLUSION

Friendship is constant in all other things
Save in the office and affairs of love:
Therefore all hearts in love use their own tongues;
Let every eye negotiate for itself
And trust no agent.

The evaluation of trust is a necessary component of making any decision, even if one decides to distrust everything and everyone. The question this thesis posed is how will trust be used on the Internet? In this sense, trust is simply information. One rarely asks if information will be used, but how. Even with problems such as limited information, information of poor quality, or information that is used improperly, people do use it. To mitigate the effects of these problems, people employ various tools to evaluate, manipulate, and communicate the information. This thesis demonstrated that there are tools for the evaluation, management and communication of trust in digital networks. A protocol for a digital security deposit was developed to demonstrate the construction and the economic operation of such a tool.

The concept of trust is an ambiguous one because it is often taken for granted in day-to-day life. To redress this lack of understanding, this thesis examined trust with respect to three disciplines: philosophy, logical systems, and decision analysis. This thesis found that a useful way to think about trust is the degree to which one expects an assertion to match reality. It also demonstrated that trust as an information product should engender markets for its evaluation and exchange. Agents will develop capabilities to judge the trustworthiness of other agents or infrastructure. To operate competitively, agents may also develop the ability to create reputations that are calculated to be utility maximizing. The term “cryptographic economy” is introduced to represent a market where agents use cryptographically enabled financial and trust instruments to exchange value and build trust relationships. Current research on game theory may provide a clue as to how these markets may operate in the future.

However, to answer the question of how will trust be used on information networks, this thesis considered two areas where the evaluation and management of trust and risk are pivotal to making a profit: finance and economics. This thesis shows that just as instruments such as contracts, futures, bonds, stocks, and letters of credit allow one to limit
risk, make assertions about the future, or increase the likelihood of conducting transactions with untrusted participants, one can do the same in the digital domain. The relationship between financial instruments and trust instruments is demonstrated, and trust instruments are shown to lessen the need for trust and lessen the likelihood of risk; a consequence of which is that trust instruments increase the likelihood of the exchange of value and decreases the associated costs of the transaction.

The main focus of this thesis was that trust will be managed in three ways with respect to these instruments. First, instruments from the real world may be extended to digital network users, such as secure credit card presentation protocols for the Internet. These protocols may be used with greater efficiency for "real world" transactions by using a digital network infrastructure. For instance, one can use an Internet bank to lower costs to the bank and increase customer satisfaction. Second, digital versions of those tools are being created that liberate them from many of the constraints of the real world. With e-cash one may be able to conduct anonymous, international commerce solely within the domain of the digital network. Cryptographic protocols allow for the creations of tools which lessen the need for trust in ways that strongly resemble, differ slightly, or are wholly different from their counterparts in the real world.

The third way to manage trust is to shift the direction of the trust relationships. Instead of a landlord being at risk and having to trust a tenant not to damage the apartment, a security deposit displaces the risk onto the tenant. Such an action can be considered to be the simple shifting of trust, or the coercion of a users’ behavior depending on how sensitive the tenant is to risk. (If a tenant is just as likely to trash the place as before, the trust relationship has been shifted, if the tenant will now be more careful, the trust relationship has been shifted and the risk lessened.) To demonstrate how this can be accomplished a protocol is developed that is similar to a security deposit. The digital security deposit prevents users from redistributing digital objects (known as characteristics) by creating an economic incentive not to do so. The protocol is essentially a digital e-cash note, "blinded" in a zero knowledge proof. Two limitations of this protocol are its rather expensive computational requirements, and a lack of an understanding on how to properly price a digital security deposit in an electronic market. Also, the protocol does not provide a reimbursement to the person who is betrayed, but it does create an incentive for the principals of the relationship to behave fairly. The result of the protocol is that one can sell an access token to a user on a digital network and be assured that the user will not give (or sell) that token to others.

Having demonstrated how trust is created and managed, one must consider how trust is communicated. This thesis examines various certificate schemes which communicate trust relationships on a computer network. It is argued that if trust relationships are to be
communicated and managed at a sophisticated enough level to be useful, the certificates must be able to represent complex relationships and policies while remaining easy to use and secure. For the majority of the users and purposes of the digital world, this thesis predicts that a solution to this problem will be the creation of well tested default expressions and trust relationships. In this way, unsophisticated users can easily adopt certified default trust relationships – for instance, a user could grab the default trust relationship for conducting commerce from the Digital Better Business Bureau.

Finally, the thesis examined policy with respect to the digital instruments mentioned above. The term precedent dependency is introduced to describe regulatory structures that become dependent on the characteristics of previous technologies to set the requirements of future technologies. The danger is that policy makers will wish to shape digital instruments, information technologies, and communications networks in the form of previous technologies and policies in order to maintain the status-quo. Unfortunately, this may actually cripple new technologies since the distinctions that generated old policies are no longer relevant.

However, one may hope that the development of the Internet, both technically and culturally, will not be hindered by external policy processes. Rather, I hope that the cooperative ethos of the Internet will affect the responsiveness of corporate and political organizations to the needs of their communities. Part of the ethos of the Internet is – amusingly – a rather trusting attitude towards ones' environment. The Internet is a network of networks with few guarantees – one trusts that the next node will pass ones' packets on. Even so, it seems to have worked rather well! So, while the digital world has much to learn from the real world with respect to economics and trust management, the real world may be able to learn something with respect to trust, communication, and community from the Internet.
REFERENCES CITED


