Chapter 1

Introduction

... Detection is, or ought to be, an exact science and should be treated in the same cold and unemotional manner. ...

Arthur Conan Doyle

1.1 Motivation

The highly technical nature of computer crimes facilitated a wholly new branch of forensic science called digital forensics. Instead of dead bodies, digital forensic scientists collect and analyse data produced, transmitted, and stored by digital devices. The aim of digital forensic analysis remains the same – to clarify events of the incident and, ultimately, identify its perpetrators.

At the time of writing, the field of digital forensics is rapidly evolving. Despite having a variety of practical techniques and tools, it provides little theoretical basis to support correctness of investigation findings. The development of such a theoretical basis is seen as an important research problem. In particular, the first Digital Forensic Research Workshop (2001) stated in its report that

> "What is missing in the digital realm is any real theoretical data about the details of transformations involved in moving from reality to a digitally processed representation. ... Trained and certified forensic serologists can comment on the correctness of DNA evidence via explanations that incorporate findings from molecular biology, population genetics, and probability theory. Most analysis in Digital Forensic Science[1] cannot make similar claims." [4]

Since then, significant progress has been made in some areas, such as testing of information copying tools [60] and specification of data examination and analysis tools [22]. Little work, however, has been done on the theory of event reconstruction.

Event reconstruction is the process of determining the events that happened during the incident. It is a fundamental activity in any investigation, because unless investigator determines what happened and how it happened, there is simply no basis for determining why it happened or who may have done it.

In "ordinary" forensics the link between evidence and perpetrator's actions is often straightforward – a fingerprint on the wall indicates that someone has touched the wall, the unique shape of papillar lines can be used to identify the person in question. Common sense reasoning is usually sufficient to analyse events of the incident.

In digital forensics, however, the link between evidence and perpetrator's actions is more complex. A single push of a button triggers a chain of events inside one or more digital devices that produce the digital evidence. Informal, unaided reasoning is not always sufficient to comprehensively analyse this chain of events. According to Stephenson [76], the problem of "inconsistencies in interpreting digital evidence in complex attacks" is a specific problem to be

[1] Currently there is no universally agreed term for digital forensics. Some authors call it digital forensic science, computer forensics, or forensic computing.

solved by the digital forensic community. This research contributes to solving this problem.

1.2 Research objectives

One way to make event reconstruction more rigorous is to employ mathematics. In order to do it, the task of event reconstruction has to be formalised as a mathematical problem. Once this is done, the reconstruction can be performed completely in the formal domain and, possibly, automated.

The aim of this research is (1) to formalise event reconstruction in a general setting, that is, assuming nothing specific about the digital system under investigation or about the purpose of event reconstruction, and (2) to show that this formalisation can be used to describe and automate spelected examples of digital forensic analysis.

1.3 Research idea

To see how the research aim can be achieved, consider the following idea. Many real-world digital systems, including digital circuits, computer programs, and network protocols, can be described mathematically as finite state machines. A finite state machine can be depicted as a graph, whose nodes represent possible system states, and whose arrows represent possible transitions from state to state. All possible computations leading to a particular state can be determined by backtracing transitions leading to that state (see Figure 1.1). In theory, the investigator could perform event reconstruction as follows:

1. Obtain a finite state model of the system under investigation.
2. Determine all possible scenarios of the incident by backtracing transitions from the state in which the system was discovered.
3. Discard scenarios that disagree with the available evidence.

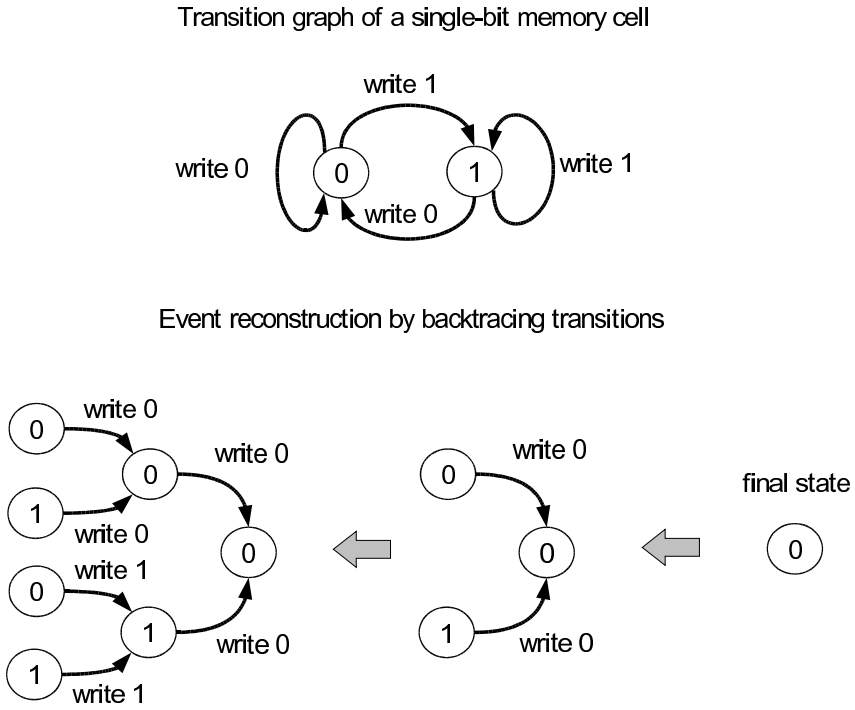

Figure 1.1: Event reconstruction by backtracing transitions

This vague approach is generalised and formalised in this dissertation. The results of this work have been published in [39], which is given in Appendix D.

1.4 Dissertation structure

The structure of the dissertation is schematically shown in Figure 1.2. First, the background of forensic event reconstruction is researched.

Chapter 2 describes relevant concepts from the legal theory. It reviews concepts of legal process, evidence, and proof, and the role of forensic expert in the legal process. It also gives definition of digital evidence and highlights difficulties associated with its use in litigation.

Chapter 3 describes the technical side of digital forensic investigation. It reviews the digital forensic investigative process, classifies its analysis techniques, and highlights the need for effectiveness and efficiency of digital forensic techniques.

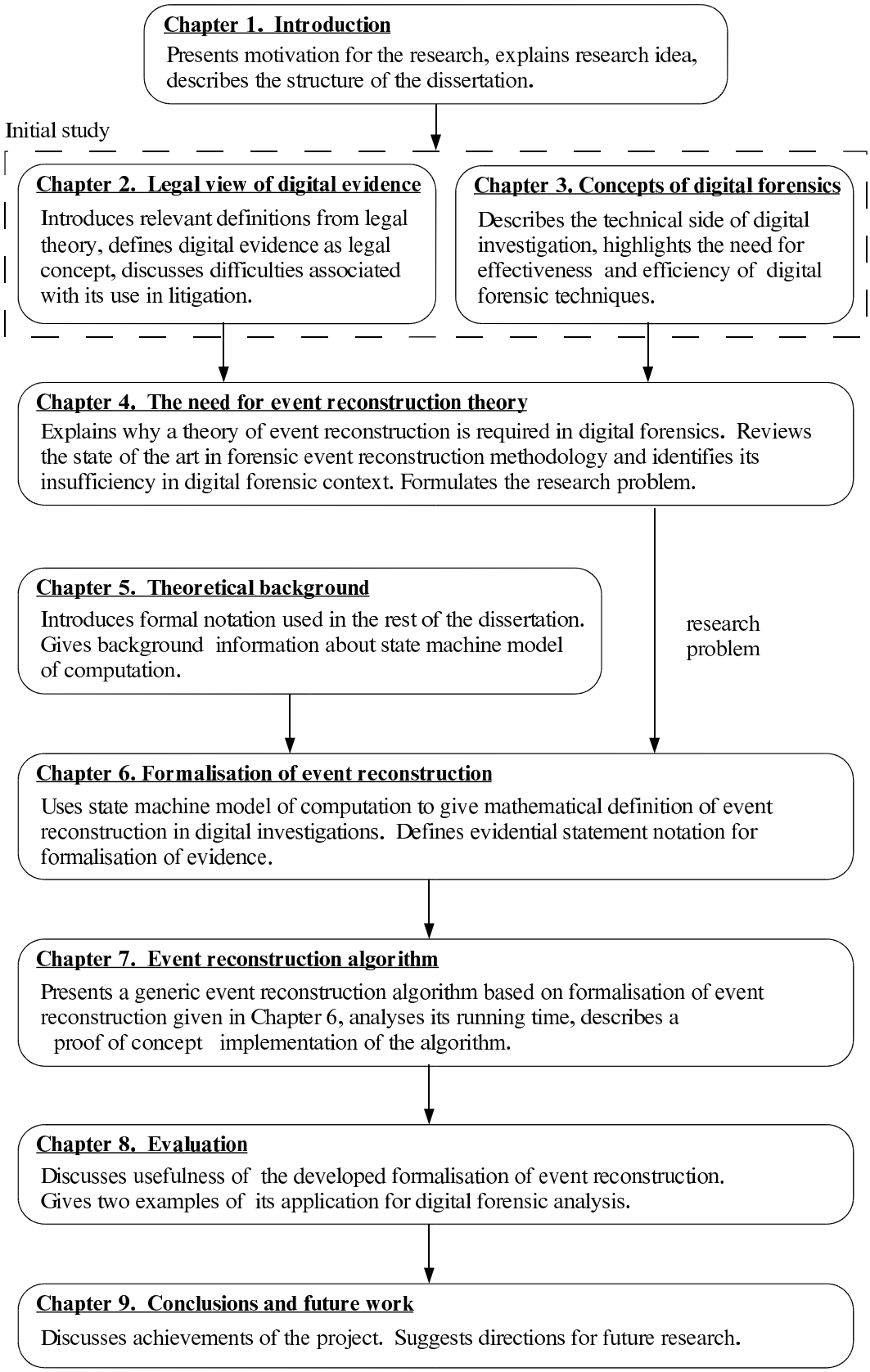

Figure 1.2: Dissertation structure

Second, once the background study is completed, the problem statement is given in Chapter 4. The chapter reviews existing semi-formal reconstruction techniques from related forensic sciences and identifies their deficiency in digital forensic context. Based on this analysis the research problem is formulated at the end of the chapter.

Third, the mathematical model of event reconstruction is developed. After the formal notation and necessary theoretical background is given in Chapter 5, Chapter 6 develops a mathematical definition of event reconstruction problem. It builds a system of formal objects that describe the system under investigation, transition backtracing, and the evidence.

The key result presented in Chapter 6 is the evidential statement notation. It describes the evidence as a system of observations about the past behavior of a finite state machine. The problem of event reconstruction is then defined as finding all possible explanations for the given evidential statement with respect to the given finite state machine.

Fourth, the usefulness of the developed formalism needs to be demonstrated. Chapters 7 and 8 serve that purpose. Chapter 7 constructs a generic event reconstruction algorithm, which is based on the developed formalisation of event reconstruction. It also analyses the running time of the algorithm, and describes a "proof of concept" implementation of the algorithm in Common Lisp. Chapter 8 then uses that algorithm and its implementation to formalise and automate selected examples of digital forensic analysis.

Finally, the conclusions for the entire work, as well as directions for future research are given in Chapter 9.

1.5 Summary of achievements

To conclude this chapter, given below is a list of the key achievements of this project.

- For the first time, a precise mathematical definition of event reconstruction in digital investigations has been given. Apart from providing basis for automation of event reconstruction, it contributes to the development of digital forensics theory as a discipline.
- A generic event reconstruction algorithm has been designed and implemented. Unlike many digital forensic tools, it has solid theoretical foundation, which can be used to defend its admissibility in legal proceedings.
- As example applications of the developed formalisation, two instances of digital forensic analysis have been formalised and automated. They demonstrate why and how formal event reconstruction can be used in practical investigations.
- It has been demonstrated on a practical example, that formal approach to event reconstruction can discover weak points and hidden assumptions in informal event reconstructions.