US

University of Sussex

REPRESENTING ARCHAEOLOGICAL UNCERTAINTY

IN

CULTURAL INFORMATICS

by

Maria Sifniotis

A thesis submitted in fulfilment of

the requirements for the degree of

Doctor of Philosophy

at the

University of Sussex

July 2012

Declaration

I hereby declare that this thesis has not been and will not be, submitted in whole or in part to another University for the award of any other degree.

Signature:

Maria Sifniotis

UNIVERSITY OF SUSSEX

MARIA SIFNIOTIS

Representing Archaeological Uncertainty in Cultural Informatics

This thesis sets out to explore, describe, quantify, and visualise uncertainty in a cultural informatics context, with a focus on archaeological reconstructions. For quite some time, archaeologists and heritage experts have been criticising the often too-realistic appearance of three-dimensional reconstructions. They have been highlighting one of the unique features of archaeology: the information we have on our heritage will always be incomplete. This incompleteness should be reflected in digitised reconstructions of the past.

This criticism is the driving force behind this thesis. The research examines archaeological theory and inferential process and provides insight into computer visualisation. It describes how these two areas, of archaeology and computer graphics, have formed a useful, but often tumultuous, relationship through the years.

By examining the uncertainty background of disciplines such as GIS, medicine, and law, the thesis postulates that archaeological visualisation, in order to mature, must move towards archaeological knowledge visualisation. Three sequential areas are proposed through this thesis for the initial exploration of archaeological uncertainty: identification, quantification and modelling. The main contributions of the thesis lie in those three areas.

Firstly, through the innovative design, distribution, and analysis of a questionnaire, the thesis identifies the importance of uncertainty in archaeological interpretation and discovers potential preferences among different evidence types.

Secondly, the thesis uniquely analyses and evaluates, in relation to archaeological uncertainty, three different belief quantification models. The varying ways that these mathematical models work, are also evaluated through simulated experiments. Comparison of results indicates significant convergence between the models.

Thirdly, a novel approach to archaeological uncertainty and evidence conflict visualisation is presented, influenced by information visualisation schemes. Lastly, suggestions for future semantic extensions to this research are presented through the design and development of new plugins to a search engine.

Acknowledgements

Short contents

Contents

List of Publications

Refereed Conferences

- **Sifniotis**, **M.**, Jackson B., Mania, K., Vlassis, N., Watten, P. L., White, M.: *3D visualisation of archaeological uncertainty*. In Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization, APGV 2010: 162

- **Sifniotis, M.,** Felicetti, A. Needle in the Haystack: finding, indexing and semantically querying image collections of ancient coins. In M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis (Editors). Full Paper Proceedings of the 14th International Conference on Virtual Systems and Multimedia (VSMM), pp. 165—171, October 2008, Limassol, Cyprus.

- **Sifniotis**, **M.;** Watten, P.; Mania, K. & White, M. (2007), *Influencing Factors on the Visualisation of Archaeological Uncertainty*, in D. Arnold; F. Niccolucci & A. Chalmers, ed.,'VAST07: The 8th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage', pp. 79—85.

- **Sifniotis, M.;** Mania, K.; Watten, P. & White, M. (2006), *Presenting uncertainty in archaeological reconstructions using possibility theory and information visualisation schemes*, in M. Ioannides; D. Arnold; F. Niccolucci & K. Mania, ed.,'VAST 2006 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage-Short papers.', pp. 198—203.

- **Sifniotis, M**.; Jackson, B.; White, M.; Mania, K. & Watten, P. (2006), *Visualising uncertainty in archaeological reconstructions: a possibilistic approach*, in 'SIGGRAPH '06: ACM SIGGRAPH 2006 Sketches', ACM, New York, NY, USA, pp. 160.

- White, M., Liarokapis, F., Mourkoussis, N., Basu, A., Darcy, J., Petridis, P., **Sifniotis, M.**, Lister, P.F., ARCOlite — an XML based system for building and presenting Virtual Museums using Web3D and Augmented Reality, IEEE Proc. Theory and Practice of Computer Graphics 2004.

Refereed Journals

- **Sifniotis, M.**, Jackson B., Mania, K., Vlassis, N., Watten, P. L., White. 2011. *Mathematical Modelling and Visualization of Archaeological Uncertainty.* To be submitted to ACM Transactions on Computer Graphics and Visualisation

- Liarokapis, F., Mourkoussis, N., White, M., Darcy, J., **Sifniotis, M.**, Petridis, P., Basu, A., Lister, P.F., Web3D and Augmented Reality to support Engineering Education, World Transactions on Engineering and Technology Education, UICEE, Vol.3, No.1, 11-14, (2004).

Invited Talks

- **Sifniotis, M.** The Good, the Bad and the Ugly: The Wild West of On-line Coin Sales. CHIRON workshop, VSMM 2008, Limassol, Cyprus

- **Sifniotis, M.** COINS: The Search Engine. Workshop presentation (COINS workshop). EVA Florence, 18 April 2008

- Hermon, S., **Sifniotis, M.** Communication of Cultural Heritage in Museum Environment: the challenges of the digital era. In Open Digital Cultural Heritage Systems Conference, EPOCH Final Event, Rome, February 2008 2007

- White, M.; **Sifniotis, M.**; Petridis, P. & Suciu, C. (2007), *'Virtual reality applications for digital heritage: from multimodal interfaces to interactive environments.'* Computer Applications and Quantitative Methods in Archaeology, CAA UK 2007, Chapter Meeting.

- White, M., Petridis, P., **Sifniotis, M.**, (2005) "ARCO–Augmented Representation of Cultural Objects", CAA UK 2005: Computer applications and quantitative methods in archaeology

- Oosterlynck, D., Pletinckx, D., White, M., **Sifniotis, M.**, Petridis, P., Thalmann, D., Clavien, M., "EPOCH ShowCase 2.4.2: Multimodal Interface for Safe Preservation of Valuable Objects", EPOCH Documentation, VAST 2004, Belgium, Brussels.

Invited Articles

- Felicetti, A., **Sifniotis, M.** Numismatic Web Search Tool. COINS Project deliverable D9. September 2008

- **Sifniotis, M.** An Overview of 3D Visualisation Using Game Platforms. 3dVisA bulletin September 2007 issue. (JISC 3D Visualisation in the Arts Network)

List of Tables

List of Figures

Glossary

3D Three-Dimensional

AR Augmented Reality

BECTA British Educational Communication and Technology Agency

CAD Computer-Aided Design

CAVE Cave Automatic Virtual Environment

CDSS Clinical Decision Support Systems

CVE Collaborative Virtual Environment

CSG Constructive Solid Geometry

DST Dempster-Shafer Theory

EBM Evidence-Based Medicine

FHW Foundation of the Hellenic World

FOSS Free and Open-Source Software

FOV Field Of View

KBS Knowledge-Based System

GDK Game Development Kit

GIS Geographical Information Systems

GBT Generalised Bayesian Theorem

HMD Head Mounted Display

HSI Hue - Saturation - Intensity

NPR Non - Photorealistic Rendering

NURBS Non Uniform Rational B-Splines

LOD Level Of Detail

PR Photorealistic Rendering

TBM Transferable Belief Model

UNESCO United Nations Educational, Scientific & Cultural Organization

VR Virtual Reality

VRML Virtual Reality Modeling Language

X3D eXtensible 3D

"Where shall I begin, please your Majesty?" he asked. "Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."

Alice's Adventures in Wonderland

CHAPTER 1

Introduction

Cultural informatics is a broad research area which describes the application of information technology to the heritage sector. The area has been increasingly active during the last twenty years due to the lowering costs of technology as well as the lower barrier of entry. More particularly, the archaeological community has witnessed a sharp increase in Three-Dimensional (3D) reconstructions of ancient structures. Assisted by both advances in technology and lower technical costs, museums, archival institutions and heritage attractions have been using 3D reconstructions more and more often.

However, archaeologists and computer scientists have urged caution [1] in the abundant use of virtual reconstructions because of the possibility of misleading the public. Existing literature highlights the need for visualisations indicating where archaeological finds end and reconstruction begins.

Archaeology is both an uncertain discipline and a destructive process. An archaeological site can never be excavated and interpreted to its exact ancient dimensions. This is a one-off process, since once the site is dug and its evidence removed and processed, the excavation cannot be repeated. Consequently, archaeological hypotheses and interpretations also contain the element of uncertainty.

The archaeological community has stressed the need to acknowledge the availability of other possible hypotheses as well as the difference between what was found and how it is interpreted. As a result, new approaches have come forward that take these alternatives under consideration, and attempt to include them in a virtual reconstruction.

This thesis makes contributions in the area of cultural informatics and specifically in the use of 3D reconstructions as a hypothesis evaluation tool.

The main question of the thesis is:

How can an archaeologist's belief in a reconstruction be quantified, and how this belief can be visualised alongside the 3D reconstruction itself?

Before an answer to the question itself can be explored, a number of key concepts pertaining to the question have to be explained. Firstly, one has to consider, what is archaeology? How do archaeologists interpret their evidence? What are computer graphics? What is the relationship that archaeology has with computer graphics?

Section 1.1 serves as an introduction to archaeology, its changing role through the last century and covers the basics of archaeological theory and inference. Section 1.2 offers a short primer in computer graphics principles such as what makes up a 3D model and the ways that it can be created. Section 1.3 explores the relationship between the discipline of archaeology and computer graphics; how graphics have been used as a tool for various archaeological purposes. The criticisms and suggestions that arose from the uses of computer graphics in archaeology are summarised in Section 1.4. Lastly, Section 1.5 revisits the main question of the thesis and expands on the different contributions the thesis makes.

1.1 Archaeology

Archaeology is the scientific study of human cultures through the excavation, collection and interpretation of cultural and environmental remains [2]. The width and breadth of archaeological study is wide, encompassing areas such as cultural history, human evolution, cultural and individual behaviour, and ecology. This wide expansion of research areas has had two effects: firstly, a difficulty to pinpoint exactly what the aims of the discipline are; secondly, the inclusion and collaboration with a variety of other disciplines such as anthropology, palaeobotany, history, geology, mathematics, computer science, psychology, medicine etc. Even though scientific methods are used, as mentioned above, archaeology forms the study of human cultures; thus this makes the discipline also humanistic-oriented in nature.

The most common conception of the first archaeologists brings to mind antiquarians of the 18th—19th century. In reality, archaeology has existed for thousands of years. Early examples include Flavio Biondo in Europe (15th century), Shen Kuo in

the East (1088 AD), and in Islamic Egypt Ibn Wahshiyya (9th century AD). However, until the 19th century, archaeology was mostly expressed as antiquarianism and the hoarding of strange and ancient objects for personal collections. The first systematic scientific approach to archaeology by means of empirical examination and development of theories, was formulated by J. Winckelmann (early 19th century)–he was the one who first distinguished between Greek, Greco-Roman and Roman art. Pitt Rivers (mid 19th century) was one of the first to develop a typology, a method of classifying artefacts according to their characteristics. William Petrie and Mortimer Wheeler (early and mid 20th century) were two of the most noted British archaeologists due to their meticulous approach to excavation, detailed record keeping and the dissemination of archaeology to the public. They developed the methods of seriation (a relative dating system based on assemblage of artefacts) and the grid system excavation, respectively.

One of the most important technological developments for archaeology in the 20th century is considered to be radiocarbon dating (Libby, 1949). It brought a revolution in archaeological dating, allowing archaeologists to reasonably date organic material and reassess past discoveries. Also about that time, a diversity of ideas emerged, regarding what archaeology is, and how it should be applied. This gave rise to the term *archaeological theory*. In essence, the term encompasses all the intellectual frameworks through which archaeologists interpret discovered material remains.

1.1.1   Archaeological theory

The major schools of archaeological thought over the last two centuries are Culture History, Processualism and post-Processualism:

Culture History (19th—mid 20th century)

Culture History involves the grouping of archaeological sites into distinct cultures, the determination of geographical spreads of population as well as any interaction between different cultures [3]. Each culture in essence describes human behaviour and changes in behaviour could be explained by social and economic exchanges between cultures.

Processualism (1960s—1970s)

Processualism, or New Archaeology was proposed by Lewis Binford in the 1960s [4]. It advocated a more scientific view of archaeology where culture was a set of

behavioural processes. Influenced by other scientific disciplines, the Processualists advocated the use of hypotheses, scientific method and testing against evidence. To them, the view of Cultural History put too much importance on culture over the people themselves.

Post-Processualism (1980s—1990s)

This school of thought, formulated in the 1980s by Hodder, Shanks, Tilley and Miller, rose as a debate to Processualism [5]. Since theories on cultural change can not be independently or experimentally verified, then what is considered true is simply what seems the most reasonable to archaeologists as a whole [6, 7]. Additionally, there are many cases where the same pattern, or data set, could be interpreted in different ways. Hodder, who highlighted this issue based on discoveries on his gathered data, termed it *equifinality*. Lastly, post-Processualism stated that since archaeologists are not perfectly objective, the conclusions they reach will always be influenced by personal biases [8].

Different though they may be, Processualism and post-Processualism do share some similarities. They both try to understand whether the knowledge obtained through study of the past represents the actual past or a possible reconstruction of the past. For this reason, they both promote the idea that interpretation of the past always carry biases which have to be acknowledged and ideally, removed [9]. This element of subjectivity that arises from archaeological inference is quite important and deserves to be analysed further.

1.1.2 Archaeological inference & subjectivity

Consider the position an average archaeologist is in: she[1] has to discover, observe, catalogue, and analyse present material remains (or, the *archaeological record)* belonging to ages long past. The remains, as found today, might be strikingly different than when actually used. The archaeologist has to assign a place, a time, and a recovery context [10] to those remains in order for them to have some sort of meaning. This results in the archaeologist gathering *data* from the remains–the remains are now a source of data. While place and time can (most of times) be measured quite objectively, by for example, stratigraphy or radiocarbon dating, the context is a different matter. She is asked to interpret those remains, and reconstruct the habits and meanings of past societies, but she has never met them and can not observe them. She can

---

[1]*She* equally means *he*.

only make deductions in the here and now of the present. This tentative link, continuously looping, between material remains, their data, and their use as evidence for a conclusion, sums up the archaeological inferential process [11].

Thompson [12] suggests that archaeologists use two stages of inference. The first stage occurs when the material evidence is collected; its significance has to be weighted. Is this piece of evidence important? Does it carry *indicative* [12, :327] *quality*? One can immediately see that this depends on the individual. What one archaeologist might deem as important, perhaps a less experienced one, will not. This is the first element of subjectivity.

The second level of inference is *probative analogy* [13]. Much like its legal counterpart, in the archaeological context it suggests a relationship, a correlation, between some material evidence and a behaviour or event. However, unlike law, where current human behaviour can be correlated with current statistics and data, the archaeologist deducts a past behaviour based on present observations. Thus, the archaeologist has firstly to identify something to correlate her evidence to, and secondly has to prove that the correlation is strong. This is the second element of subjectivity.

The process of archaeological inference is a process that involves subjectivity; it is directly related to the individual, their level of knowledge and expertise, as well as personal biases. This is acknowledged in the archaeological literature while the various theoretical frameworks, despite their different approaches, agree on the subjectivity involved in interpretation.

Around the time of the post-Processualism development, computers began to be used in various areas of archaeology such as quantitative analysis, database management of archaeological data, Geographical Information Systems (GIS), on-site recording methods, computer learning and 3D representation. As a discipline, archaeology focuses a lot on illustration as a medium of presenting results, approaches and deductions. Computer graphics, and the capability for 3D representation mostly excited the imagination, since there was an opportunity to visualise what until now only existed as photographs of ruins, or at best an artist's interpretation. The next section introduces computer graphics principles.

1.2 Computer graphics principles

This section explains notions such as modelling, lighting, texturing and rendering. It also explores the most common ways to create a 3D model. This provides the set-

ting for the subsequent analysis of archaeological projects that make use of computer graphics.

1.2.1 From two to three dimensions

The lowest unit of a two-dimensional picture is a pixel; it represents a single drawing point. It consists of three sub-regions each corresponding to the basic colours: red, green and blue (RGB), such as in Figure 1.1. Thus, a graphics image is made of a rectangular grid of independent pixels. This type of display mode is also known as a raster image. In the late 1970s, Smith and Catmull [14] proposed that the opacity/transparency of a pixel is as important as its colour. They extended the RGB representation by adding the *alpha-channel* value which represents transparency, thus resulting to the RGBA pixel.

(a) A pixel with the basic colours

(b) A pixel with an alpha level.
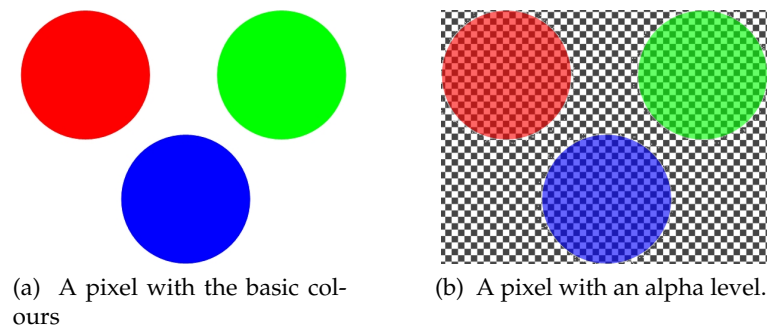
Figure 1.1: Pixels with RGB and RGBA values

Another 2D display method is vector graphical representation where the image is modelled as a set of commands. While no solids can be drawn with this method, one can draw graphs, wireframe images, etc. Vector graphics store the image as a set of commands, rather than pixels. Nowadays, most vector graphics systems actually run on pixel based hardware. As a result, lines are approximated with pixels. Nonetheless, numerous vector systems are able to draw text, curves, and fill regions.

A step further is polygon modelling. This display mode is most commonly used to represent three-dimensional shapes, by constructing a set of connecting polygons to represent an object. The most basic operation of polygonal modelling is the creation of a triangle. Interconnecting polygons representing an object are known as *meshes*. Three-dimensional modelling is not restricted to polygonal representation, albeit being the simplest, but not always memory efficient, technique. Others include parametric cubic curves and subdivision modelling.

Polygon modelling offers linear approximation to surfaces. As a result, modelling a curve would require storing a large number of coordinates in order to achieve some degree of accuracy. Parametric cubic curves (or splines) use higher degree functions that approximate the shape by using only a few points, thus utilising less storage and being easier to alter. Common used techniques in computer graphics are Bézier curves, B-spline curves, and Non Uniform Rational B-Splines (NURBS). However, parametric forms are unable to represent some simple curves such as circles. Bézier curves and B-splines can only represent what polynomial parametric forms can.

By introducing homogeneous coordinates they are generalized to rational Bézier curves and NURBS. Thus, rational Bézier curves and NURBS are more powerful than Bézier curves since the former can represent circles and ellipses. Consecutively, NURBS are more powerful than B-splines.

Bézier curves are defined by control points. Two points represent the start/end of the curve, while the rest specify the curve's direction (Figure 1.2). Bézier curves can be combined, subdivided and be made more complex. Consecutively, Bézier surface patches arise from combining curves in two dimensions instead of one, like in Figure 1.3.
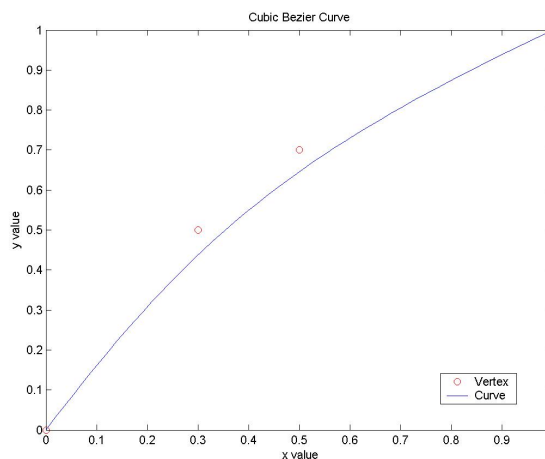
Figure 1.2: A degree three Bézier curve. The dots indicate the control points

B-splines are a generalization of the Bézier curves, and offer advantages over them. They handle an increased number of control points (or *knots*), complicated curves, sharp bend and corners. Like in Bézier curves, the interior knots define the shape of the curve-however each knot has an active range of influence. As a result, the curves can be manipulated with greater accuracy. Uniform B-splines have their
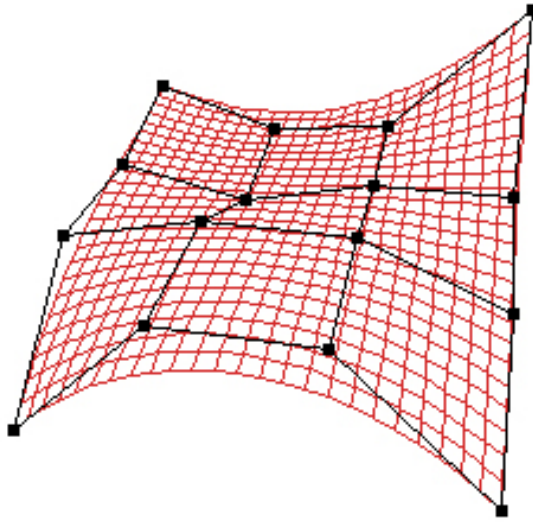
Figure 1.3: A Bézier surface patch. The black dots indicate the control points

knots equally spaced out, while non-uniform do not. NURBS extend the non-uniform B-splines in that they are rational. Homogeneous co-ordinates allow the element of weight to the control point, represented as a factor of w. As with Bézier curves, NURBS and B-splines can be combined to create respective surfaces

Lastly, subdivision surfaces algorithms are used to create smooth surfaces out of arbitrary meshes. This includes meshes composed of triangles, rectangles, and other polygons. There are numerous refinement approaches, among the first proposed being these of Catmull-Clark [15] and Doo-Sabin [16]. Refining a mesh involves inserting new points in between existing ones, therefore increasing the detail and complexity of the object (Figure 1.4).
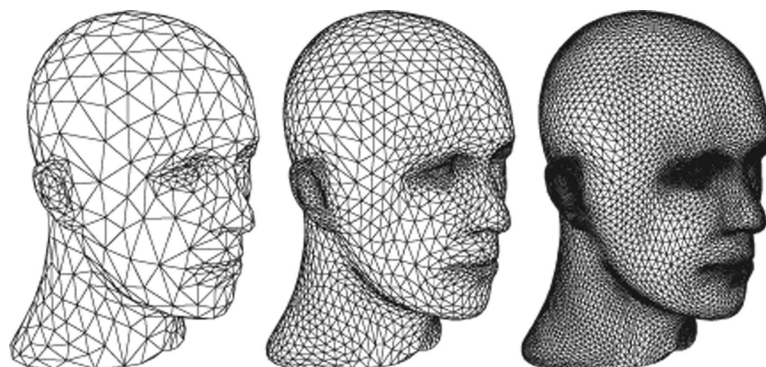
Figure 1.4: Subdivision algorithm (Image from [17])

The aforementioned techniques are commonly referred to as *surface modelling*, because they try to describe an object by its surface properties. An object modelled in such a way does not necessarily contain information about its volume or is a closed volume at all. However, there are circumstances when the object's volume has to be known, so that its geometry and interaction with its environment can be better understood. Properties such as the outside, inside and centre of mass might be required. Solid, or volumetric, modelling is an approach to this case. A solid object should have certain properties [18] to ensure that no invalid representation is created. Well-formed solid objects can be combined together to form new solids by a set of Boolean operations, such as union, intersection and difference [19]. Figure 1.5 serves as an example. This technique is also known as Constructive Solid Geometry (CSG).
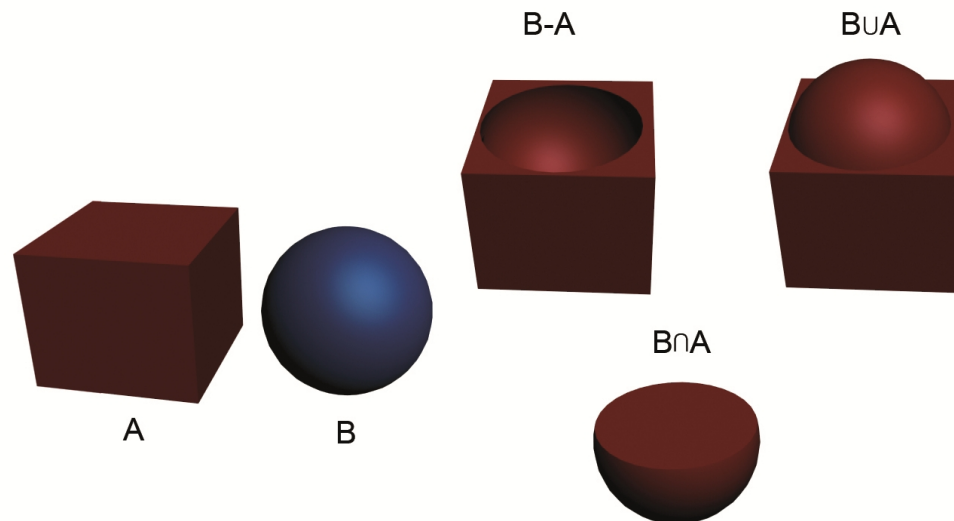
Figure 1.5: Boolean expressions

Having analysed the ways a three-dimensional object can be described, it is necessary to examine the context in which it will be displayed: its orientation in space, shading/texture features, and illumination.

1.2.2   Space orientation

When drawing, either in two or three dimensions, the position of the shape is determined by its coordinates. By specifying the vertices, or connecting points, of i.e. a triangle, we can model it in two or three dimensions.

In a two-dimensional plane, $\mathfrak{R}^2$, a position is specified by its x and y coordinates. Consecutively, a position in a three-dimensional plane $\mathfrak{R}^3$ includes the z coordinate. However, the positioning of the three-dimensional plane's axes is different between computer graphics and the usual mathematical one. As illustrated in Figure 1.6 in computer graphics, the z-axis gives the illusion of depth; it points towards the viewer.
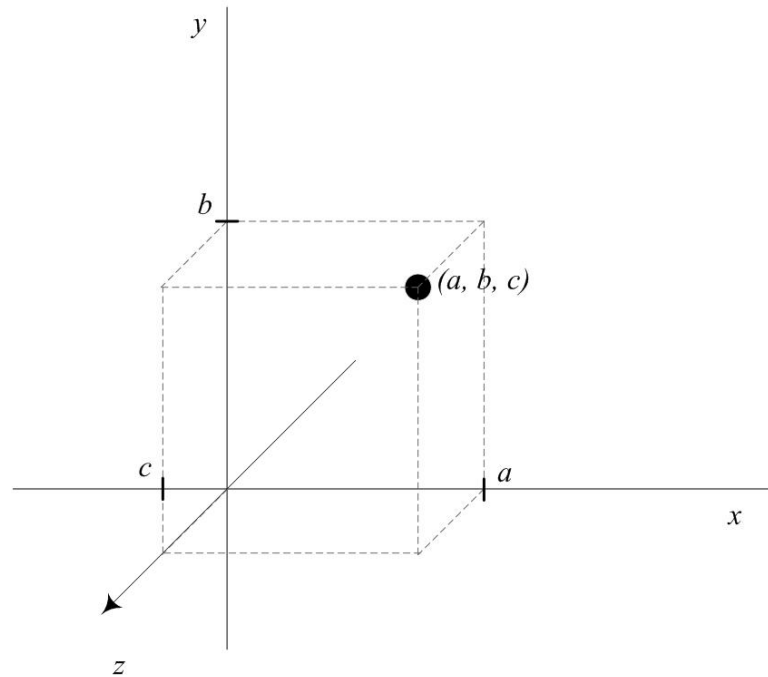
Figure 1.6: Coordinates at $\mathfrak{R}^3$; z-axis points towards the viewer

1.2.3   Texturing

Once a three-dimensional object is created, it undergoes a surfacing procedure. This involves anything affecting its appearance in rendering, and it is composed of *texture* and *material*. Texturing is the procedure of adding an image to an object in order to simulate greater realism without the cost in processing power. A representative of this approach, is bump mapping[20]; as illustrated in Figure 1.7 the coarseness of an orange is represented by a bitmap. There are many methods to achieve this; including advanced texturing that involves combining different textures for one area.

The *material* properties affect the diffuse colour, specularity, ambience, transparency, reflectivity (and other properties) of an object. Diffuse colour is the colour reflected by an object when light falls on it. Specularity defines the glossiness of an

normal texture   bump texture

bump & normal texture   without texture
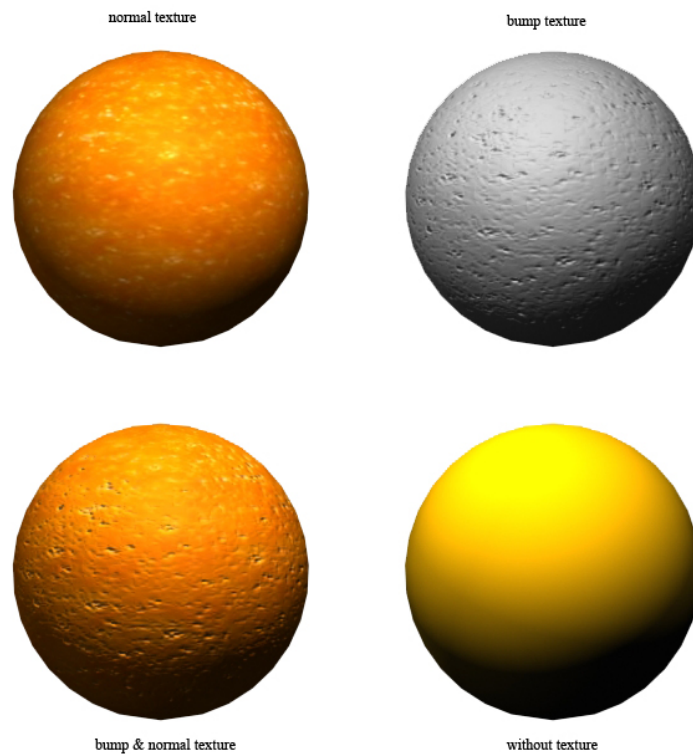
Figure 1.7: Demonstration of basic texturing principles

object. Diffusion and specularity can be thought of as matte and glossy properties. A material's ambience reflects an inner glow emitted by the object. Transparency refers to the object's opacity values, while reflectivity defines how the surrounding world is reflected onto the object.

1.2.4 Lighting

Lighting is an important aspect that adds to the overall visual quality of the 3D graphics scene by enhancing it with shadows and illumination. It is divided into two broad categories: local and global.

Local lighting considers only sources of light shining directly on a surface and then reflected on the viewpoint. A very popular model is the *Phong*[21] lighting and shading model. It is very flexible and can be efficiently implemented both in hardware and software. It works only with point lights and deals with the ways a light

reflects off surfaces. A particular point on a surface is being illuminated by a light source and is viewed from a viewpoint, as demonstrated in Figure 1.8. Having a point light with intensity I^{In}, the model calculates the amount of light that reaches the eye through surface reflection.
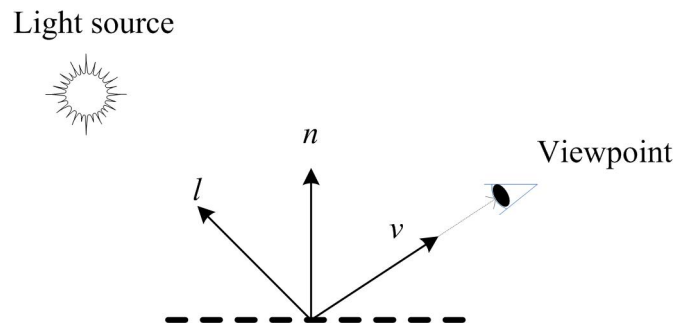
Figure 1.8: The Phong lighting model: n represents the surface normal, l the direction towards the point light source, v the direction towards the viewpoint (after [22]:70)

Thus, it offers two types of reflection: Diffuse and Specular. Diffuse reflection occurs when the light is reflected evenly in all directions, away from the surface and is mostly used on matte (Lambertian) surfaces. Specular reflects the light in mirror-style, according to the shininess of the material. Light is reflected primarily in the direction with the angle of incidence equal to the angle of reflection. The further one rotates away from this direction, the lower the specularity. Figure 1.9 illustrates the two concepts.

Other models for local lighting exist, such as the Cook - Torrance [23], and Schlick [24].

While they are efficient and fast, local lighting models do not consider secondary reflections, for example when light bumps on several surfaces before reaching the viewpoint. Additionally, the shadows they cast are not representative of actual light-casting shadows. Global lighting overcomes some of these limitations by offering two major approaches: ray tracing and radiosity.

Ray tracing [25], attempts to track the specular movement of light. In other words, it follows each individual ray received by the viewer as it reflects off surfaces (or moves through transparent areas [26]) and calculates the reflection direction. This is repeated until the ray reaches the source. Radiosity aims to analyse the balance of lighting in a scene [27] by taking under account the light energy transfer between two surfaces. While ray tracing performs extremely well with specular effects, radi-

Light source   Light source

(a) Diffused reflection: The light is reflec-
ted equally in all directions. The dotted
line indicates the incoming light while the
arrows the diffused light

(b) Specular reflection: The dotted line
indicates the incoming light. The arrows
show the outgoing light; the longer the
arrow, the more intense the reflection to-
wards this direction
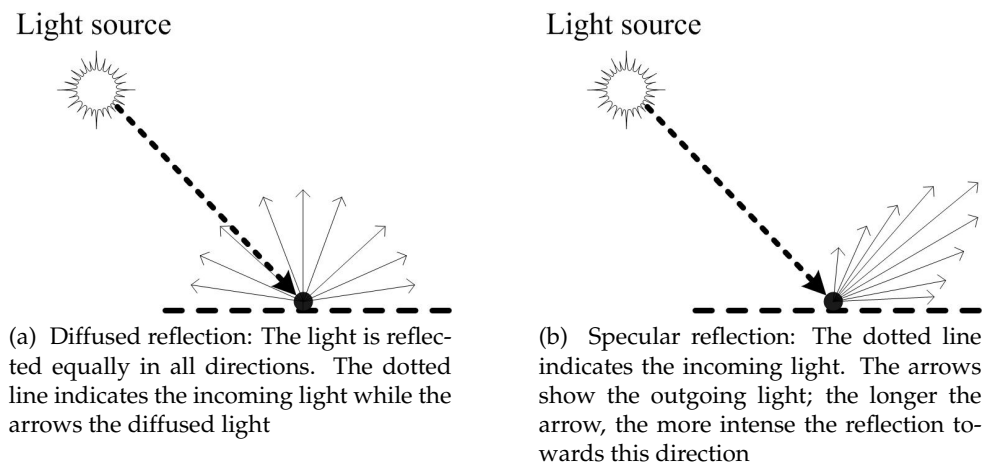
Figure 1.9: Phong diffusion and reflection

osity excels at calculating diffusion, and these two perform very well in combination.
However the computation cost is extremely heavy, often taking hours for proper ra-
diosity/ray tracing calculations.

Until now the aspects considered how a scene is populated with objects and how
these objects are displayed on the screen. The last major property of a scene is its
viewpoint. The viewpoint represents what the observer sees and is often likened to a
camera view, with a direction, position and Field Of View (FOV). It determines what
parts of the 3D objects will be visible in the final image, in what order and which will
be clipped from view.

1.2.5   3D content creation

There is an abundance of software that can be used to create 3D models. Despite this
variety, however, the software broadly falls under two categories:

Programming & Procedural

In order to draw a 3D model on the screen, the user has to actually enter the mathem-
atical equivalents of each vertex or other information (e.g. colour) of which the object
is comprised. Special manoeuvres including translation/rotation/scaling of an ob-
ject must also be represented mathematically. As the scene becomes more complex,
the programming difficulty increases. Essentially, it is a set of functions that aid a

programmer into creating 3D scenes from simple primitives. Well-known representatives of this category are OpenGL and Direct X. For example, the following lines in OpenGL create a polygon and assign a green colour to it:

```
glBegin(GL_POLYGON);    /* Begin issuing a polygon */
glColor3f(0, 1, 0);     /* Set the current color to green */
glVertex3f(-1, -1, 0); /* Issue a vertex */
glVertex3f(-1, 1, 0);  /* Issue a vertex */
glVertex3f(1, 1, 0);   /* Issue a vertex */
glVertex3f(1, -1, 0);  /* Issue a vertex */
glEnd();
```

An extension of this is procedural modelling. It is a technique that uses algorithms and rules to produce textures and complex models. An example of this is the generation of terrain from a greyscale heightmap. Procedural modelling allows textures and models to be created in the run-time of an application. These rules and algorithms describe the properties of the model. For example, an algorithm for a procedurally-generated house may define a range for the height of its walls, number of windows, entrances, etc. This technique is often used in computer games to generate random content and introduce variety at a much lower cost.

Modeling Software

In this category, modelling is achieved mainly by using specialised software and rarely any hard-coded programming. Creation is achieved by using artistic techniques in combination with an in-depth knowledge of the software package. Users can add, subtract, stretch and otherwise change the object as desired. Models can be viewed from a variety of angles, usually simultaneously. An example modeller is 3ds Max by Autodesk. Modelling capabilities include polygon, mesh, NURBS, CSG and others. It also features lighting, texturing and rendering techniques. Figure 1.10 shows the working area of the software. Other popular options include Autodesk Maya and Blender.

Scanning

To create a 3D model with scanning, specialised equipment, such as a laser scanner has to be used. The scanners work with a real-world object, gather its surface data and with customised software, extrapolate the 3D shape of the object. The use of 3D scanners has found widespread appeal from manufacturing, to medicine and cultural heritage.
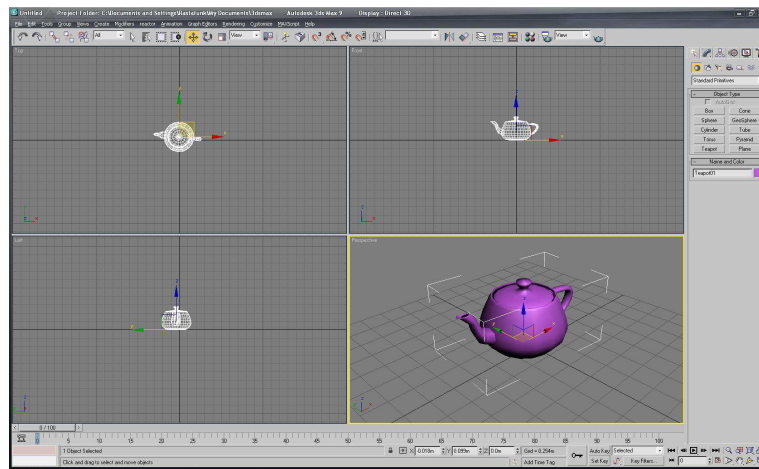
Figure 1.10: 3ds Max modelling environment; the four panels show different views of the object simultaneously

Stereophotogrammetry

Specialised software can perform 3D data acquisition by using photographs from accurately placed cameras, or video sequences. This technique is known as stereophotogrammetry.

1.3 Computer graphics & archaeology

The previous section presented the main elements that make up a 3D model, as well as the ways it can be created. This one examines the uses that computer graphics has found in archaeology. The section looks at the what, how, and why. More specifically:

- What projects and research paths have developed that blend computer graphics with archaeology?

- How are the project being created? Which are the main technologies that have been used?

- Why are computer graphics useful for archaeology? What did the archaeologists think about it?

1.3.1 Projects

Projects relating to archaeology and computer graphics can be categorised into four broad areas: Preservation and collection; site enhancement and promotion; education

and learning; hypotheses and evaluations. These areas are not mutually exclusive, in fact, numerous projects can be considered to belong in two or three concurrently. To illustrate the breadth of the research, a number of representative projects of each area are presented.

Preservation and collection

Many reconstructions deal with digital preservation of the ancient site. In such cases, a reconstruction is made in order to keep a digital archive of the site/object and bring together elements that may be dispersed in various locations (i.e. different museums). For example, in [28] the authors use CSG to preserve a Japanese temple along with several unique items. In [29] the authors created digital models from the Parthenon's sculptures scattered around the world, and assembled them together in a digital environment.

Digital collections also exist from archaeological excavations. When an area is excavated, the data is meticulously catalogued and often incorporated in GISs and databases. Examples can be found in the excavations at Petra, Jordan ([30], [31]) and Çatalhöyük, Turkey [32].

ARCO (Augmented Representation of Cultural Objects) [33, 34], was a European project which provided a tool chain for museums to digitise and catalogue their collections. Through the tools the museums were able to create virtual exhibitions with an online presence. explorable through 3D representations and augmented reality.

Site enhancement and promotion

Projects that enhance a visitor's experience to the site fall under this category, as well as those dealing with heritage promotion and dissemination. At least two European projects, Archeoguide [35] and LifePlus [36] have developed mobile systems for enhancing a visitor's experience of an archaeological site. In the first case, the visitors personalise their tour and explore the site aided by information provided through a Head Mounted Display (HMD), which includes superimposed reconstructions on top of actual remains. In contrast, LifePlus, transmits information to a customised PDA. In [37] the authors introduce an immersive, mobile, Augmented Reality (AR) environment, where visitors to Pompeii are able to witness through a HMD virtual flora and fauna of the Roman period within the context of the real buildings.

Turning to site promotion, UNESCO, the United Nations organisation protecting and disseminating the world's heritage, offers virtual tours around the word's pro-

tected monuments available on the Web [38]. Additional examples can be found in [39] which promotes a virtual set of ancient Olympia including an informative web site.

Education and learning

Educational projects and those that promote learning of heritage material belong in this category. The Foundation of the Hellenic World, a non-profit cultural institution located in Greece, aims to educate and project learning on Greek history and culture. The institution deals with 3D reconstructions [40], immersive Virtual Reality (VR) and table top interactive environments [41]. Its strongest educational medium is considered to be *Kivotos*, a fully immersive 3x3x3 Cave Automatic Virtual Environment (CAVE) in which users navigate with the use of a wand.

The Ename Centre for Public Archaeology fostered the *Ename 974* project. *Ename 974* focused on educating and understanding of Flanders history and culture [42] through VR and multimedia techniques. This was achieved by an on-site VR system located at the ruins of the Benedictine abbey church at Ename. The first prototype, functioning as a kiosk outside the church, allows visitors to view a superimposed three-dimensional model of the church over the actual remains. The second prototype was installed outside a standing monument–an early Romanesque church closed to the public for restoration. The visitors were able to follow the excavations and restoration procedures of the monument as well as virtually visit the closed space.

Hypotheses and evaluation

This category is the one most relevant to the topic of this thesis. Under this category belong projects examining archaeological questions and providing specific tools to aid archaeological research. A very useful tool for examining ritual spaces is Photorealistic Rendering (PR). By introducing elements of accurate lighting and shadows in a space, researchers are able to witness the visibility of the area. One of the first projects was INSITE [43] which examined Bronze Age ritual complexes in Malta and tested two hypotheses for the organisation of ritual performance based on the visibility between priests and audience. A further prospective of PR is the simulation of ancient light sources. In [44], the authors simulated the lighting of various candle types by measuring the actual spectral data and incorporating them to illuminate the preserved frescoes of House Vetii in Pompeii.

Another example can be found in the modelling of the Roman theatre in Can-

terbury. The computer model revised and highlighted faults in earlier popular illustrations and drawings of the theatre, specifically, the seating area was wrongly proportioned [45].

In [46] the authors developed STRAT, a tool to aid archaeologists recording actual stratigraphic data and part of the 3D Murale European Project. STRAT allows archaeologists to record the stratigraphic data layer by layer (including any artefacts) as the excavation occurs, and visualise the data both in two and three dimensions.

1.3.2 Technologies

A little while after new technologies are introduced, there are to be found in virtual archaeology projects. This section looks at some of the most common technologies that have been used, by focusing on those that feature a higher level of interaction. Highly interactive projects include those that might utilise some sort of avatar, or sense of being there to the user. The 3D environment is navigable and changes might occur due to user input. The ones featuring a low level of interaction might include specially-created websites, digital repositories, hosting static, rendered images, or videos of 3D renders and are not examined here.

The early years

Some of the first examples can be found in the reconstructions of the temple precinct from Roman Bath [47] and the Roman bath houses at Caerleon [48]. Both were created by using CSG, and used as case studies to test the features of ray-casting algorithms. In a research project by IBM, solid modelling was used to reconstruct the Saxon Old Minster of Winchester [49], while similar techniques were employed in the reconstruction of Furness Abbey [50].

While the previous examples constructed the model using area plans and elevations, the reconstructions of the Dresden Fraüenkirche [51] and the Visir Tomb in Saqqara, [52], utilized photographs of existing buildings and items as basis for texturing.

VRML

The Virtual Reality Modeling Language (VRML) is an ISO-standardised, nonproprietary, text markup language that describes interactive 3D scenes. Originally developed in the 90s, it had its first version out in 1995 [53] and its second, standardised version (VRML 2.0 [54]) in 1997.

The structure of a VRML scene consists of nodes that describe what is in the scene. For example, the following code snippet constructs a 3D shape (a sphere) and applies material properties (RGB colour and transparency). It can be thought of as a higher level of programmable modelling, where the user can both include custom 3D models or create complicated models from primitive solids. Scenes created with VRML can be interacted with through special viewers and can be accessed both through the web and off-line.

```
#VRML V2.0 utf8

# A Sphere
Shape {
        appearance Appearance{
                material Material {
                        diffuseColor 0.8 0.8 0.8
                        transparency 0.9
                }
        }
        geometry Radius {
                radius 1.5
        }
}
```

VRML, around the time of its standardisation, started attracting a number of archaeological research projects that were being active on the Internet [55]. VRML's most appealing features were its ability to run on the web, its interactivity and low cost of implementation. Amongst the first applications were cases studies using archaeological data in VRML [56], the reconstruction of Canterbury theatre [57], and the English Heritage/Superscape's reconstruction of Stonehenge. The next 10 years that followed saw a plethora of VRML reconstructions, some of them being the tomb of Sen-nedjem [58], the Appian Way [59] and the Avebury project[60]. VRML was also used as a means of navigating virtual collections of online museums, examples can be found in the ARCO project [61]. With the release of X3D, the new 3D standard based on XML, the focus slowly shifted.

X3D

X3D [62] is a specification and description language for 3D objects based on XML. First introduced in 1999 and standardised by ISO in 2004, it also describes 3D objects and scenes. X3D was created to be a replacement of VRML, fixing a number of issues and bringing numerous extensions (such as animations and NURBS) to the format. It

has a well-defined XML syntax but one can also use the VRML 97 syntax if so desired. An example X3D document of a sphere is presented.

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE X3D PUBLIC "ISO//Web3D//DTD␣X3D␣3.2//EN" "http://www.web3d.org
    /specifications/x3d-3.2.dtd">
<X3D profile="Interchange" version="3.2" xmlns:xsd="http://www.w3.org
    /2001/XMLSchema-instance" xsd:noNamespaceSchemaLocation="␣http://www
    .web3d.org/specifications/x3d-3.2.xsd␣">
<Scene>
        <!-- A Sphere -->
        <Shape>
                <Sphere radius='1.5'/>
                <Appearance>
                 <Material diffuseColor='0.8␣0.8␣0.8' transparency='0.9'
                    />
                </Appearance>
        </Shape>
</Scene>
</X3D>
```

The strong relationship of X3D with XML and other open source standards like DOM/XPath made it much easier than before for users to create importers and exporters between their preferred 3D modelling software and X3D. It also enabled them to create custom attributes and nodes that could accompany 3D models. Archaeologists soon saw benefits in this new format. Nicolluci [63] proposed that X3D, because of its capacity to include custom nodes, would be an excellent candidate for describing archaeological models. Hetherington [64] applies XML and X3D technology to a 3D model of Stonehenge. The application allows the user to choose and navigate in real-time through different temporal phases of the model. By choosing to visualise buildings of Sao Polo with X3D, Cabral [65] suggested that the design and validation process was much more efficient. It also allowed for rapidly prototyping candidate models and offer them to the expert historians and archaeologists for validation.

Game Engines[2]

The use of game engines as a medium for 3D representations has been a relatively new direction for archaeology, but not for other disciplines. Their application in scientific modelling [67], simulations and education ([68], [69], [70]) has been well documented.

---

[2]This section is a very condensed summary of an invited technology report written by the author [66]. It offers an introduction to the workings of a game engine and an extensive description of archaeological projects using them.

Jacobson [71] created a set of modifications for the *Unreal* engine and extended it for use in a customised CAVE environment. He then applied it on a visualisation of an Egyptian tomb [72]. The Notre Dame Cathedral project, based on the *Unreal 2.0* engine, was the first to show how a complex historical architectural structure could be imported into a game engine [73]. Erik Champion explored how a virtual cultural heritage scenario can evoke a sense of belonging to its users, by focusing on an ancient Mayan site at Palenque, Mexico [74, 75, 76].

In the AERIA project [77] the *Quake 2* game engine was used to reconstruct a Bavarian Roman fortress and a classical Athenian farmhouse. Additionally, the *Half Life* and *Morrorwind* engines were used to reconstruct the palace of Nestor in Pylos and the throne of Apollo, respectively. The authors postulated that game engines can actually serve as a low cost but powerful tool for heritage visualisation.

This sentiment is followed by Anderson [78] who suggested that as the game industry keeps investing in its own technology, it will also keep it updated for heritage use. He used the *Quake 3* engine to recreate a Pompeian house from archaeological plans. The same engine was also used by researchers from the University of Aizu, Japan [79].

The Rome Reborn 2.0 project has used the City Engine [80] to reconstruct ancient Rome with about 7000 buildings[81]. By using a blend of manually-created and procedurally-generated buildings [82] it has been one of the longest in duration and biggest in scope heritage projects featuring a game engine.

1.4 Criticisms & suggestions

The previous sections demonstrated the wealth and breadth of projects and technologies used in virtual heritage. If one looks at the history, along with advances in the world of IT, the archaeological community has been exploring the benefits of 3D modelling since the 1980s.

In 1990, P. Reilly firstly proposed the concept of *virtual archaeology* [83], referring to 3D models of archaeological formation. The word *virtual* was to represent a replica, a surrogate that would efficiently describe the original formation. Reconstructions were made in effect to describe the original artefact as closely as possible, by putting more and more effort in realistic representations and detail.

Ambitious though they were, these first attempts were soon met with criticisms from the archaeological community; the archaeologists feeling that the reconstruc-

tions promoted advances in computer technology rather than archaeological information. While visualisation could offer much in research areas of such a visual discipline as archaeology, until that time it was used to present already existing knowledge to the public, offering little insight to archaeological questions. These criticisms were effectively summarised in a well-known publication by Miller and Richards in 1995 [1].

Furthermore, due to the high cost, time and effort in producing 3D graphics content at that time, most of the projects were undertaken by IT specialists; the archaeologists had no direct control over the models created. There was concern that the past was presented as a known reality, especially when the models looked extremely realistic [84]. If a rigid and very realistic model of the past is given, two major drawbacks are faced: one politically correct view of the past is projected (thus dismissing alternate realities) and the public is restricted in its ability to criticise and evaluate the interpretation. When compared to hand drawn illustrations, reconstructions carried a greater degree of authority, thus leading to singular views of the past. However, efforts to represent alternative models, or difficulties in interpretation were minimal [85].

Now, a bit less than twenty years after the criticisms by Miller and Richards were drawn, the stage is not yet clear. The rapid technological development has made tools of the 3D trade more common and low-cost than before and brought advanced graphics hardware to the masses. As a result, there has been a wide spread increase in projects dealing with *virtual archaeology*.

Some conferences and journals are particularly dedicated to the area (VAST, CAA, Internet Archaeology), while others (VSMM, EG, IEEE Computer Graphics) actively receive input from virtual heritage research. Since the 5th framework onwards, the European Community fosters a wider interest for cultural heritage and has actively funded projects that cross between heritage and informatics.

Gillings [86] draws attention to the usual development pattern of archaeological visualisation:

> *"Sophisticated VR models are created largely because we can, and are always generated as finished and freestanding, in effect, end products"* ([86, 18])

He suggests that archaeological VR[3] (Virtual Reality) in archaeology should not be

---

[3]His use of the term *virtual reality* is not just restricted to the definition of VR that involves a strong amount of user presence and the use of head-mounted display or stereoscopic glasses. It is a blanket

seen as *eye candy,* an expensive and marginalised technology displaying only visual expectations. Rather, he proposes a change of focus towards explorative and flexible environments. Archaeological VR should encompass imitation, representation as well as involvement and creativity.

In the same fashion, Goodrick, Earl, and Wheatley [60, 87] caution against the reluctance of the archaeologists to facilitate the use of visualisation technologies as standard techniques in their work. They define three misconceptions: virtual archaeology is costly; there is a lack of readily available information for archaeologists to create VR projects; virtual archaeology techniques require great technical skills. On the contrary, they suggest that tools, hardware and software has evolved to a point where it can be available to archaeologists.

It is also encouraging that numerous bodies have come forward with propositions for a more rigorous and documented representation of 3D objects. Some of these include: the AHDS Guides to Good Practice for CAD (2002) and Virtual Reality (2002), Virtual Archaeology Special Interest Group (VASIG) and the Cultural Virtual Reality Organisation (CVRO). What all of the above share is a belief that 3D visualization methods should be applied only where needed, with a scholarly rigour and distinctions should be made between evidence and hypothesis.

In 2005, The London Charter (TLC) was created. It is a document that establishes a set of *"principles for the use of three-dimensional visualisation by researchers, educators and cultural heritage organizations"*. These principles are briefly summarised here:

1. Subject Communities: Aims and objectives of TLC are valid across domains in which 3D visualisation can be applied to cultural heritage.

2. Aims and Methods: A 3D visualisation method should normally only be used to address an aim when it is the most appropriate available method for that purpose. Evaluation of methods is required to determine suitability.

3. Sources: Relevant sources pertaining to a 3D visualisation should be identified and evaluated in a structured way. Sources are defined as all information, digital and non-digital, considered during, or directly influencing, the creation of the 3D visualisation outcomes.

---

term, analogous to Reilly's Virtual Archaeology. The term *archaeological VR* encapsulates all realistic-looking archaeological reconstructions–from simple video to areas navigable by keyboard/mouse as well as those that require extensive use of HMD.

4. Transparency requirements: Information should be provided to allow 3D visualisation methods and outcomes to be understood and evaluated appropriately in relation to the contexts in which they are used and disseminated.

5. Documentation: Process and outcomes of 3D visualisation creation should be sufficiently documented to enable the creation of accurate transparency records, potential reuse of the research conducted and its outcomes in new contexts.

6. Standards: Appropriate standards and ontologies should be identified, at subject community level, systematically to document 3d visualisation methods and outcomes to be documented, to enable optimum inter- and intra-subject and domain interoperability and comparability.

7. Sustainability: Ensure long-term sustainability and preservation of 3D representations by digital archiving or other means.

8. Access: Ways in which the outcomes of a 3D visualisation could contribute to the wider study, understanding, interpretation and management of cultural heritage assets.

While a number of positive opinions have been expressed, only a few offer insight on practical applications of TLC. Nevertheless, TLC has brought to the limelight the need to accurately document and justify the sources and choices used when creating 3D representations of cultural spaces and objects.

In summary, the archaeologists' critique is that in their majority of virtual heritage applications, computer graphics have been used recklessly and mostly for the purposes of just using a cutting edge technology. This technology, however, has great potential to be used as a research and documentation tool, that will be transparent and make use of best practices. Three-dimensional models should be created to solve actual archaeological problems and they should make their construction process clear [88]. Archaeological theory, too often forgotten about, should be inherently related to uses of computer science in archaeological problems.

This thesis takes those concerns and suggestions under consideration and makes contributions in this area. To reinstate the thesis question:

How can an archaeologist's belief in a reconstruction be quantified, and how this belief can be visualised alongside the 3D reconstruction itself?

The inferential process, the ambiguity, that are involved in archaeological interpretations, how can they be included, encapsulated, in 3D reconstructions? How can one quantify their own belief? The next section examines the thesis contributions analytically.

1.5   Thesis contributions

The contributions belong to the area of cultural informatics and specifically the uses of digital reconstructions as a research tool for archaeology. As discussed in previous sections, archaeologists piece together available information derived from evidence into a speculative view of the past. This view becomes more certain as evidence increases.

The thesis researches methods to quantify and visualise the archaeological expert's uncertainty in the reconstructed interpretation. This approach offers visualisations based on an archaeologist's knowledge and evidence on reconstructed parts. The major and minor contributions are presented in Tables 1.1 and 1.2.

Contribution	The design, development and distribution of a novel questionnaire that examines archaeological uncertainty in interpretation during excavations of structures.
Importance	There had been no previous work relating to perception of, and evidence related to, archaeological uncertainty, even though it had been identified as a matter of importance to archaeologists. (Published in [89])
Contribution	Archaeologists across different disciplines are equally concerned about uncertainty in interpretations and reconstructions.
Importance	This indicates that uncertainty is an important issue that cuts through different areas of archaeological training. (Published in [89])
Contribution	Archaeologists indicate preference among different categories of evidence.
Importance	This indicates that an element of bias can exist towards certain evidence and should be made transparent in reconstructions. (Published in [89])
Contribution	Application, for the first time, of three mathematical uncertainty models to archaeological uncertainty/belief representation for the purposes of reasoning visualisation.
Importance	Previous approaches to visualisation of uncertainty in 3D reconstructions offered no justification to the values, or meanings of uncertainty. (Parts published in [90, 91])
Contribution	Innovative visualisation of belief and conflict elements on 3D archaeological reconstructions by using information visualisation schemes.
Importance	For the first time, a way is provided with which experts can represent belief in the reconstruction as well as conflict that weak or contradicting evidence brings to an interpretation. A number of information visualisation schemes are applied and evaluated to both an archaeological case study and experimental models. (Parts published in [90, 91, 92])
Contribution	Design and development of new extension to a search engine for searching for archaeological image and text data.
Importance	The extensions to the engine allowed for the aggregation and indexing of images which was unavailable before. Further extensions allow for the integration of semantic information that will be useful in visualisations of uncertainty. (Published in [93, 94])

Table 1.1: Thesis major contributions

Contribution	A previous approach to archaeological reliability representation through fuzzy logic is not conceptually valid.
Importance	Different kinds of uncertainty exist. It is very crucial to choose the right model for the right uncertainty. Fuzzy logic is well suited for other purposes, but not subjective belief reasoning.
Contribution	Suggestions that a probabilistic approach is not capable of representing archaeological uncertainty are unsustainable. It can be done through a Bayesian approach.
Importance	The Bayesian representation of subjective belief is used through a wide number of disciplines. The previous approach examined conditional probability but without Bayes' rule which updates with evidence.
Contribution	Explanation and illustration of the quantification models through novel archaeological paradigms.
Importance	Relevant archaeological examples were crucial for placing uncertainty quantification in the right context.

Table 1.2: Thesis minor contributions

1.6 Thesis structure

The thesis contains eight chapters. This Chapter offered a general introduction to the area, pinpointed the thesis's area of interest, and analysed the contributions. Chapter 2 focuses on the representation of uncertainty in archaeology and other disciplines. It looks at previous approaches to the visualisation of archaeological uncertainty in reconstructions. Other disciplines, such as law, GIS and medicine are also examined for their approaches towards representing belief. Lastly, the chapter examines the visualisation methods used to represent uncertainty, such as information visualisation schemes.

Chapter 3 includes the structured evaluation of archaeological uncertainty based on a questionnaire, the evidence factors, interviews conducted with archaeologists, questionnaire design and derived results. It concludes with a discussion and definition of archaeological uncertainty related to reconstructions.

Chapter 4 presents a mathematical method for quantifying uncertainty based on Subjective Probability. The model is then applied to an archaeological scenario. Chapter 5 presents two further mathematical methods of quantifying uncertainty, based on Evidence Theory and Possibility theory. It then proceeds on applying both of them to archaeological examples.

Chapter 6 begins by explaining why previous approaches based on fuzzy logic are not adequate to handle uncertainty and why probability theory is one way of doing so. The Chapter also contains a series of experiments conducted on possibilistic, probabilistic, and belief results in order to determine if a relationship exists between them. Lastly, it presents and analyses a series of uncertainty visualisations based on the three uncertainty quantification models.

Chapter 7 presents research results on the design and development of a search engine with a focus on identifying ancient coinage. Discussion of results includes a connection between semantic annotation of uncertainty information and semantic descriptions in search engines.

The last Chapter presents the thesis conclusions and discusses future work.

1.7 Published Work

Parts of this thesis have been published or presented in:

- **Sifniotis, M.,** Jackson B., Mania, K., Vlassis, N., Watten, P. L., White, M.: 3D

visualization of archaeological uncertainty. APGV 2010: 162

- Felicetti, A., **Sifniotis, M.** Numismatic Web Search Tool. COINS Project deliverable D9. September 2008

- **Sifniotis, M.,** Felicetti, A. Needle in the Haystack: finding, indexing and semantically querying image collections of ancient coins. In M. Ioannides, A. Addison, A. Georgopoulos, L. Kalisperis (Editors). Full Paper Proceedings of the 14th International Conference on Virtual Systems and Multimedia (VSMM), pp. 165—171, October 2008, Limassol, Cyprus.

- **Sifniotis, M.** The Good, the Bad and the Ugly: The Wild West of On-line Coin Sales. CHIRON workshop, VSMM 2008, Limassol, Cyprus

- **Sifniotis, M.** COINS: The Search Engine. Workshop presentation (COINS workshop). EVA Florence, 18 April 2008

- **Sifniotis**, M.; Watten, P.; Mania, K. & White, M. (2007), *Influencing Factors on the Visualisation of Archaeological Uncertainty*, in D. Arnold; F. Niccolucci & A. Chalmers, ed.,'VAST07: The 8th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage', pp. 79—85.

- **Sifniotis, M.** An Overview of 3D Visualisation Using Game Platforms. 3dVisA bulletin September 2007 issue. (JISC 3D Visualisation in the Arts Network)

- White, M.; **Sifniotis, M.**; Petridis, P. & Suciu, C. (2007), *'Virtual reality applications for digital heritage: from multimodal interfaces to interactive environments.'* Computer Applications and Quantitative Methods in Archaeology, CAA UK 2007, Chapter Meeting.

- **Sifniotis, M.**; Mania, K.; Watten, P. & White, M. (2006), *Presenting uncertainty in archaeological reconstructions using possibility theory and information visualisation schemes*, in M. Ioannides; D. Arnold; F. Niccolucci & K. Mania, ed.,'VAST 2006 7th International Symposium on Virtual Reality, Archaeology and Cultural Heritage-Short papers.', pp. 198—203.

- **Sifniotis**, M.; Jackson, B.; White, M.; Mania, K. & Watten, P. (2006), *Visualising uncertainty in archaeological reconstructions: a possibilistic approach*, in 'SIGGRAPH '06: ACM SIGGRAPH 2006 Sketches', ACM, New York, NY, USA, pp. 160.

Knowing ignorance is strength.
Ignoring knowledge is sickness.

_____

Lao Tzu,*Tao Te Ching*,Ch.71

CHAPTER 2

Uncertainty Literature

2.1   Introduction

In this Chapter, the literature background of uncertainty is examined. Firstly, a background to the research of archaeological uncertainty visualisation is provided, and the lack of uncertainty representation models is highlighted. Then, the concept of uncertainty as well as its forms of expressing it are examined from a philosophical and cognitive point of view. Following, advances in information visualisation are presented with relation to uncertainty. Lastly, a number of different visualisation and quantification approaches to uncertainty by a variety of disciplines such as law, medicine and GIS, are presented.

2.2   Uncertainty and archaeological reconstructions

Chapter 1 introduced the concerns raised by archaeologists with regards to three-dimensional reconstructions of archaeological structures. This section expands on these issues by looking at previous approaches to alternative reconstructions and the topic of uncertainty.

In [57], the authors create an interactive model of the Roman Canterbury theatre using VRML. Not only is the user able to change views and navigate in the model, but also to interact and manipulate objects in the world. In this way, alternative interpretations of the theatre are embedded in an interactive model. By manipulating side sliders, the user is able to change the theatre's height, the seat height, and the seat depth.

(a) The walls follow a uniform texture

(b) A white line differentiates between remains and reconstruction
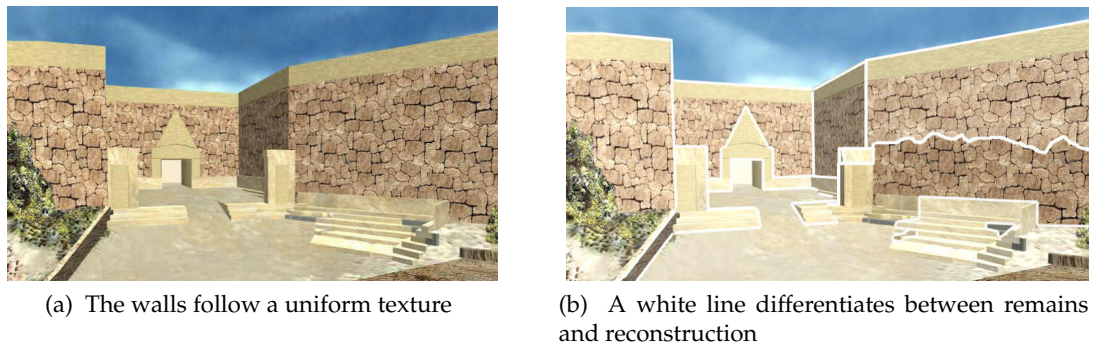
Figure 2.1: The Acropolis Propylaion reconstruction [96]

H. Eiteljorg ([95],[96]), was amongst the first researchers who demonstrated concern on the level of detail of heritage virtual worlds. He suggests that photorealistic visualisations impose their descriptions to the viewer and pass as a certain view of the past. Together with G. Tressel they stress the need to build accurate models which distinguish between the real and the hypothetical. His work on visualising parts of the Acropolis Propylaion demonstrates this concept, where renderings show the difference between the actual remains and the virtual parts (as shown in Figure 2.1).

Roussou [97] expands on the differences between PR and Non - Photorealistic Rendering (NPR) for archaeological reconstructions. It is suggested that both approaches if used sensibly can greatly aid heritage visualisation. The choice of visualisation techniques usually depends on the targeted audience: archaeologists and researchers or the general public. The former usually require a set of tools to evaluate hypotheses and aid interpretation, and as such the use of NPR techniques perhaps would aid more than PR ones. In the latter case, it is suggested that photorealism is a necessary part to transmit the information into engaging presentations destined for the public. As part of the ARCHEOS European project, an interactive VR setting from the Argos Agora has been developed. Specifically, the Tholos monument is displayed through which the users are able to alter between two different hypotheses. Additionally, the users are able to interactively manipulate archaeological fragments found around the area and snap them in their appropriate positions, as illustrated in Figure 2.2.

Kensek [98] presents a series of thoughts on ambiguity and uncertainty in 3D reconstructions. She stresses the importance of distinguishing between the real and the hypothetical and proposes a number of ways to achieve this through visualisa-

Figure 2.2: The Tholos monument: At the top two images, the remains on the ground differ from the reconstructed columns. At the bottom the two different hypotheses available to the user are shown [97]

tion. Drawing from non-digital techniques she suggests the use of different colours, transparency and rendering as various ways of representing ambiguity, as shown in Figure 2.3. Additionally, a project is demonstrated where the user is able to construct and evaluate different versions of Egyptian columns by combining various shafts, bases and capitals. The system shows both the certainty of the construction as well as prevents the user from making wrong combinations.

In [99] the reconstructed portions are displayed as sketch-lines and transparent overlays. Methods to present uncertainty are introduced in [98] where the use of colours, transparency and rendering are suggested as ways of representing ambiguity.

Another approach is given in [100] which deals with temporal uncertainty in archaeology. The research compares a variety of information visualisation schemes and uses visual cues (such as transparency and wireframe) to separate uncertain from recovered areas.
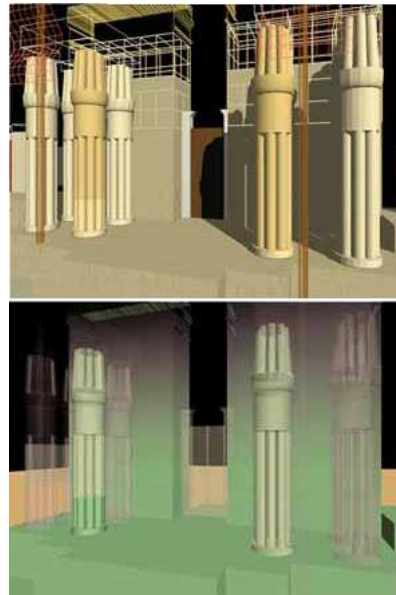
Figure 2.3: Four representations of uncertainty: Colour: red indicates less confidence while green more; Transparency: more transparent is less confident; Hybrid: colour and transparency; Rendering type: wireframe is less confident, shaded more).

In [101] fuzzy logic is introduced as a basis for quantifying reliability in virtual reconstructions. A reliability number is assigned, ranging from 0.0 (low reliability) to 1.0 (highest reliability), for each reconstructed part of the structure. By combining the values using a fuzzy logic union, an overall reliability index of the reconstruction is calculated. This approach is further analysed in Sections 6.2.1 and 6.2.2.

The majority of those approaches clearly distinguish between parts of the building that the expert knows about, and parts that they are partially (or fully) ignorant. However, there has been no attempt to explore, or formalise, the reasoning process which results in that uncertainty visualisation. For example, why would a column be 50% transparent? Why, in Figure 2.3, the lower end of the column is greener than its upper? How are those numbers decided? What evidence supports this visualisation?

This thesis suggests that the area of archaeological 3D reconstructions does not have sufficient research on which to build upon with regards to uncertainty. As a result, a top-down approach is taken to examine uncertainty, represented in Figure 2.4.

The concepts of ignorance and uncertainty are analysed and a number of definitions are presented. This is followed by different quantification and visualisation
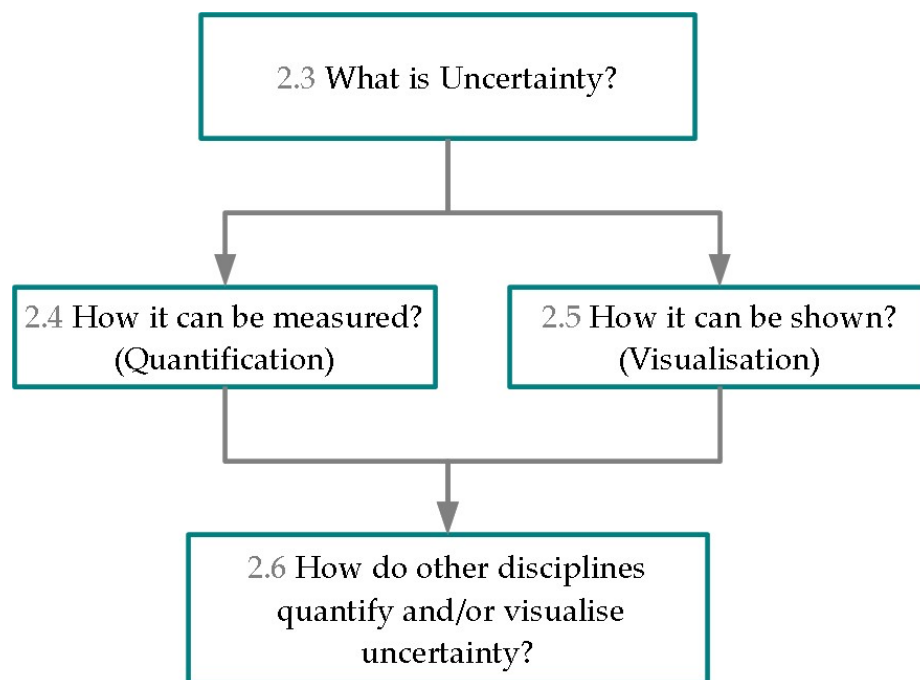
Figure 2.4: Top-down approach for analysis of uncertainty. Numbers refer to sections in the thesis.

methods. Lastly, other research fields are also examined. These fields use inferential reasoning and incomplete evidence to formally express beliefs in order to reach judgement or evaluate a hypothesis. Examples are spatial modelling, law, medicine, and geology. Some to a greater, and others to a lesser, extent they have explored the notions of mathematically and visually expressing beliefs and judgement under uncertainty. Their approaches and the mathematical models used are also introduced.

2.3 What is uncertainty?

As archaeologists have warned about the possible misuses of 3D reconstructions, artificial intelligence theorists and researchers in automated reasoning have warned about the misuse of reasoning/uncertainty representation models [102, 103, 104]. Not all models are appropriate for all cases.

In order to express a form of ignorance about a proposition, a hypothesis, it is useful to understand the category of ignorance that is dealt with. Unfortunately, a fixed definition of uncertainty does not exist. Numerous definitions do exist amongst the research literature, some less generic than others. This section presents a num-

ber of definitions and sub-categorisations of uncertainty along with examples where appropriate.

Bonissone and Tong [105] identify three general categories for expressing ignorance.

1. Incompleteness: refers to cases where the value of a variable is missing

2. Imprecision: refers to situations where the value of a variable is given, but not with the required precision

3. Uncertainty: covers cases where an agent can construct a personal subjective opinion on a proposition that is not definitely established for him [102].

The following archaeological examples illustrate the differences between the three forms suggested by Bonissone:

Incompleteness: Consider a database of Roman coin measurements. Information includes the coin's diameter and weight. Suppose for one coin that the information for the diameter is available (e.g. 18.66mm) but not the weight. In this case, the information is incomplete in terms of weight but precise and certain in terms of the diameter.

Imprecision: Consider the value of the weight of the aforementioned coin to be 2.9 g.–3.8 g. In this case, the information is complete in terms of data available for both variables, but it is imprecise, since it cannot be said without ambiguity which is the actual weight of the coin.

Uncertainty: Suppose that entries of measurements were done by an unskilled and unknowledgeable agent. Data for the aforementioned coin is 18.66 mm. for diameter and 2.9g. for weight. However, even though the data is precise and complete, it may be wrong.

A major difference between these forms of ignorance is the participation of an agent and, in consequence, whether an objective or subjective component is involved. Incompleteness and imprecision form *objective* expressions of ignorance. They are solely related to the data, the context, forming properties of it, and exist independently of an observer. On the other hand, uncertainty, forms a *subjective* type of ignorance, since the observer is taken under consideration. Uncertainty appears when an observer is not certain about the information available, and as a result, reflects partial knowledge, or belief.

Smets [106] suggests that *imperfection* permeates all information and data. It is quite idealistic to presume that perfect, complete information will exist for a system/hypothesis/suggestion. As a result, this imperfection must be taken into account when modelling a system.

He identifies three aspects of imperfect data that can co-exist:

1. Imprecise

2. Inconsistent

3. Uncertain

Imprecision and inconsistency relate directly to the information, the data. Imprecision can indicate that more than one world may be compatible with the given information. For example consider the statement:*"The Olympian gods were at most twelve"*. This gives as possible answers any number between 1 and 12 and thus the information is imprecise. Imprecision can also manifest in erroneous data, as for example in: *"The Olympian gods were 13"*. Compared to the model suggested by Bonissone, imprecision includes incompleteness. Imprecision in data is the biggest cause of uncertainty in an agent.

Inconsistency, on the other hand, suggests that none of the worlds currently known are compatible. It usually results from conflicting information; for example: *"My flight arrives from London at 20:00."* and *"There are no flights arriving from London after 15:00"*.

Uncertainty is also defined as objective (aleatory) and subjective (epistemic). Objective uncertainty relates to randomness, variability, the chance that an event is likely to occur, for example during a repeatable experiment. Subjective uncertainty is directly related to an agent's opinion about the reliability of the data depending on said agent's state of knowledge. If the agent has all the information required, there is no subjective uncertainty involved (but there may well be objective!). Objective uncertainty is well represented by statistical, frequentist, probabilistic methods. For the case of this thesis, objective uncertainty is not relevant and has been excluded.

Klir [107] also relates uncertainty to information deficiency. Different information deficiencies (i.e. incompleteness, vagueness,imprecision) will result to different uncertainty types. He identifies:

1. Nonspecificity: There are a variety of choices, the data may be imprecise.

2. Fuzziness: The data is vague with no clear boundaries.

3. Discord: There is conflict among choosing between alternatives. The data may have discrepancies.

Klir's categorisation is analogous to the imprecise and inconsistent data of Smets. The *imprecise erroneous category* is similar to the *imprecision* of Klir, while the *non-erroneous* is akin to the *vague & fuzzy* one.

   Gershon [108] follows a similar approach to Klir. He relates uncertainty to deficiency, or imperfection, in information. The information could be incomplete, too complex, corrupt, or inconsistent.

1. Corrupt: Analogous to errors in physics and engineering.

2. Incomplete: Data is missing.

3. Inconsistent:Different types of data are conflicting with each other.

4. Incomprehensible: Difficult to understand, confusing, complicated.

5. Uncertain: Data exists, but user/expert is unsure about the quality/existence.

In 1995, the ISO standards board released the *Guide to the Expression of Uncertainty in Measurement* [109]. Solely focusing on measurements of well-defined quantities, it provided guidelines on how to measure and report uncertainty in results. Two types of uncertainty are identified:'

1. Random effect uncertainty: due to random erroneous results

2. Systematic effect uncertainty: continuous error affecting the measuring process.

Pang [110] divides uncertainty in three categories:

1. Statistical: Due to some random process. Covers probabilistic and confidence methods.

2. Error: Covers the differences between an estimate an an actual value. Imprecision.

3. Range: Covers an interval of possible values.

Longley [111, 128] defines uncertainty as an umbrella term which depicts the difference between the perceived, incomplete, state of the world (by an agent) and the actual state of the world. Increasing uncertainty lowers the quality of a representative system (Langley's specialisation is in GIS). Uncertainty manifests in different ways:

1. Error

2. Inacurracy

3. Ambiguity

4. Vagueness

As it can be seen, there is no single definition of uncertainty. Research areas, such as GIS, offer their own interpretations on uncertainty based on works from Klir, Pang, Bonissone and others, depending on what fits their field. Section 2.4 introduces mathematical models of quantifying uncertainty while Section 2.6 examines approaches to uncertainty from different disciplines.

2.4   Quantifying uncertainty

Until well into the late 19th century, scientific uncertainty was completely undesirable. Proper scientific endeavours were considered unscientific if permeated with uncertainty [112]. The development of statistical methods and its acceptance as a valid scientific tool, had two major consequences. Firstly, uncertainty was starting to be something quite normal and expected for scientific enquiry, and no longer negative. Secondly, because of the direct relationship between probability and statistics, probability was the only way of dealing with uncertainty. Probability has had more than 300 years of extensive research through contributions from scientists such as Bernoulli, Bayes, Gauss, Laplace, Poisson, De Finetti, Kolmogorov and Savage.

The monopoly of probability as a mathematical method of representing aleatory and epistemic uncertainty, lasted until the second half of the 20th century, when non-probabilistic methods begun to emerge. These methods identified a number of uncertainty paradigms where probability was not able to give sufficient answers[1]. The last 50 years have seen the emergence of a number of non-probabilistic theories. Three of those are the most established and with the most amount of research literature behind them, given the short amount of time involved. The probabilistic model of subjective Bayesian probability is also included for completeness.

---

[1]The major difference between bi-valent and multi-valent logic is extensively analysed in Chapter 4.2 where also the mathematical foundations for the theories are examined.

2.4.1   Static, dynamic, decision

What all of the following theories have in common is three components: static, dynamic, and decision.

- Static: pool of rules, previous knowledge, anything that can be considered as background information to the current problem. There could be cases where this is empty, reflecting total ignorance.

- Dynamic: New information arrives, in the form of evidence, data, etc., that updates the knowledge.

- Decision: Given all accumulated knowledge, this aids in selecting one of the possible hypotheses.

The differences arise in how each of those components is implemented.

2.4.2   Subjective Probability (1763)

Based on Bayes' Rule [113], this form of probability is equated to a quantification of subjective belief. The belief is calculated by combining what one already knows about the data, and the support (or non-support) of the available evidence.

Bayes' Rule works fine with singular variables, in other words it is capable of assigning belief on one hypothesis but not a combination of two hypotheses. For example, consider the following statement: *The column was most likely Ionic, I would value my certainty at 90%.* This can be represented with subjective probability. Now consider the following: *The column was most likely Ionic or Corinthian, I am 90% sure of that.* This can not be represented in a holistic way; the rules of probability declare that between those two the probability must be shared. The equivalent would be: *I am sure 45% that it was Ionic and 45% that it was Corinthian.* A second requirement of Bayes' Rule is a priori information, knowledge available to the expert or system that will be used as a basis for updating.

2.4.3   Evidence Theory (1976)

Evidence theory is best recognised through the Dempster-Shafer Theory of Evidence (DST). Shafer [114], who built his theory on Dempster's research [115], suggested the use of degrees of belief to represent uncertainty in a quantified manner. One of the main features of evidence theory is that a belief can be assigned to a set of

hypotheses rather than a singular one. A second feature is that it does not require *a priori* knowledge and it is more flexible in scenarios of total ignorance. Lastly, DST can contain Bayes Rule as a special case [116].

A number of different probabilistic formalisms arose that tackled some of the shortcomings of DST, such as the hints formalism by Kohlas [117, 118], and random sets [119]. The non-probabilistic ones include Shafer's original work on belief functions and Smet's [120] Transferable Belief Model. For a further analysis of the different formalisms the reader is referred to [121].

2.4.4   Fuzzy Sets & Possibility Theory (1965, 1978)

In 1965 Zadeh [122] introduced a theory of sets where the boundaries are not crisp, but fuzzy. An object could belong to more than one sets, according to its degree. This allowed for the representation of vague terms, whereas before it was impossible to do in an intuitive way. Consider the expression *"It is sunny outside."* If the amount of clouds covering the view is less than a certain percentage, then it is sunny. How much is less? Is it, 30%? If the clouds cover 31%, is it immediately cloudy? Fuzzy set theory introduced this needed vagueness by having a gradual transition between terms such as sunny, cloudy, little, a lot, etc.

Possibility theory [123] is a theory for uncertainty quantification based on fuzzy sets. It is non-probabilistic as it does not follow certain rules of probability. It does not require *a priori* knowledge.

2.4.5   Expressing uncertainty

The mathematical models that quantify uncertainty and used in decision theory all rely on a normative framework. They describe how a rational agent *should* act.

For a long time, humans expressing confidence judgements when updating their beliefs, were perceived to follow the rules and structure of subjective probability theory. However, it has been shown that in various situations, humans did not behave as rationally as probability theory would predict [124]. The *conjunction fallacy* is often commited by humans; people often assign a probability of conjunction to their reasoning that is, in some cases, greater than one or even both of the two premises.

Research results indicate that when people are asked to express their confidence, they do not actually produce a probability [125]. Under-confidence is often expressed when issues to be assessed are easy, and over-confidence is expressed when issues to be assessed are difficult. Research by Griffin and Tversky [126] explain over-

and under-confidence on the hypothesis that people rely heavily on the strength of the evidence with little regard to its weight. Example of overconfidence arise also in the field of medicine [127], where clinicians evaluating a positive mammogram result gave a probability of patient sickness at 70% where the calculated Bayesian probability was 8%. Further research [128] confirm findings by Griffin and Tversky and identify patterns where weight, quality and strength of evidence can be over- or under-weighted.

Researchers have also criticised the application of the probabilistic framework when single events are assessed [129], e.g. the plausibility that a diagnosis for a patient is correct. They showcase that the probabilistic framework should be applied to the understanding of confidence assessments expressed over a long period of time and for multiple events. Alternative normative frameworks such as possibilistic measures have not been applied to uncertain situations except in a few cases of empirical evaluation of human radiological diagnosis [130].

The theories are examined more analytically in Chapters 4 and 5. The next section introduces methods for visualising uncertainty.

2.5   Visualising uncertainty

Information visualisation techniques [131]encompass a wide range of approaches developed to help people visually interpret data. However, when uncertainty is involved, particularly in scientific data, there is a need to visualise that precise lack of information. The significance of visualising the uncertainty in a body of data has been acknowledged in different fields such as flow data [132], positional uncertainty in molecular structure [133], astrophysical data [134], terrain modelling [135] and anisotropic rock property models [136]. Reznik and Pham [137, 138] have evaluated the information uncertainty of fuzzy models and its visualization.

A systematic classification for uncertainty visualisation techniques has been surveyed by Pang [110], classifying the techniques into adding glyphs and modifying geometry, attributes, animation and sound. Pang suggests using different visual cues according to the type of data, such as scalar, vector or range. Methods that can be directly applied to a 3D model are presented [139]:

Free graphical variables: Affects the hue, texture, lighting, shading, crispness and opacity/transparency values of a visualised object. Also includes pseudocolouring, a technique for representing scalar, interval, or ordered data values by using a sequence

of colours. A well-known example would be geographical maps.

Side by side: compares two, or more, different visualisations of data i.e. simulated and actual.

Animation: Helpful for representing changes in uncertainty dynamically this might include blinking areas, and motion blur.

Integration of objects: This approach represents uncertainty with externally attached objects. These can include glyphs, error bars, text labels.

Research on the visualisation of uncertainty has been slow to catch up, being occasional and sporadic compared to the visualisation of information. The majority of the research deals with the uncertainty in a body of crisp, measurable data. Much less research has been conducted in visualisation of uncertainty involved in reasoning, thought process, and subjective belief about information [140]. However, researchers highlight the need [141] for the creation of formal error and visualization frameworks and also formally evaluate through user studies simulated and experimental 2D/3D data with uncertainty information.

2.6 Uncertainty in other fields

2.6.1 GIS

The GIS community was one of the first disciplines to acknowledge the existence of uncertainty and take efforts to visualise it. Veregin [142] identified 5 goals for managing and reducing uncertainty within a GIS system. He stressed that sources and cause of uncertainty should be understood first in order to manage its occurrence or propagation through a GIS system..

Fisher [143], by expanding on Klir's definition and categorisation of uncertainty, provides a taxonomy of GIS uncertainty as the result of error, ambiguity, vagueness or lack of information of data. In further research he [144] suggested that a large part of geographical phenomena are vague in nature and a formal recognition of vagueness should be included in geographical systems. He proposes that a way to represent vague concepts would be through fuzzy sets.

In [145] the authors use Evidence Theory to integrate subjective beliefs and spatial geographical information regarding land suitability analysis. Ferrier [146] uses a Knowledge-Based System (KBS) utilising a comparison of approximate reasoning techniques such as DST, Bayesian probability and possibility theory. The KBS is used together with a GIS to analyse sedimentary basins. When eliciting the experts' know-

ledge for use in the KBS, it became apparent that two factors had to be addressed: the vagueness and imprecision of data and the reliability of a hypothesis.

Couclelis [147] suggests that representation of uncertainty within a GIS should move away from a strict spatial data perspective. She argues that the limitation in the experts' knowledge should also be acknowledged. There are cases where the actual truth may be unknown (i.e. due to incomplete information). She proceeds by analysing how approaches from the AI community, logic, philosophy and mathematical quantifications of epistemic uncertainty (such as fuzziness, belief or Bayesian estimate) can help the GIS community to formally map out the uncertainty in spatial knowledge.

In specific archaeological applications of GIS, there have been some extensive discussions on the uncertainty involved in regards to the accuracy of DEM [148, 149]. Wheatley [150], influenced by Fisher's uncertainty work on fuzzy viewshed analysis [151], suggested that uncertainty be acknowledged in cumulative viewshed analysis and proposed that viewshed and visibility analysis should be enriched with including their uncertainty. In the area of predictive modelling, Canning [152] has used DST to represent the expert knowledge and a body of data for predicting Aboriginal sites at Victoria, Australia. Comber [153] uses DST and Bayesian analysis to predict the existence of bog habitats.

There have been a number of uncertainty visualisations in spatial applications. In [135] animations are used to represent different exploratory data visualizations. Jiang has shown how paleness, or whiteness, gives the perception of carrying more uncertainty than a more solid colour and applies it on geographic fuzzy spatial datasets. Hengl [154] applies this colour-model scheme based on Hue - Saturation - Intensity (HSI) to represent spatial data uncertainty in soil thickness and a fuzzy classification of landforms. For a further in-depth analysis of GIS visualization of uncertainty the reader is refereed to [155].

2.6.2 Law

Lawyers must continually assess the weight and strength of evidence as well as be able to combine various types of evidence in order to present a strong case to the court. On the other hand, the judge and jury have to evaluate this evidence and lend support to the guilt or innocence of the accused. They also have to accept or discard evidence according to its weight. There is a large element of ignorance and ambiguity involved in law cases; evidence might be incomplete, inconclusive, conflicting,

witnesses might be unreliable, forensic tests might have a percentage of error.

The need to assign credibility and weight has been recognised by the discipline for at least two centuries. Most commonly referred to as Evidence Law, it is quite well represented in the following quote by Sir G. Gilbert one of the founders of modern law:

> "There are several degrees from perfect Certainty and Demonstration quite down to Improbability and Unlikeness...and there are several Acts of Mind proportioned to those Degrees of Evidence...from full Assurance and Confidence, quite down to Conjecture, Doubt, Distrust and Disbelief.
>
> Now what to be done...is to range all Matters in the Scale of Probability so as to lay Weight where the Cause ought to be prepondate."[156]

There are no ground rules of weight, covered by law, for grading evidence, although it has been proposed [157]. However, there are certain categories of evidence, like eyewitness accounts, where the judge is obliged to warn the jury about potential unreliability [158]. In some legislation areas, such as Scotland, at least two witnesses are required to bring a person to court.

One of the greatest contributors of Evidence Law was J. Wigmore. Apart from major influences in Anglo-American jurisdictions, he developed the Wigmore Chart [159]. The chart aimed to represent the justification of conclusion of facts made by the jury/judge in a trial. It was a graphical method to represent the reasoning process, evidence involved as well as its strength and direction.

The indisputable existence of uncertainty was there, and it was not long before that probabilistic and non-probabilistic methods were evaluated as a means of representing belief about evidence. Since the early 80s, books have been devoted to the use of probability [160, 161, 162] in the legal process. Bayes' rule and its application to evaluating evidence and reaching decisions has been explored by Dawid [163, 164, 165]. Schum [166, 167] has explored the relationships between knowledge, probability and credibility and developed MACE (Marshalling Competence and Credibility Evidence). MACE uses subjective Bayesian Probability and a chain of questions about the witness that assists the law enforcer in evaluating the witness. While the majority of research has focused on probabilistic representations, in more recent years there has been an increase in the use of evidence theory and fuzzy systems. Dragoni et al. [168] use the Dempster-Shafer Theory to represent belief revision in deliberations of jury, including the issue of credibility. Jösang [169] uses DST belief model as an example of subjective logic involved in the courtroom. Yablon [170] suggests that

fuzzy sets could be used to determine the concept of fairness in a litigation. Chavkin suggests [171] that fuzzy sets could be used in law to represent vague concepts.

There has been limited actual practical application of probabilistic and non-probabilistic inference in court decision making. In an infamous (in lawyer circles) case of Regina v. Adams [172] the jury was instructed to use Bayesian reasoning to combine certain evidence. The ruling was so critical of the use of Bayes' theorem in the courtroom that it has not been used since. Despite the mistrust by law courts, researchers in evidence law continued advocating the need for a framework to represent evidence and belief in legal contexts. Schum, Twining, Dawid and others [173, 158, 174] have suggested that evidence and knowledge representation is an important multi-disciplinary subject and should be researched further.

2.6.3   Medicine

Health practitioners are also experts who have to face uncertainty on a daily basis by evaluating evidence (symptoms) in order to make a diagnosis. Similar to law and archaeology, evidence might be inconclusive, vague, conflicting. The knowledge available is rarely comprehensive or complete. Additionally, the doctor's expertise plays an important role in evaluating evidence and making decisions.

Uncertainty in medicine has been long recognised and accepted by its practitioners. Medical students are trained [175] in identifying and handling the existence of uncertainty. Fox, and later Light, [176, 175] identify three elements of medical uncertainty:

1. Incomplete mastery of knowledge (the knowledge of the professional.

2. Limitations of current literature (the knowledge of the profession).

3. Difficulties in distinguishing between the above.

Soon after, an influential paper by Ledley and Lusted suggested that medical reasoning could be quantified by using Bayesian probability [177, 178]. Developments in computer science and AI led to research in Clinical Decision Support Systems (CDSS). The role of these systems is to assist practitioners in reaching a diagnosis, in preventive care, and treatment [179].

The majority of the systems use a pool of knowledge that assists in decisions [180]. Medical experts contribute to this pool by adding their knowledge about symptoms, diseases, and results from clinical trials. Depending on the CDSS used, the knowledge

is expressed by rules (MYCIN [181]), subjective probabilities (de Dombal [182, 183], Internist-1/QMR [184]), fuzzy sets/logic (CLINAID [185]) and evidence theory [186].

2.7   Conclusion

This Chapter has examined the concept of uncertainty; how it is handled in different fields, its different definitions and visualisation approaches. It is quite clear that cultural informatics, and archaeological 3D reconstructions involving visualisation of uncertainty have a long, long, way to go compared to other disciplines. This thesis suggests that 3D reconstructions could be so much more than *pretty pictures*. They have the potential to be what GIS has become for geography, archaeology, and medicine: a way to represent and visualise the extent of knowledge. Examination of other disciplines has shown that they use a plethora of mathematical models, all within the three categories identified in Section 2.4, namely Bayesian subjective probability, Fuzzy Sets and Evidence Theory.

For the purposes of this thesis, three mathematical models of representing uncertainty, each belonging to one of the different categories, are evaluated and applied in the area of cultural informatics and archaeological 3D reconstruction. Bayesian Subjective Probability is included, firstly for being the oldest and most researched model and secondly for having been used extensively throughout other disciplines for the purposes of quantifying uncertainty. From the non-probabilistic category of Fuzzy Sets, Possibility Theory is also included, as an established alternative mathematical model of representing subjective uncertainty. It is a relatively new area, with active research, but fewer applied cases than Bayesian Probability. Its use of two measurements (Neccessity and Possibility) offers a range of capabilities not supported by Subjective Probability.

Bayesian Subjective Probability and Possibility Theory are two of the most popural models in their respective categories (probabilistic and non-probabilistic/fuzzy). However, the most known representative of Evidence Theory, the Dempster-Shafer Theory, has not been applied in this thesis. Rather, a non-probabilistic interpretation of it, the Transferable Belief Model [187], is evaluated and applied. There are three main reasons to choosing a less popular model:

Firstly, the implicit rejection of Probability Theory (on which DST is based) allows the TBM to represent conflict. Conflict can arise between two (or more) contrasting pieces of evidence. The capability to quantify conflict would be of great use in a field such as archaeology where often evidence may be encountered that is ambiguous,

conflicting, or untrustworthy.

Secondly, the TBM allows for the mathematical representation of an 'open world'. In other words, while evaluations of evidence are being made, the expert is allowed to enter new hypotheses in the list of possible interpretations. This feature is not supported by either Subjective Probability or Possibility Theory.

Thirdly, the TBM is a very new model, developed in the beginning of the 1990s with limited research and almost no applied case studies. Its evaluation and application in this thesis is the first application of the model in the context of Cultural Informatics.

Before those models can be applied and evaluated, uncertainty must be identified and defined in its archaeological context. The following Chapter identifies uncertainty in a cultural informatics context, and presents research on archaeologists' attitude towards uncertainty and evidence.

> Knowledge is of two kinds. We
> know a subject ourselves, or we
> know where we can find
> information on it."

<div align="right">

Samuel Johnson

</div>

CHAPTER **3**

Identifying Uncertainty

3.1 Introduction

Chapter 2 evaluated how other disciplines approach uncertainty, presented different definitions and classifications and higlighted the lack of similar approaches in the area of cultural informatics. Cultural heritage uncertainty representation is a novel research area and this thesis explores foundation questions. Figure 3.1 illustrates the case. Stage A is where the thesis makes contributions in, while Stages B and C are envisioned future interest areas within archaeological uncertainty research.

In order to represent archaeological uncertainty in 3D reconstructions, the first step is identifcation: Firstly, does it exist? To what end? Is uncertainty something important, or not? How does it affect archaeologists? How can it be defined, in relation to reconstructions?

Research conducted as part of this thesis involved the creation and distribution of a questionnaire focused on expert archaeologists. The questionnaire explored the perception of uncertainty among expert archaeologists, queried on different types of evidence that might influence uncertainty and examined possible differences between different archaeological groups. This Chapter presents the questionnaire design, development, and research results. Finally, by discussing research results and evaluating previous work on uncertainty it gives a definition of archaeological uncertainty in 3D reconstructions.
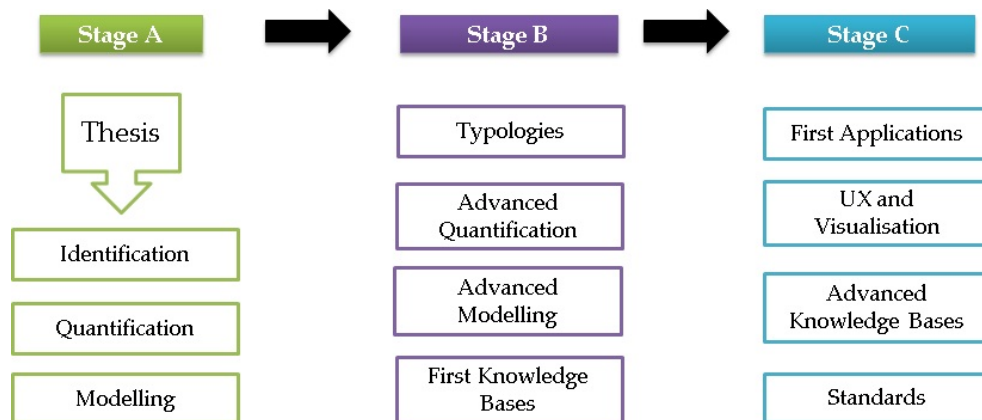
Figure 3.1: Stages of uncertainty research

3.2 Questionnaire on Archaeological Uncertainty and Evidence

There is no set list of evidence types that turns up in every excavation. Additionally, every historical period may show an abundance of a specific type of evidence but a complete lack of another. Of particular interest to this thesis are types of evidence that influence the interpretation of structures.

In order to identify evidence types, the first step was to conduct open structured interviews with expert archaeologists. Discussions were done on an individual basis and the panel consisted of five archaeologists. The discussion focused solely on the interpretation of structures. Three of the archaeologists were Romano-British specialists, one focused on Eastern Europe and Byzantium, and the last on Anglo-Saxon England. All of them have in excess of 10 years experience in their fields. Three were interviewed personally, one by phone, and one by email. From these discussions a list of evidence factors that can influence uncertainty was created. The factors identified are:

1. Features: all elements of man-made structures; these can range from ditches to wall remains, to post-holes of wooden structures.

2. Artefacts: any object made, affected, used, or modified in some way by human beings; this usually includes pottery, glass, lithic, etc.

3. Biofacts: Biofacts (or ecofacts) constitute of human, animal and plant remains which are not changed by human interaction.

4. Textual evidence: ancient texts and documents which may provide information on architecture, decoration, lifestyle etc. of a specific culture.

5. Absolute comparisons: a structure is compared to a structure of similar proportions.

6. Contextual comparisons: a structure and its context is compared to similar structures with similar contexts.

7. Topography: natural features of the landscape in which the building is located; elevation of the area, characteristics of the region and landform data in general.

8. Peer review: the interviews indicated that interpretation is heavily based on discussions and re-discussions with other archaeologists and architecture specialists were considered important.

The next step involved the creation of a questionnaire with an aim to gather more information on how archaeologists perceive uncertainty as well as the evidence.. The aim of the questionnaire was to determine if:

- There is any difference between the results obtained by archaeologists with different expertise. The initial hypothesis was that a difference should be evident.

- There is any perceived preference among these different factors.

3.2.1   Design

There is an abundant bibliography on the design of questionnaires for factors involving perception, immersion [188], and computer interfaces evaluation [189]. However, in the case of this thesis, there is no similar work conducted relating to questions and factors on archaeological uncertainty. As a result, the questions were designed progressively with consultation by the archaeologist panel described above and by consultation of questionnaire design textbooks [190, 191]. The questionnaire was designed in three steps: Questionnaire type selection, response scale selection, and question wording.

Questionnaire type

There are a number of questionnaire types suitable for human factors testing and evaluation. These include multiple choice, rating scales, semantic differential and

open ended questions and are illustrated in Table 3.1. The most commonly used is a rating scale. The chosen questionnaire type should be based on the way the data would be eventually analysed as well as the appropriateness of the area in question. For example, do multiple choice answers reflect the area or not.

Usually, for exploratory purposes, open ended questions or structured interviews are used. Through these questions, the participants offer written responses with no restrictions. This allows the researcher to retrieve opinions on a topic and is a useful precursor for data gathering for descriptive or exploratory purposes. An issue with open ended questions is that they are difficult to quantify and are prone to subjectivity on the researcher's part; she must decide whether an open-ended answer signals an agreement or disagreement with a topic.

For descriptive purposes, multiple choice questions and rating scales are mostly used. Since those can be quantified at the interval scale, they provide access to a wider range of statistical options than the open-ended nominal ones.

As mentioned in Section 3.2, structured interviews/open ended questions were used with the archaeological panel to determine the factor list. For the purposes of the questionnaire, since it focuses on attitude measurement towards different factors, a multiple choice option was considered to be restrictive. As a result, a rating scale was used through most of the questions. Multiple choice options appear in the personal information pages.

Response scale

The response scale simply represents the distribution of the responses by providing firstly the number and secondly the type of allowable answers to a question. There are various issues to consider, such as the balance, the polarity and the number of values. Like the example in Table 3.1, a scale that has an equal number of positive and negative alternatives is known as balanced. Such balanced scales are commonly chosen because they tend to produce nearly normal distributions. On the contrary, unbalanced scales are used when it is expected that responses will be selected from the scale's extremes.

Another point to consider is the presence or absence of a scale midpoint. Denying the ability of neutrality will probably increase the variability towards one extreme or the other. Additionally, it is sometimes considered impolite to bereft respondents of the choice to be neutral and force them to select a choice.

Lastly, the number of response alternatives tends to be based on the degree of

Questionnaire types	Rating scale	Semantic differential	Multiple choice	Open ended
Description	Used for most evaluation scenarios	Used for measuring values, attitudes or complex relationships. This is represented by simultaneous ratings across multiple dimensions	Used for collection of demographic data / screening respondents.	When there are too many possible responses to be listed.
Advantages	Structured, specific questions. Can be quickly answered. Reliable data, can be used statistically	Provides data on the similarity of various concepts	Answers easy to summarise	May obtain unanticipated answers. Easy to write questions
Disadvantages	Structure leaves little room for unanticipated answers. Questions should be well structured out	Limited use for testing and evaluation scenarios	Difficult to ask complex questions. May force the respondent to make a choice even if she does not want to	More time to complete, difficult to summarise and statistically analyse.
Examples	Rate your overall reaction to every-day commuting: strongly dislike | dislike | neutral | like | strongly like	Place an X in each of the lines to describe your commuting experience: Relaxing———Tiring Quick———Slow Tube on time———Late	How many times do you commute to work / week? _1 _2 _3 _4 _more than 4	Please write: What do you like most about commuting? What do you dislike most about commuting?

Table 3.1: Questionnaire types (afer [191, :229])

Ice cream is good for breakfast

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree

Table 3.2: Typical Likert scale statement and possible answers

discrimination required. For example, by increasing the number of alternatives, it is possible to decrease the uncertain responses (neutral). However, this also increases the questionnaire administration time because the respondent has to examine all alternatives. Research suggests [192] that up to seven alternatives aids discrimination between the answers– more than that and increased variability and small discrimination is faced between answers. Most evaluation questionnaires utilise a balanced, bipolar scale with 5 to 7 points.

For the purposes of this questionnaire, in parts where a rating scale is used, a balanced, bipolar scale with seven points was chosen. The aim was to offer the most conveniently possible range of answers, in a balanced way and with the inclusion of a neutral standpoint–an issue that was considered extremely important by the archaeological advisory panel.

A Likert [193] likelihood scale is a rating scale used for attitude measurement. In place of a numerical scale, answers to a statement are given on a scale ranging from complete agreement on one side to complete disagreement on the other side. The original version of the scale uses five points of agreement or disagreement, as shown in Table3.2:

Sometimes a four-point scale is used; this is a forced choice method since the middle option of "Neither agree nor disagree" is not available. This is often considered as forcing the choice of the reader one way or another. Likert scales may be subject to distortion from several causes. Respondents may avoid using extreme response categories (central tendency bias); agree with statements as presented (acquiescence bias); or try to portray themselves or their organization in a more favourable light (social desirability bias).

Question creation and wording

The third step in constructing a questionnaire involves the wording of questions. Attention must be paid to:

1. Vocabulary: It must be clear and precise, avoiding jargon and confusing terms.

2. Negatives: Negative phrases such as "*Indicate how often you do not eat breakfast during a week*" may appear confusing and misunderstood by the reader–the negative word should usually be left out.

3. Double-barrelled questions: These are questions involving two questions at the same time i.e. *"Do you like watching cartoons or eating ice cream"*. The respondent may like cartoons but not ice-cream or vice versa and is thus confusing to answer one way or another.

4. Leading questions: These presuppose an event e.g. The statement *"Indicate the sourness of the ice-cream"* presumes that the ice cream is *de facto* sour.

3.2.2 Prototype

A total of 35 related questions were created. The archaeologists were asked to carefully read through questions and make suggestions on their validity, appropriateness, easiness of understanding, and order, as well as the types of answers available. The resulting questionnaire can be found in Appendix A.

The questionnaire is divided in nine parts. The first part deals with perceptions of uncertainty in the discipline and how alternative hypotheses are handled. Parts two to six contain questions on the different identified factors. Part seven provides combinations of factors (for example features and artefacts, or artefacts and biofacts) and queries the expert as to how strong she considers this combination of evidence to be. In the eighth part, the list of factors is provided and the expert is requested to assign an order of importance for each. Finally, in the last part, the expert enters his/her personal data such as their specialisation field, time spent on excavations etc. Most importantly she is asked to give an opinion on the completeness of the factor list.

The answers to parts 1—6 are chosen from a seven point ordered response Likert scale. While prototyping the questionnaire, the first decision was to use a likelihood Likert scale, which would indicate the agreement or disagreement with a statement. However, feedback received from archaeologists showed that a frequency scale was better understood in accordance to the specific questions. As a result, a frequency scale was used.

Values include never, almost never, sometimes, often, very often, and always. The expert is asked as to how often she encounters each statement when making interpretations and reconstructions of a structure. Table 3.3 illustrates the concept:

I consider two or more valid interpretations.						
Never	Almost never	Seldom	Sometimes	Often	Very often	Always

Table 3.3: Question example

The participants were aware under which group each question belonged since each part is clearly marked out as belonging to uncertainty, or a specific factor.

In part 7 the expert is asked to rate a list of factor combinations from 1 (weak) to 9 (strong). Lastly, part 8 follows a ranking system where the expert must rank the eight factors according to how important she believes they are in an interpretation. This is done by using numbers from 1 to 8 only once and assigning them to the factors.

Once the questions and scaling were completed, it was again tested by five archaeologists and five non-archaeologists for its coherence, layout and structure. The questionnaire was distributed in two versions, printed and electronic.

Both versions, printed and electronic, had the same content and in the same order. The printed version was personally distributed to archaeologists and collected after completion. The electronic version was created with a simple design using XHTML and PHP technology, while also being assistive and interactive. Features include error handling, data retention and real-time data collection. Before the actual distribution of the electronic questionnaire to any experts, fifteen people were asked to evaluate it in terms of: understanding of the instructions, text length readability, font size, questionnaire flow and error handling. They were also asked to identify their computer expertise, browser choice and screen resolution. 73.3% were male and 26.7% female. 26.7% were less than 25 years old, 53.3% were between 25–33 and 20% were above 33. 33.3% considered themselves to be adequate or knowledgeable with IT while the rest were identified as experts or professionals. While 100% considered the length of text and instructions easy to read, 33.3% faced issues with the error handling and 20% with the default font size. About 73% of them were navigating with Mozilla Firefox and the rest with Internet Explorer. 60% had a resolution of 1024x768 and 40% had 1280x1024 or more.

Following their suggestions, more robust error handling was implemented, which introduced the ability to change the text size, gave instructions for all popular browsers, and ensured full conformance with W3C XHTML/CSS specifications.

Figure 3.2: Electronic questionnaire page

3.2.3   Participants

Participants were recruited across the archaeological discipline. Because the case study is based on a Romano-British structure, half of the participants were Roman archaeology experts. A requirement for inclusion was the participation in archaeological excavation and interpretation of structures. Twenty experts participated, 50% specialising in Roman archaeology and the other half in different fields. As far as experience is concerned, 50% of the archaeologists had been working for 10 years or more on their particular field while 30% had also been excavating for more than 10 years. 60% of the archaeologists involved were between 26-33 years old, 15% between 34-41 and the rest above 41. 50% of the participants were male and 50% female.

3.2.4   Procedure

Before answering the questionnaire, the participants had to read an information page that explained the format of the questionnaire and how to answer questions. In the case of the printed questionnaire, this was also verbally explained. In the electronic version this was ensured by having to specifically tick a box in order to accept that they fully understood the instructions. At the end of the questionnaire participants

answer questions about the questionnaire itself such as if they considered the factor list complete, or if the questionnaire was difficult to fill in.

3.2.5 Results

The data was examined in three parts. The first part was related to the Likert-style questions (parts 1—6), the second part with the evidence combinations (part 7) and the last one with the ordering (part 8).

In order to establish the appropriate tests for the data, it was tested for parametric conformance assumptions. These assumptions presume that data: is normally distributed; has same variance through it; is measurable at least at the interval scale; is independent (results from one participant are independent from another). While the last two statements hold true, the first two were tested by normality and homogeneity tests respectively. Kolmogorov-Smirnov analysis indicated (sig. <0.05) that the data is not normal, while Levene's test (sig. >0.05) indicated homogeneity between the Roman and non-Roman groups. The violation of one assumption leads us to use non-parametric tests. The parts will be now described more analytically.

Parts 1—6

Thirty-five questions belong to these parts. Usually, one-way ANOVA is used to check for differences between independent groups (such as the case here); however ANOVA requires parametric data. A non-parametric alternative to ANOVA is the Mann-Whitney approach. This test examines whether there are differences between the results of groups. The data was tested with exact significance values, an option which gives more accurate results in small samples like ours. The Kolmogorov-Smirnov-Z (KGZ) test was chosen, which behaves similarly to the Mann-Whitney test but is much more appropriate for smaller samples. It appears that answers were not different between the Roman and the non-Roman group (Sig>0.05).

The following step involved averaging the answers for each factor and examining the results. Terms that have a reverse phrasing were examined. For example, in the questions measuring the attitude towards uncertainty, the statement "*There is only one true interpretation of a site*" is a reverse-phrased one: if the attitude towards uncertainty is positive, the answer to this question should be negative. However, this would give a less weight in the scale, and vice versa. For this reason, six negative-phrased questions had their scores reversed before they were averaged. As a result, out of the 35 variables, 8 resulted (Table 3.4).

Resulting variable	#	Resulting variable	#
Uncertainty	7	Ancient texts	3
Features	7	Biofacts	3
Topography	3	Peer support	5
Artefacts	4		
Total	**35**		

Table 3.4: Aggregated variables

This data was again tested for parametric properties, and it was found to be homogeneous but non-parametric; Biofacts and Ancient texts failed the Kolmogorov-Smirnov parametric test (Table 3.5). Table 3.6 shows the results for the homogeneity test.

A KGZ test was conducted to test for any difference between the two groups. The results, demonstrated in Table 3.7 show that no difference was observed. The above results suggest that no difference can be observed between answers given from the Roman and non-Roman groups.

Looking at the means of the different factors (Table 3.8 and Figure 3.3) it can be seen that archaeologists primarily place importance on the suggestions of their peers (Peer support). Then, evidence from existing sites (Comparisons) and Features score quite strongly, followed by Artefacts and Topography, Biofacts and Ancient texts.

Part 7

Twenty-two options belong to this Part. Having established in the beginning of Section 3.2.5 that the data is non-parametric, a KGZ test was conducted to test for any differences between the two groups. The results, included in the appendix, show no significant difference between groups and allow us to examine the means as a whole group. Table 3.9 shows the results from the combinations of Part 7. It is interesting to note that combinations involving features, artefacts, contextual and absolute comparisons and topography score more strongly than other combinations. Additionally, combinations involving texts and biofacts score the least in the table.

Roman experience?		Kolmogorov-Smirnov(a)		
		Statistic	df	Sig.
Uncertainty	yes	.255	10	.064
	no	.242	10	.099
Features	yes	.174	10	.200(*)
	no	.200	10	.200(*)
Artefacts	yes	.202	10	.200(*)
	no	.158	10	.200(*)
Biofacts	yes	.285	10	.021
	no	.184	10	.200(*)
Topography	yes	.185	10	.200(*)
	no	.214	10	.200(*)
Comparisons	yes	.138	10	.200(*)
	no	.178	10	.200(*)
Ancient Texts	yes	.302	10	.010
	no	.303	10	.010
Peer Support	yes	.202	10	.200(*)
	no	.139	10	.200(*)

Table 3.5: Normality test for Parts 1—6

	Levene Statistic	Sig.
Uncertainty	.491	.492
Features	.658	.428
Artefacts	1.590	.223
Biofacts	.179	.677
Topography	1.097	.309
Comparisons	1.300	.269
Ancient Texts	2.353	.142
Peer Support	.275	.606

Table 3.6: Homogeneity test for Parts 1—6

	Kolmogorov Smirnov Z	Exact Sig. (2-tailed)	Most Extreme Diff.
Uncertainty	0.447	0.958	0.2
Features	0.671	0.708	0.3
Artefacts	0.894	0.332	0.4
Biofacts	0.671	0.526	0.3
Topography	0.224	1	0.1
Comparisons	0.894	0.181	0.4
Ancient Texts	0.894	0.288	0.4
Peer Support	0.447	0.981	0.2

Table 3.7: KGZ test for Parts 1—6

	Mean	Std. Deviation	Minimum	Maximum
Peer Support	5.25	0.75359	4	6.6
Comparisons	5	0.71737	3.67	6.67
Features	4.9625	0.47659	4.25	6
Uncertainty	4.8429	0.52006	4.29	6
Artefacts	4.7625	0.66627	3.75	6.25
Topography	4.15	0.73727	2.67	5.67
Biofacts	3.9	0.72628	2.67	5.33
Ancient Texts	3.6333	0.7327	2.33	5

Table 3.8: Means for Parts 1—6

	Sum	Mean	Std. Deviation	Variance
Features & artefacts.	155	7.75	1.02	1.039
Features & contextual comparison.	138	6.9	1.071	1.147
Features & peer input.	138	6.9	0.852	0.726
Features & absolute comparisons.	137	6.85	1.424	2.029
Features & topography.	131	6.55	1.638	2.682
Artefacts & contextual comparison.	130	6.5	1.821	3.316
Features & biofacts.	129	6.45	1.638	2.682
Artefacts & absolute comparisons.	128	6.4	1.635	2.674
Artefacts & biofacts.	127	6.35	2.033	4.134
Artefacts & peer input.	124	6.2	1.508	2.274
Topography & absolute comparisons.	124	6.2	2.093	4.379
Topography & contextual comparison.	121	6.05	1.791	3.208
Artefacts & topography.	121	6.05	2.038	4.155
Topography & peer input.	118	5.9	1.586	2.516
Features & textual evidence.	115	5.75	1.97	3.882
Topography & textual evidence.	111	5.55	2.481	6.155
Biofacts & peer input.	111	5.55	1.504	2.261
Artefacts & textual evidence.	111	5.55	2.114	4.471
Biofacts & contextual comparison.	109	5.45	2.328	5.418
Biofacts & absolute comparisons.	106	5.3	2.319	5.379
Biofacts & topography.	98	4.9	2.469	6.095
Biofacts & textual evidence.	98	4.9	1.997	3.989

Table 3.9: Means for Part 7
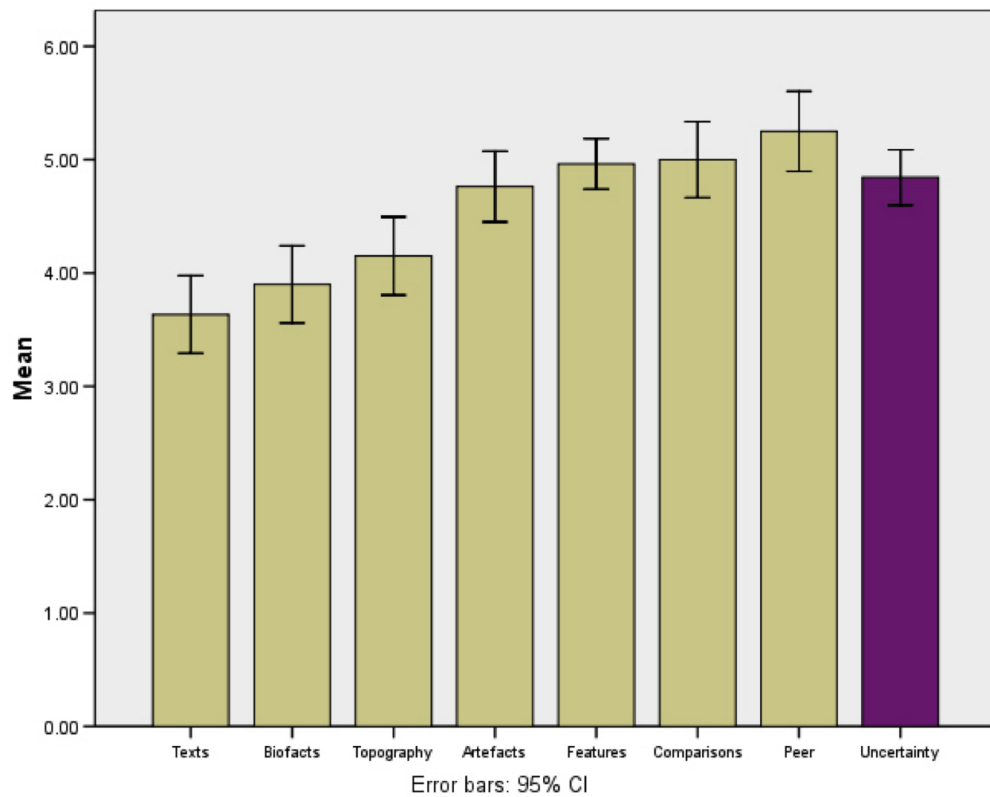
Figure 3.3: Means of averaged variables

Part 8

The last part involves eight (8) variables, each accounting for a specific identified factor. The KGZ test (Table 3.10) indicated that there was no difference between the two groups (sig>0.05).

The next step was to examine the group means. Table 3.11 and Figure 3.4 show the means calculations. Features and artefacts indicate higher scores for their means (6.85 and 5.70 out of 8 respectively) while the comparisons and peer input are between 4.50 and 4.90. Lastly, the biofacts and textual evidence are located at the lower end of the scale.

Factor	Kolmogorov-Smirnov Z	Exact Sig. (2-tailed)	Most Extreme diff
Features	0.671	0.498	0.3
Artefacts	0.447	902	0.2
Biofacts	0.447	0.953	0.2
Topography	0.224	1	0.1
Contextual comparisons	0.447	0.783	0.2
Absolute comparisons	0.447	0.969	0.2
Textual evidence	0.224	1	0.1
Peer input	0.447	0.978	0.2

Table 3.10: KGZ test for Part 8

Factors	Sum	Mean	Std. Deviation	Variance
Features	137	6.85	2.059	4.239
Artefacts	114	5.7	2.055	4.221
Absolute comparisons	98	4.9	2.1	4.411
Contextual comparisons	90	4.5	1.539	2.368
Peer input/outside support	88	4.4	2.479	6.147
Topography	78	3.9	1.553	2.411
Biofacts	63	3.15	1.843	3.397
Textual evidence	52	2.6	1.847	3.411

Table 3.11: Means of influencing factors from Part 8
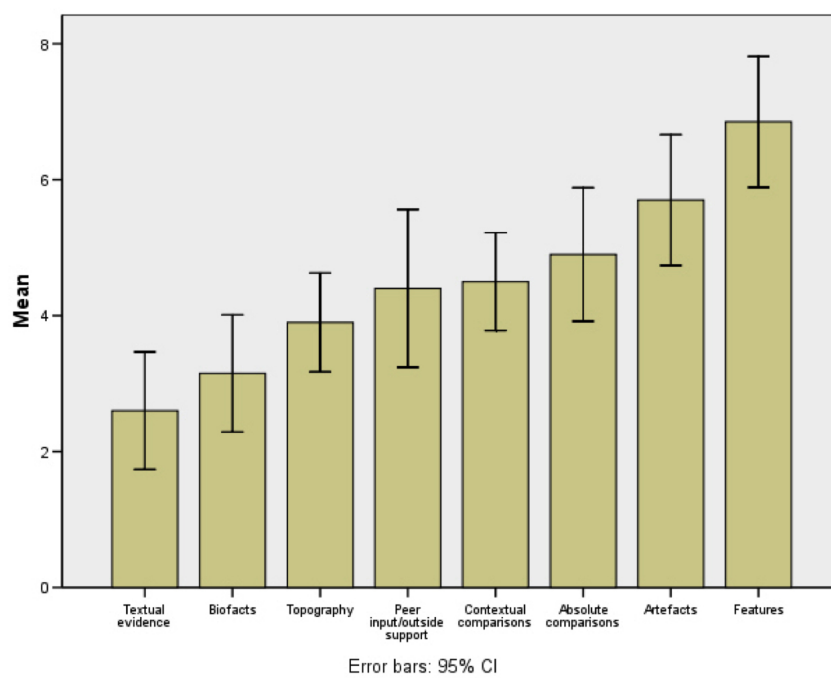
Figure 3.4: Graphs of means for Part 8

3.2.6 Discussion

The questionnaire examined the attitude of archaeologists towards uncertainty and whether there is any difference between Roman archaeologists and those of other backgrounds when interpreting archaeological structures. Additionally, it examined whether there is any preferential order between the identified factors.

Evidence

Throughout the three parts, results indicate that there is no significant difference between the two groups. Furthermore, the archaeologists seem to place more importance on peer feedback and evidence like features and artefacts. Accordingly, less importance seems to be placed to textual evidence and biofacts.

Feedback received from the evidence combination parts equally suggests that combinations of the above strong factors are considered more favourably when making interpretations. The results obtained show that there can be preference amongst different evidence which may be influential in visualisations of uncertainty.

Going back to the discussion in Section 2.6.2 regarding the weight and strength of evidence in law, perhaps this preference amongst archaeologists is something similar. It would be interesting to explore whether any overconfidence or under-confidence (such as mentioned in 2.4.5) occurs when making judgement with different kinds of evidence.

Uncertainty

Figure 3.5 shows the results on questions archaeologists were asked about uncertainty. A large majority considers two or more equally possible interpretations, when interpreting structural remains. They also acknowledge that other interpretations could be equally viable to the ones they make themselves. Furthermore, they tend to agree that they maybe more than one interpretation of a site while they are concerned about the prevalence of uncertainty in archaeology. Lastly, even if just more than half of them feel sure about their interpretations given the evidence they have collected, 70% of them end up with one interpretation.

These results strengthen the importance uncertainty has on the archaeological discipline. It is interesting to note that even though uncertainty is prevalent in their opinions, most of them end up with one interpretation more often than not. This is to be expected as a decision-making procedure; the available evidence gives weight

towards one hypothesis rather than another, and is subsequently supported. The concept of equally valid interpretations is also understandable; given that the evidence is never complete, archaeologists accept the fact that other experts may interpret the data differently.

Lastly, this gives the impression that archaeologists may consider hypotheses in an *open world* state of mind. Consider an archaeologist who excavates a site and accepts as valid scenarios A,B, or C. Due to evidence her judged interpretation is C. However, new recovered evidence points to A,B,C and also D. Provided that the new evidence is strong and credible enough to account for this, the expert has to re-evaluate the given evidence and consider a new hypothesis.

3.3   Conclusion

The questionnaire results highlighted some important issues. Firstly, archaeologists are concerned about uncertainty. Secondly, even after they evaluate all evidence and provide interpretations, the confidence in said interpretation is not 100%. Thirdly, there are indications that more credibility may be placed in some types of evidence; as in being more indicative to provide the correct interpretation. All these should be taken under consideration in the modelling of archaeological uncertainty.

In Sec. 2.3 different definitions and categorisations of uncertainty and ignorance were presented. It is clear that there is no fixed definition that can be given. However, the definitions have something in common. There is a dividing line between what is known, measured, vague, fuzzy or missing in regards to the data, and what is believed about the data. Smet's [106] model represents this concept well. Figure 3.6 indicates the relationships and differences between the aspects of imperfect data as identified by Smets.

Consider the concept of archaeological uncertainty involved in a 3D reconstruction. A reconstruction reflects the belief of the expert in how the building might have looked like, given existing knowledge and her interpretation of evidence at hand. There is no *knowing* for certain, rather a quantity of *belief, based on accumulated evidence,* that the reconstruction is a representation of the past. If new evidence/facts/information arrives (i.e. missing data is found), then the belief can change. In other words, the archaeologist/expert is expressing belief about a statement (the reconstruction) based on a body of knowledge that might include precise, imprecise, incomplete, vague and/or conflicting data.

Often or more Sometimes Seldom or less

(a) Archaeology is fraught with uncertainty; you cannot be too sure about anything.

(b) Other alternatives to my interpretation could be equally viable.

(c) I consider two, or more, valid interpretations.

(d) I feel sure about my interpretation given the available evidence.

(e) There is only one true interpretation of a site.

(f) I end up with one interpretation and no alternatives.

Figure 3.5: Questions on Uncertainty
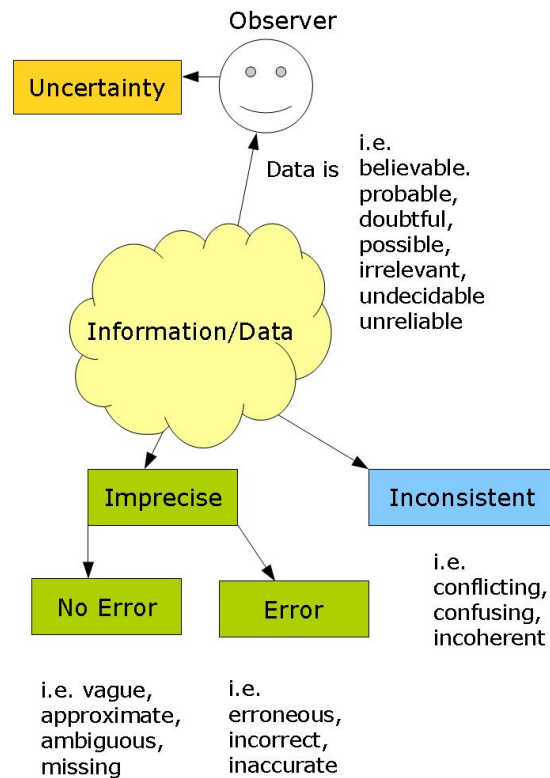
Figure 3.6: Representing imperfect knowledge (based on [106])

An example is given to illustrate the concept, in Figure 3.7. A plain column shaft has been discovered in an ancient Greek site. Consider Sally, a fresh archaeology graduate. Based on the data available, and her background knowledge, to suggest an interpretation, she would not be able to choose between Ionic and Corinthian orders. Despite the measurement of data being precise and complete, the uncertainty is reflected on her interpretation. A more experienced archaeologist, John, suggests that Corinthian columns were less common than Ionic ones–his accumulated knowledge includes a related fact that Sally was not aware of. As a result, he favours that this

was possibly an Ionic column.

State of Knowledge

Data ◄──── Artefact

24 flutes
220mm.
diam
1 m. heigh
Ancient
Greek loc.

+

Archaeologist's
Accumulated
knowledge

Interpretation

What is it?	A shaft from a column	Accumulated knowledge
What kind of column?	Ionic, Doric, Corinthian?	Possible worlds
Re-evaluate?	Doric usually has 20 flutes. It is not a possibility.	Accumulated knowledge
What kind of column?	Ionic, Corinthian?	
Suggest?	*Both are equally possible and I cannot choose.*	**Uncertainty reflected on state of knowledge**

Figure 3.7: Representing uncertainty on a state of knowledge

While both are using the same recovered data, each has different accumulated knowledge and they use both to express their belief. Both of their suggestions will be valid until new evidence is found that would support Corinthian or Ionic.

Great care must be taken to distinguish precision and certainty from truth. The expert/observer can be precise and certain about an event P even though in actuality she may be wrong. For example, consider the following proposition P: *"I am certain*

the Olympian gods were thirteen in number". While the statement is precise and certain, it is false since the gods are twelve and the data fed to the statement is erroneous. The certainty in a proposition can be reflected as saying *"I believe that* P*"* but not as "I know that P". To believe in P does not necessarily make P true.

This is the fundamental difference between the concept of belief and knowledge; saying "I know that P" means that not only there is a belief in P but P is also true. Conclusively, subjective uncertainty deals directly with beliefs *about the knowledge*.

During the course of this thesis the terms *belief* and *degrees of belief* are used interchangeably. As a result, the following phrase defines the approach of this thesis towards archaeological uncertainty:

> *Archaeological uncertainty can be defined as the degree of belief of the expert in the*
> *interpretation/reconstruction of an archaeological structure or object. The degree*
> *of belief depends on the expert's state of knowledge at the given moment.*

The following two Chapters apply mathematical methods of quantifying an expert's belief based on knowledge, to archaeological scenarios. Subjective probability, fuzzy sets and possibility theory as well as evidence theory are explained and applied to simulated archaeological cases.

Lastly, attention must also be placed in the *state of knowledge.* The extent, and quality, of relevant knowledge is what reduces uncertainty and results in better-supported hypotheses. As discussed in Chapter1, a large part of the archaeological inferential process is making analogies with similar discoveries. This is also evident with the absolute and contextual comparison categories identified by the archaeologists in this Chapter.

Consider the following scenario: During an excavation, Sally, the archaeology graduate, unearths a piece of pottery but has troubles identifying it. She uploads a picture of it to a database of archaeological information, and compares her find with other pottery shards that are perhaps better identified by expert archaeologists. Having completed her inferential process, her uncertainty is now at a reduced state. The concepts of searchable knowledge states, and describable evidence are further explored in Chapter 7.

It is a truth very certain that
when it is not in our power to
determine what is true, we ought
to follow what is most probable.

R. Descartes

Quantifying Uncertainty: Probability Approach

One of the contributions of this thesis is the application of subjective probability as a means of quantifying archaeological uncertainty. The mathematical foundations are set by discussing classical logic, mathematical axioms of probability theory, and the derived Bayes' Theorem. Then, the application to archaeological scenarios is demonstrated. This is followed by a discussion evaluating potential benefits and drawbacks of using subjective probability.

4.1   Introduction

The word *probability* in the Oxford English Dictionary has three definitions [194]:

1. *The property or fact of being probable, esp. of being uncertain but more likely than not; the extent to which something is likely to happen or be the case; the appearance of truth, or likelihood of being realized, which a statement or event bears in the light of present evidence.*

2. *An instance of the property or fact of being probable; a probable event or circumstance; a thing judged likely to be true, to exist, or to happen.*

3. *As a measurable quantity: the extent to which a particular event is likely to occur, or a particular situation be the case, as measured by the relative frequency of occurrence of events of the same kind in the whole course of experience, and expressed by a number between 0 and 1. An event that cannot happen has probability 0; one that is certain to*

happen has probability 1. Probability is commonly estimated by the ratio of the number of successful cases to the total number of possible cases, derived mathematically using known properties of the distribution of events, or estimated logically by inferential or inductive reasoning (when mathematical concepts may be inapplicable or insufficient).

As the above quotation shows, to give a definition of probability one must consider the expected values, or outcomes, of probability. Will the values measure a tendency of something to occur, or are they a measure of how strongly one believes it will occur? These three examples demonstrate which probability might be used when faced with uncertainty:

1. Roll a dice, toss a coin.

2. Predict the weather, the team that will win the FA Cup.

3. Debate whether there is life on other planets, or the existence of a gold deposit in an area.

The first scenario is what people mostly associate with probability–games of chance, dice, roulette, the flip of a coin. Otherwise known as objective, physical, or frequentist, for the purposes of this thesis, the term *Frequentist Probability* (FP) will be used to refer to this approach. In such scenarios, an event such as "The coin will land on heads" occurs at a persistent rate in a long run of trials. Thus, FP makes sense only when dealing with well-defined, random, and repeatable experiments.

The second scenario implies that data must be used in order to assign probabilities. Taking the weather scenario under consideration, if it has been observed that with current weather conditions the day is sunny 10% of the days, the probability of the sun coming out today will be 10%. This value will, of course, change, as more weather data is gathered. Probability in this kind of scenario is termed Statistical Probability (SP).

The third scenario can be assigned to any kind of statement, even when no random process is involved. The difference from the previous two scenarios is that there is already an outcome, there is no awareness, however, what it is. A gold deposit may exist or it may not. An expert's opinion, could indicate a 60% certainty–this does not mean that the deposit is there 60% of the time and the others moves away, rather it reflects the expert's opinion that the chances are a bit better than half. In other words, it reflects the degree to which the statement is supported by the available evidence. In this form, subjective (also referred to as evidential, or epistemic) probability can be

expressed as *degrees of belief*–how likely one would be to gamble against certain odds. Bayesian Probability is a form of subjective probability.

Regardless of the different philosophical aspects of probability, there exist mathematical foundations which are followed in all scenarios. The following sections present an introduction in classical logic, probability laws, conditional probability and Bayes' theorem.

4.2   Principles of classical logic

Classical logic sets a number of formal rules for the handling of logic both in a mathematical and a philosophical manner. It is not in the purpose of this thesis to expand on the evolution of classical logic from the philosophical years of Aristotle to the laying down of the mathematical foundations in the 20th century–relevant information can be found in [195]. Three are the major laws of Classical logic, the law of bivalence, contradiction, and excluded middle.

The law of bivalence states that any proposition P must *either* be true or false.

The law of contradiction judges as false any proposition P asserting that both proposition P and its denial, proposition ¬P, are true at the same time and in the same respect. In the words of Aristotle, "*One cannot say of something that it is and that it is not in the same respect and at the same time*"[196]. For any proposition P, it cannot be both the case that P and −P (not P) is true. Symbolically, this is expressed as:

$$\neg(P \wedge \neg P) \tag{4.1}$$

For example, if P is *Tim is bald*, then the inclusive conjunction *Tim is bald, and Tim is not bald* is false.

The law of excluded middle states that for any proposition P, it is true that either P, or its denial ¬P is true at the same time and in the same respect. Symbolically:

$$P \vee \neg P \tag{4.2}$$

For example, if P is *Tim is bald*, then the inclusive disjunction *Tim is bald, or Tim is not bald* is true. This is not quite the same as the principle of bivalence, which states that P must be either true or false. It also differs from the law of non-contradiction, which states that $\neg(P \wedge \neg P$ is true. The law of excluded middle only says that the total $(P \wedge \neg P)$ is true, but does not comment on what truth values P itself may take.

In any case, the semantics of any bivalent (or two-valued logics, such as classical) logic will assign opposite truth values to P and ¬P (i.e., if P is true, then ¬P is false), so the law of excluded middle will be equivalent to the principle of bivalence in a bivalent logic. Classical logic and the derived Boolean relationships are very relevant to probability theory.

However, the same cannot be said about non-bivalent logics, or many-valued logics. These logic systems may have either different or analogous laws, or reject the law of excluded middle in its entirety. Fuzzy sets, and the derived fuzzy logic, reject the law of bivalence and consequently, the law of excluded middle. This is further analysed in Chapter 5, however the following simple example illustrates the logical differences between the two.

Consider the statement P *The glass is full of water.* With classical logic, the glass could either be full or totally empty. A half-full glass, however, counts as both full and empty at the same time. This particular multi-valent logic, assigns a matter of degree to each proposition instead of complete truth or falsity.

4.3   Probability calculus

It is worth briefly expanding on important properties of set theory as it defines the rules and conventions of how to handle *probability events.* This aids in the further discussions of probability and the consequent differences between classical sets and fuzzy sets.

Classical set theory

A set is a collection of objects, concrete, or abstract, hereby termed as *elements*. If S is a set and x an element of it, this is expressed as $x \in S$; if x does not belong to S then $x \notin S$. For each use of set theory, all objects that are relevant constitute a set that is called a *universe*. A common way to define any set S that consists of some objects of universe Ω is to assign the number 1 to each member of Ω that is also a member of A and to assign the number 0 to the remaining members of Ω. This assignation is otherwise known as *characteristic function*. Given that the characteristic function of each $x \in \Omega$, a set A can also be described as in Eq. 4.3:

$$A(x) = \begin{cases} 1 & \text{if } x \in \Omega \\ 0 & \text{if } x \notin \Omega \end{cases} \tag{4.3}$$

It is important to note that the range $[0, 1]$ is purely symbolic; any other meaningful expression could have been used such as $[\mathtt{false}, \mathtt{true}]$. For example, the characteristic function f of the set of real numbers from 10 to 20 is:

$$fA(x) = \begin{cases} 1 & \text{if } 10 \leqq x \leqq 20 \\ 0 & \text{if } \mathtt{otherwise} \end{cases}$$

Figure 4.1 illustrates the concept; notice the sharp boundaries that distinguish between membership and non-membership of the set.
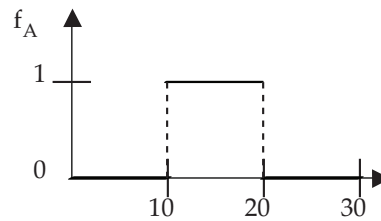
Figure 4.1: The characteristic function f of the set of real numbers from 10 to 20

A set is considered *finite* and/or *countable* if the number of elements is finite and the contained elements are countable by any conventional labelling. For example, set $A = \{1, 2, 3...n\}$ is finite to $|A| = n$; with the same logic, the set of natural numbers, $\mathbb{N}$ is not finite but is countable.

Let A and B be sets that form part of a universe X. The symbol "|" is to be read as *given that*; for example $\{x \mid x \ \mathtt{satisfies} \ P\}$. Briefly, the operations that can be performed on sets are:

$$
\begin{array}{rrll}
\text{Union} & A \cup B = & \{x \mid x \in A \text{ or } x \in B\} & (4.4) \\
\text{Intersection} & A \cap B = & \{x \mid x \in A \text{ and } x \in B\} & (4.5) \\
\text{Complement} & A^c = & \{x \mid x \notin A\} & (4.6) \\
\text{Difference} & A \setminus B = & A \notin B^c & (4.7) \\
\text{Contradiction law} & A \cap A^c = & \emptyset & (4.8) \\
\text{Excluded middle law} & A \cup A^c = & X & (4.9) \\
\text{Commutativity} & A \cap B = & B \cap A & (4.10) \\
& B \cup A = & A \cup B & \\
\text{Associativity} & (A \cap B) \cap C = & A \cap (B \cap C) & (4.11) \\
& (A \cup B) \cup C = & A \cup (B \cup C & \\
\text{Distributivity} & A \cap (B \cup C) = & (A \cap B) \cup (A \cap C) & (4.12) \\
& A \cup (B \cap C) = & A \cup (B \cup C) & \\
\text{Idempotence} & A \cup A = & A & (4.13) \\
& A \cap A = & A & \\
\text{De Morgan Laws} & (A \cap B)^c = & A^c \cup B^c & (4.14) \\
& (A \cup B)^c = & A^c \cap B^c &
\end{array}
$$

Two sets are considered **disjoint** if their intersection is empty–in other words, if they have no elements in common. A **partition** of a set S is a collection of sets that are disjoint and their union is S. Figure 4.2 illustrates in practice sets and associated operations:
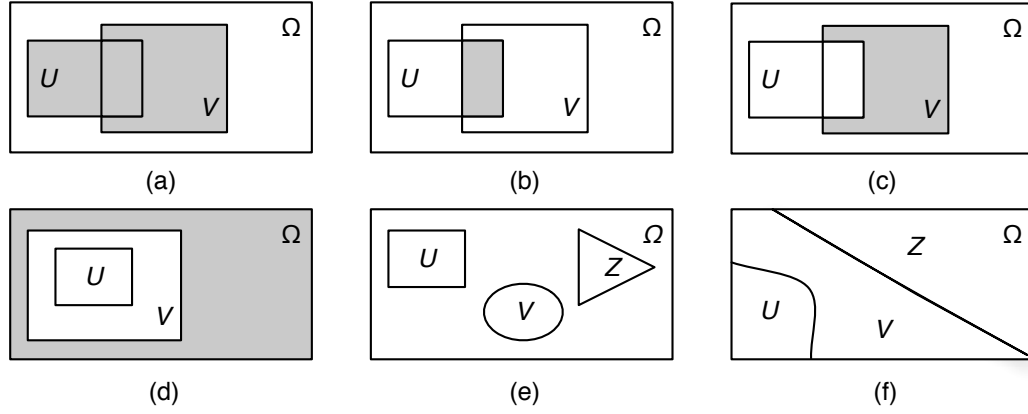
Figure 4.2: Crisp sets examples: (a) The shaded region is $U \cup V$. (b) The shaded region is $U \cap V$. (c) The shaded region is $U \cap V^c$. (d) $U \subset V$ and the shaded region is V^c. (e) The sets U, V, Z are disjoint. (f) The sets U, V, Z form a partition of set Ω.

4.3.1 Probability axioms

A probabilistic model is a mathematical description of an uncertain situation. It is composed of two main elements, the *sample space* Ω and the probability law. The sample space defines the set of all possible *outcomes* of an experiment. The probability law assigns to a subset A (or event A) of possible outcomes a non-negative number Pr(A) which represents the belief about the likelihood of the elements of A. Thus, the mathematical expression Pr(A) is read as the "probability of (the event/statement) A". The probability axioms are summarised in Table 4.1.

Non-negativity:

For every event A: $Pr(A) \geqq 0$

Additivity:

For non-disjoint events A and B: $Pr(A) \cup B = Pr(A) + Pr(B) - Pr(A \cap B)$

Additivity:

For disjoint events A and B,

the probability of their union satisfies: $Pr(A) \cup B = Pr(A) + Pr(B)$

For a sequence of disjoint events: $Pr(A_1 \cup A_2 \cup \cdots A_n = Pr(A_1) + Pr(A_2) + \ldots + Pr(A_n)$

Normalisation:

The probability of the entire

sample space Ω is equal to one $Pr(\Omega) = 1$

Complement:

For an event not happening $\bar{A}$ $Pr(\bar{A}) = 1 - Pr(A)$

<div align="center">Table 4.1: Probability axioms</div>

From the normalisation and additivity axioms, it is also derived that $1 = Pr(\Omega) = Pr(\Omega) \cup Pr(\emptyset) = Pr(\Omega) + Pr(\emptyset) = 1 + Pr(\emptyset)$–and thus, the probability of an empty event is zero $Pr(\emptyset) = 0$.

The following example is provided as an illustration of the axioms. Let Ω be a collection of 50 Roman denarii (coins). 15 of them have a high silver content, 20 a high bronze content, 5 have both silver and bronze, while 10 are mostly made of bulion. The Venn diagram in Figure 4.3 describes the scenario. Let $Pr(A)$ be the coins with silver content, and $Pr(B)$ the coins with bronze content.
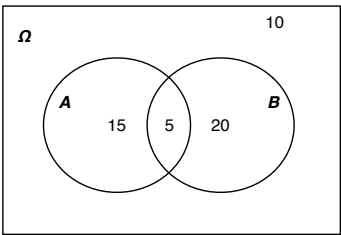
<div align="center">Figure 4.3: Venn diagram of Roman denarii</div>

The Karnaugh table (Table 4.2) explains the probabilities.

	A	Ā	Total	A	Ā	Total
B	$\Pr(A \cap B)$	$\Pr(\bar{A} \cap B)$	$\Pr(B)$	$5/50$	$20/50$	$25/50$
B̄	$\Pr(A \cap \bar{B})$	$\Pr(\bar{A} \cap \bar{B})$	$\Pr(\bar{B})$ $\Rightarrow$	$15/50$	$10/50$	$25/50$
Total	$\Pr(A)$	$\Pr(\bar{A})$	1	$20/50$	$30/50$	1

Table 4.2: The probabilities of coins

4.3.2 Conditional probability, independence, and total probability

Conditional probability refers to the probability of an event occurring if some condition has been applied. We write $\Pr(B \mid A)$ to mean $\Pr(B$ occurs given that A has occurred). The conditional probability formula is given in Eq. 4.15.

$$\Pr(B \text{ given } A) = \frac{\Pr(B \text{ AND } A)}{\Pr(A)} == \Pr(B \mid A) = \frac{\Pr(B \cap A)}{\Pr(A)} \tag{4.15}$$

Re-arranging Eq. 4.16 gives us the *multiplication law*:

$$\Pr(B \cap A) = \Pr(A)\Pr(B \mid A) \tag{4.16}$$

If events A and B are independent (i.e. A happening has no effect on B whatsoever or vice verca), $\Pr(B \mid A) = \Pr(B)$ and thus Eq. 4.16 reduces to $\Pr(B \cap A) = \Pr(B) \cdot \Pr(A)$. Similarly, $\Pr(A \mid B) = \Pr(A)$ and $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$.

Looking back at Fig. 4.2, a partition of a set Ω is formed by events $A_1...A_n$ if these events are disjoint ($A_i \cap A_j = \emptyset$; for $i \neq j$) and their union sums up to Ω ($A_1 \cup ...A_n = \Omega$). As a result, for any conditional event B, the *total probability law* holds:

$$\Pr(B) = \sum_{j=1}^{n} \Pr(A_j)\Pr(B \mid A_j) \tag{4.17}$$

The following example illustrates the law: An Early Anglo-Saxon graveyard consists of 60% males, 30% females, and 10 % children. Burials containing votive deposits were 45% in male graves, 35% in females and 20% in children. What is the probability that a randomly selected grave contains a votive deposit?

Let Ω be the cemetery population, with $\Pr(Male) = 0.6$, $\Pr(Female) = 0.3$ and $\Pr(Child) = 0.1$. Let B be the probability of a votive offering $\Pr(B \mid Male) = 0.45$, $\Pr(B \mid Female) = 0.35$, $\Pr(B \mid Child) = 0.20$, Using Eq. 4.17:

$$\begin{aligned} \Pr(B) &= \Pr(B \mid Male)\Pr(Male) + \Pr(B \mid Female)\Pr(Female) \\ &\quad + \Pr(B \mid Child)\Pr(Child) \\ &= 0.6 \cdot 0.45 + 0.3 \cdot 0.35 + 0.1 \cdot 0.2 = 0.395 \end{aligned}$$

4.3.3   Bayes' Theorem

Bayes' theorem [113] deals with *posterior probability*, a conditional probability of H given E, where H actually occurs first. In other words, given that E occurred, what is the probability that it happened through H_j? Equation 4.18 is the Bayes' theorem. Let $H_1...H_n$ be disjoint events, forming a partition of sample space Ω and $Pr(H_j) \geqq 0$ for all j. Then, for any event E, $Pr(E) \geqq 0$:

$$
\begin{aligned}
Pr(H \mid E) &= \frac{Pr(E_j \cap H)}{Pr(E)} \\
\text{multiplication law (eq. 4.16)} &= \frac{Pr(H_j)Pr(E \mid H_j)}{Pr(E)} \\
\text{total probability law (eq. 4.17)} &= \frac{Pr(H_j)Pr(E \mid H_j)}{\sum_{j=1}^{n} Pr(H_j)Pr(E \mid H_j)}
\end{aligned}
$$

$$
Pr(H_j \mid E) = \frac{Pr(H_j)Pr(E \mid H_j)}{Pr(H_1)P(E \mid H_1) + ... + Pr(H_n)P(E \mid H_n)} \tag{4.18}
$$

The concept behind Bayesian reasoning is the use of a prior condition in order to calculate a posterior–termed as *a priori* and *a posteriori*. For example, questions such as: "*What is the probability that a person has illness X given symptoms Y occurred?*" or, from the example in Section 4.3.2 "*What is the probability that bones from grave X belong to a male, given that a votive deposit was found?*" illustrate the concept of $Pr(Hypothesis|Evidence)$. Bayes' Theorem is used to calculate the probability that a hypothesis is true based on the available evidence.

To calculate this, the following are required: the probability of getting the evidence if the hypothesis is true, the probability of getting the evidence if the hypothesis is false, and how likely it would be that the hypothesis is true if the particular evidence *was not available*. The last requirement is the prior condition, this could represent a subjective belief, or even a state of uncertainty, that, through the feedback of evidence it is further strengthened or weakened. It should be stressed that evidence does not affect the actual probability that the hypothesis is true; rather, it alters the state of knowledge about it.

Taking the Anglo-Saxon graveyard example, calculating $Pr(Male|Votive)$ is:

$$
\begin{aligned}
Pr(M|V) &= \frac{Pr(M) \cdot Pr(V|M)}{Pr(M) \cdot Pr(V|M) + Pr(F) \cdot Pr(V|F) + Pr(C) \cdot Pr(V|C)} \\
&= \frac{0.6 \cdot 0.45}{0.6 \cdot 0.45 + 0.3 \cdot 0.35 + 0.1 \cdot 0.2} = 0.683
\end{aligned}
$$

Calculating $\Pr(F|V)$ and $\Pr(C|V)$ yields 0.258 and 0.051 respectively. In other words, the probability distribution of an Anglo-Saxon graveyard was {0.6, 0.3, 0.1} before recovering a votive deposit where it changed to {0.683, 0.258, 0.051} for our current dig. These two distributions are the *prior* and the *posterior*, respectively.

4.4 A probabilistic approach to archaeological uncertainty

This thesis proposes that one way to approach archaeological uncertainty is through subjective (Bayesian) probability. The term *probability* is used here strictly within its subjective meaning, representing belief and likelihood. In other words, when creating an archaeological reconstruction, the expert places her belief on the reconstruction's process. Belief increases if the evidence is corroborative, or concrete, and decreases if it is conflicting, or lacking.

The introduction to Bayes Theorem has shown that simple conditional probability by itself is not sufficient to handle the impact added by evidence in a model. The following (archaeologically simple) example illustrates the difference between these two concepts.

In an archaeological dig, a column base was discovered. Subsequently, pieces of a column top were also unearthed. A reconstruction must be produced. The archaeologist has the following knowledge about columns *before* unearthing anything: Columns belong to one of two orders, M and N. Either is as likely to be found {0.5,0.5}. These columns are distinguished usually by their bases (types A and B) with a distribution (0.95, 0.05) for order M and (0.25,0.75) for order N. Figure 4.4 illustrates the concept.

Column bases are always followed by pillars (two styles, C, D) and tops (two styles, E, F). It has been noted that column M has pillar C most of the times (0.8,0.2) and top F (0.3, 0.7). Column N has pillar D more often (0.1,0.9) and top E (0.6,0.4). The recovered evidence are a base of type A and a top of type F.

The example is now displayed by means of a Karnaugh table (Table 4.3). This style of representation is quite important and will be used extensively through the thesis.

If one were to (erroneously) rely only on simple conditional probability, the results, presented in Table 4.4 arise for the belief in a model fitting the style of column M. By using the Karnaugh table, it is easy to calculate the belief by multiplying between the respective rows. For example, for a model in the M style with Base B, Pillar D and Top F, the belief would be $0.05 * 0.2 * 0.7$.
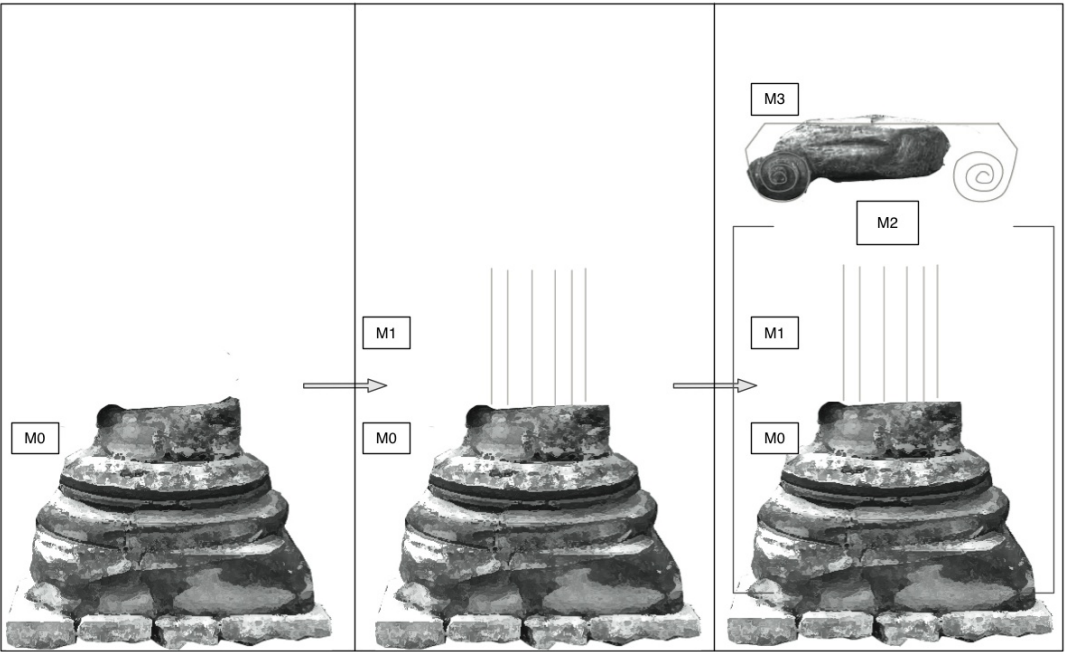
Figure 4.4: Column example

	Column M	Column N
Base A	0.95	0.25
Base B	0.05	0.75

	Column M	Column N
Pillar C	0.8	0.1
Pillar D	0.2	0.9

	Column M	Column N
Top E	0.3	0.6
Top F	0.7	0.4

	Column M	Column N
Colour Y	0.1	0.7
Colour Z	0.9	0.3

Table 4.3: The probabilities of two column styles, M and N.

Column M			
ACE	0.228	BCE	0.012
ACF	0.532	BCF	0.028
ADE	0.057	BDE	0.003
ADF	0.133	BDF	0.007

Column N			
ACE	0.015	BCE	0.045
ACF	0.01	BCF	0.03
ADE	0.135	BDE	0.405
ADF	0.09	BDF	0.27

Table 4.4: Conditional options

Even if a base of type A, a pillar C and a top F are discovered, the best probability for a column M interpretation is 0.532. The reliability result for the discovery of a base

A and a top F is 0.665–by considering both pillar types.

It is easy to observe, that by adding more evidence, this approach leads to a lesser and lesser belief. The reason is that alternative scenarios where the suggested hypothesis does not hold are not taken under consideration.

The Bayesian approach will be now considered. The question asks for a belief to be placed on the evidence. Given that a base A and consequently a top F were found, what is the likelihood that the column belongs to style M? To consider this question, the prior belief is needed (a non-committing value of $0.5, 0.5$), and the conditional values used before, found in the Karnaugh table. In more detail:

$$
\begin{aligned}
\Pr(M \mid A) &= \frac{\Pr(M)(\Pr(A \mid M)}{\Pr(M)\Pr(A \mid M) + \Pr(N)\Pr(A \mid N)} \\
&= \frac{0.5 \cdot 0.95}{0.5 \cdot 0.95 + 0.5 \cdot 0.25} \\
&= 0.791
\end{aligned}
$$

The belief that a column of style M has been discovered, has increased to 0.791 after the discovery of base A. The new belief values for M and N respectively are $\{0.791, 0.209\}$. In order to add the new discovery of top F, a similar approach is followed, but the prior belief this time comes from the new values incorporating evidence of base A:

$$
\begin{aligned}
\Pr(M \mid F) &= \frac{\Pr(M)(\Pr(F \mid M)}{\Pr(M)\Pr(F \mid M) + \Pr(N)\Pr(F \mid N)} \\
&= \frac{0.791 \cdot 0.7}{0.791 \cdot 0.7 + 0.209 \cdot 0.4} \\
&= 0.868
\end{aligned}
$$

It is also possible to combine multiple evidence together - for example in a scenario where two different parts are found simultaneously. Re-examining the base/top

scenario:

$$
\begin{aligned}
\Pr(M \mid A, F) &= \Pr(M).\Pr(A \mid M).\Pr(F \mid M) \\
&= (0.5, 0.5).(0.95, 0.25).(0.7, 0.4) \\
normalisation &= \frac{0.5\text{x}0.95\text{x}0.7,\ 0.5\text{x}0.25\text{x}0.4}{0.5\text{x}0.95\text{x}0.7 + 0.5\text{x}0.25\text{x}0.4} \\
&= \frac{0.3325}{0.3825},\frac{0.05}{0.3825} \\
belief\ in\ \{M,\ N\} &= 0.868,\ 0.132 \\
\Pr(M \mid A, F) &= 0.868
\end{aligned}
$$

Lastly, to demonstrate how belief decreases in the face of contradicting evidence, let it be supposed that a pillar of type D was found (most commonly associated with columns of style N):

$$
\begin{aligned}
\Pr(M \mid D) &= \frac{\Pr(M)(\Pr(D \mid M)}{\Pr(M)\Pr(D \mid M) + \Pr(N)\Pr(D \mid N)} \\
&= \frac{0.868 \cdot 0.2}{0.868 \cdot 0.2 + 0.132 \cdot 0.9} \\
&= 0.59
\end{aligned}
$$

The confidence in an interpretation of column M has decreased with the contradicting evidence.

Lastly, consider two archaeologists (Mr X and Mr Y) who are judging the interpretation. Each has his own subjective prior belief about the distribution of column types M and N. Mr X assigns a prior belief of {0.7, 0.3} for {M,N} while Mr. Y assigns {0.3, 0.7}. After accumulating the evidence of base A and top F, the calculated subjective beliefs are {0.939, 0.06} for Mr. X and {0.74, 0.26} for Mr. Y.

It is interesting to note that through the accumulation of evidence, Mr. X's already strong belief in the existence of M column types has been strengthened significantly. He would be very surprised if it turns out to be a column N. Similarly, the low belief in the existence of M columns shown by Mr. Y has turned to a high confidence, equal to the *a priori* belief of his colleague. If further evidence is discovered, the beliefs of these two people will actually be very similar. The latter example illustrates that even with different *a priori* beliefs, supporting evidence will eventually stabilise the confidence in the interpretation.

4.5 Conclusion

This Chapter has shown how it is possible, by using a Bayesian approach, to evaluate the belief in a reconstruction, or parts of it. The Bayesian approach demonstrates how beliefs can change in an interpretation with the addition of supporting (or contradicting) evidence. The requirement of a prior belief might be an issue–what if the *a priori* belief is biased towards one interpretation? While the examples have shown that with the accumulation of evidence the belief would adjust accordingly, other issues might also arise. Through the application of Bayes' Theorem to archaeological examples, the following issues have been identified:

Prior belief and ignorance

The prior belief requirement fits very well in cases where the expert has adequate background information on the existence of similar scenarios. What if, however, the expert is ignorant? Does not know? Classical Bayesian approaches assign an equi-probability to ignorance. For example, for two cases A and B with no prior knowledge, the expert should assign a prior of $\{0.5, 0.5\}$. However, semantically, this is not correct as $\{0.5, 0.5\}$ can also reflect knowledge that either A or B are likely to occur.

Strict totality of distribution

The laws of Probability (which Bayes' Theorem inherits) dictate that between two cases, A and B if one assigns a confidence of 0.95 to A, they should assign a confidence of 0.05 to B. However, they may be cases where this may be too restrictive or force the archaeologist to accept assumptions that they would not be comfortable with.

The effect of 0 likelihood

If one assigns a likelihood of 0 to a hypothesis, this hypothesis may never be resurrected no matter the evidence. In other words, 0 will not represent total ignorance but a total improbability.

Forced granularity

Consider a scenario which relates a particular Roman roof tile to two different types of roofs. The archaeologist knows with 100% certainty that the only recorded association for this tile has been roofs of type A and B but is not willing to give a sub-probability.

The additivity law forces us to assign 0.5 and 0.5 to each case. There is no way to assign a confidence in a less granular option.

Closed world

Probability theory (and a large number of theories of ignorance) postulate the existence of a closed world. The hypotheses tested are exhaustive and inclusive. However, there may be cases where one would need the option of adding a new hypothesis.

Bayes' Theorem is an good choice of mathematical tool to be used if one wishes to quantify subjective belief with well-measured conditions and distributions of evidence. The next Chapter examines non-probabilistic approaches to uncertainty which tackle the majority of the issues identified above.

Alice laughed. "There's no use trying," she said: "one can't believe impossible things."

Through the Looking Glass

<div align="right">

CHAPTER 5

</div>

Quantifying Uncertainty: Non-Probabilistic Approaches

The previous Chapter discussed Subjective Probability and how it can be used to quantify archaeological uncertainty. Evaluation indicated a number of issues that may arise in certain cases. This Chapter explores non-probabilistic methods to uncertainty. The contributions presented include two applications of quantifying archaeological uncertainty. These are based on two different non-probabilistic models: Possibility Theory and the Transferable Belief which has its roots on Evidence Theory. The mathematical foundations of both are explored followed by their applications to archaeological scenarios. An evaluation is provided as conclusion.

5.1 Epistemic Possibility Theory

As probability is based on crisp sets (described in Section 4.3) so is Possibility Theory based on fuzzy sets. Possibility theory is a theory of uncertainty specialising in the handling of incomplete information [123]. The term *Possibility Theory* was invented by L. Zadeh, who in turn, was inspired by a paper by Gaines and Kohout [197]. Possibility theory does not contradict probability theory; rather, it complements it. The major difference between the two is that Possibility theory uses a pair of dual-set functions, known as possibility and necessity, instead of the single function used by probability.

While during its conception, Possibility Theory was created to deal with graded semantics to natural language statements [123], during the years of its development

[198, 199] it has evolved to a representation of partial belief that parallels and complements probability.

In Section 4 it was discussed how probability has different meanings and uses depending on the context, even if the mathematical foundations remain consistent. A similar case occurs with Possibility Theory–it can support various interpretations [200], such as:

1. Logical: This approach deals with consistency of available information. By giving a possibility to a proposition, we accept that related information is not contradicted.

2. Feasibility: The ease of achievement; how easy it is to solve a problem given that certain constrains exist.

3. Plausibility: This reflects the belief or tendency of events to occur. A regular expression might be *It is plausible that ...*

Hacking [201] distinguishes between possibility as an objective or subjective (epistemic) notion. In the first case, it reflects properties of the physical world, while on the second it represents the state of knowledge of an agent. As with probability, this thesis is related to the subjective notion of possibility. Possibility measure reflects the idea of plausibility and its dual function, necessity relates to certainty or belief. In other words, the certainty of an event, reflects a lack of plausibility of its opposite. This is a major difference to probability theory which is self-dual.

In Probability Theory:

$$It\ is\ probable\ that\ "A" = It\ is\ not\ probable\ that\ "not\ A"$$

In Possibility Theory:

$$It\ is\ possible\ that\ "A"\ \neq\ It\ is\ not\ possible\ that\ "not\ A"$$

This section provides a brief overview of fuzzy sets in order to further explain how Possibility Theory works, and how it differs from fuzzy logic.

5.1.1   Fuzzy set theory

Fuzzy sets were introduced in 1965 in a seminal paper by L. A. Zadeh[122]. Fuzzy sets, contrary to classical sets, are not required to have sharp boundaries that distin-

guish members of a set from non-members. A membership in a fuzzy set does not signify concrete belonging to the set; rather it reflects a matter of *degree* of belonging. A fuzzy set A of a universe X is defined by a function that assigns to each object x in X a *membership degree* of x in A. As classical sets have the characteristic function, fuzzy sets display the membership function.

In order to express membership degree, a number is usually used from the unit interval $[0, 1]$. During the description of classical sets, it was mentioned that the interval $\{0, 1\}$ is arbitrary and does not hold any numerical significance. In the case of fuzzy sets, however, the interval does have a numerical significance. By allowing degrees of membership during this interval, fuzzy sets are able to express a gradual transition, or belonging, from being a member to not being a member of a set. Due to this ability, one of the first usages of fuzzy sets was to describe linguistic functions such as *very*, *a lot*, *a little*, a concept which is difficult to describe with classical sets.

Definition: A fuzzy set F is a pair (Ω, μ_F) where Ω is a set and μ_F is the mapping of Ω to the unit interval $[0, 1]$.

$$\mu_F : \Omega \to [0, 1] \tag{5.1}$$

$$\text{degree of membership of } \omega \text{ in } F \quad \mu_F(\omega) \text{ for } \omega \in \Omega \tag{5.2}$$

The following archaeological example serves as an illustration. Consider the expression *"The Roman wall was very high."* How high, is very high? If one accepts that, for example, above 4 meters is very high, would that imply that 3.95 meters is not high at all? In the concept of classical (or crisp) sets, one would have to concede to the latter. Figure 5.1 illustrates four membership functions, all describing the wall. In order to relate the example with the aforementioned definition, set Ω is considered to represent the heights of Roman walls. The membership of the function in a fuzzy set, will express how much the value (or the object) ω is compatible with the concept of *"heights"*.

Functions A and B are derived from classical sets. Set A contains a single object, the number 4. This can be expressed as *"High is four meters."* Set B contains numbers between 3 and 5, which can be expressed as "High is between three and five meters."

Functions C and D describe fuzzy sets. Set C gradually begins from 2 and gradually ends at 6. This can mean *"High is around four meters."* The difference with the classic set B is that while both consist of numbers around 4, the boundaries are extremely sharp at set B thus anything more than 5 meters is not considered *high* anymore. Set
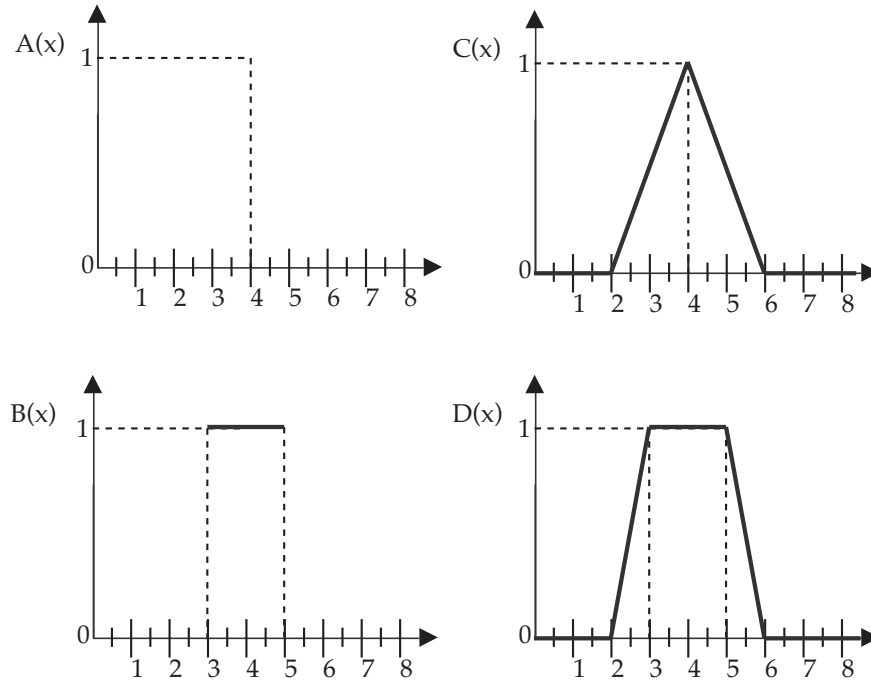
Figure 5.1: Classic and fuzzy sets; Sets A&B demonstrate classic sets while C&D are fuzzy sets.

D represents an interval-valued fuzzy set. An interval-valued fuzzy set is a set whose membership function is many-valued and forms an interval in the membership scale [202, :86]. In other words, the membership functions of an interval-valued fuzzy set, maps objects to intervals of real numbers.

Basic operations on fuzzy sets are defined by their membership function, as classical sets by their characteristic function. In brief, for two fuzzy sets A and B specified by membership functions $\mu_A(x)$ and $\mu_B(x)$:

- Inclusion: $A \subseteq B$ if $\mu_A(x) \leq \mu_B(x)$ for all $x \in X$

- Complement: $A^c = 1 - \mu_A(x)$ Elements that negate the assessment of A.

- Union: $(A \cup B(x)) = max[\mu_A(x)\ \mu_B(x)]$ Elements that belong to either A or B

- Intersection: $(A \cap B)(x) = min[\mu_A(x),\ \mu_B(x)]$ Elements belonging to both A and B

Two important properties of fuzzy sets that clearly distinguish them from classical ones, is the non-adherence to the law of excluded middle (Eq. 4.9) and the law of

contradiction (Eq. 4.8). Firstly, the intersection of a fuzzy set A with its complement A^c is not equal to an empty set:

$$A \cap A^c \neq \varnothing \qquad (5.3)$$

Secondly, the union of a fuzzy set A with its complement A^c is not equal to the fundamental set X.

$$A \cup A^c \neq X \qquad (5.4)$$

Graphs in Fig. 5.2 illustrate operations on fuzzy sets:
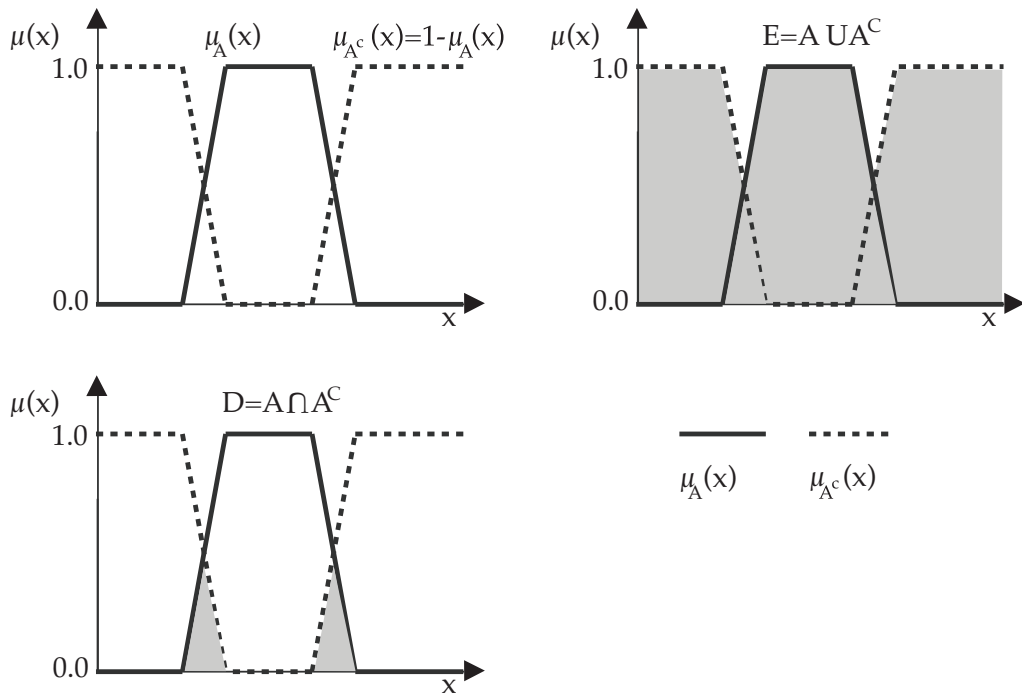
Figure 5.2: Fuzzy set operations

The remaining properties, described in Section 4.3 that hold for crisp sets, also hold for fuzzy sets.

Concluding, it should be stressed that while some properties of fuzzy sets look similar to probability expressions, they are built to describe different concepts. Fuzzy sets are analogous to crisp sets and probability theory is analogous to Possibility Theory. It is erroneous to confuse $Pr(A)$ with $\mu_A(\omega)$ (the probability of A with a mem-

bership grade in A). When $\text{Pr}(A)$ is considered, the set A is well defined but the value of the variable x to which Pr is attached is unknown (and perhaps random). On the other hand, with the membership grade, x is known while the set is ill defined.

Fuzzy sets form the basis for fuzzy logic and Possibility Theory. Before expanding on the mathematical principles of these two methods, the conceptual differences will be examined. While both methods are based on fuzzy sets, they handle different problem areas. Fuzzy logic deals with *degrees of truth*, while Possibility Theory tackles *degrees of uncertainty* and incomplete knowledge. This difference is crucial and an example is necessary to illustrate it, at least in a conceptual level:

Consider a full glass of water, similar to the example used for classical logic. In terms of binary logic, the glass is full or empty. Accounting for the quantity of water in the glass, it can be said that the glass is half full. The word *full* is actually a fuzzy predicate and the degree of truth of the expression *the bottle is full* represents the amount of water in the glass. On the other hand, consider an expression of ignorance about whether the glass is full or not. Expressing an ignorance of 0.5 does not imply that the glass is half empty. Looking back at the discussion of ignorance (Section 2.3) *degrees of truth* can be associated with imprecision while *degrees of uncertainty* can be related to belief and the uncertainty aspect of ignorance.

5.1.2 Fuzzy logic

Fuzzy logic [122] can be viewed as a special kind of many-valued logic. The truth value of a proposition P instead of assuming two values $(false, true)$ or $(0, 1)$ it can assume any value between $[0, 1]$. This value indicates the degree of truth of the proposition. For instance, let $P(y)$ represent the height of a tall column. The truth values of $P(10)$ and $P(0.5)$ are certainly 1 and 0 respectively. A truth value of $P(4)$ maybe something in the range of 0 and 1, such as 0.4. The mathematical representation:

Let A and B be propositions that accept truth (Tr) values to the range of $[0, 1]$:

$$\text{Tr}(A) = \text{Tr}(B) \text{ if } B = A \tag{5.5}$$
$$\text{Tr}(A) = 1 - \text{Tr}(B) \text{ if } A = \bar{B} \tag{5.6}$$
$$\text{Tr}(A \vee B) = \min[\text{Tr}(A), \text{Tr}(B)] \tag{5.7}$$
$$\text{Tr}(A \wedge B) = \max[\text{Tr}(A), \text{Tr}(B)] \tag{5.8}$$

As can be seen from Eqs. 5.7 and 5.8 fuzzy logic reflects the properties of fuzzy sets and as a result, the non-conformity to the laws of excluded middle and contradiction.

There is a natural connection between the degrees of membership in a fuzzy set, and the degrees of truth in fuzzy propositions. Fuzzy logic can be considered as a gradual representation of belonging.

5.1.3 Possibility theory

An expression such as X is F where X is a variable and F a fuzzy set (for example, wall is high) can be used in two different types of situations. Both of them take under consideration the fuzziness of F. If the value of X is precisely known and what is estimated is whether this value is in fact compatible with the fuzzy set F, we are interested in the *gradual* or *elastic* nature of the expression X is F. For example, if we are looking for a person to fit the description of *old*, we estimate to what extent the person fits this requirement and can be qualified as such (i.e. Paul, whose age is known to be 63 can be regarded as *old* to the degree 0.8 in a context of {70-100}). This gradual nature of properties is expressed through fuzzy logic.

On the other hand, the expression X is F in other scenarios could also mean all that is *known regarding the value* of X, is that X is F. In other words, not knowing precisely the value of X. When fuzzy sets are used in that context, the degree attached to the value of X is the level of *possibility* that it is indeed the value of the variable. As a result, the fuzzy set F is expressed as a *possibility distribution* [123]which offers various shades of plausibility on the values of the variable X. For example, if it is only known that Paul is old (but not his precise age) where the meaning of old is described by the membership function of a fuzzy set μ_{tall}, then the greater $\mu_{tall}x$ is, the greater the possibility that $age(Paul) = x$ (and vice verca). This is the essence of *Possibility Theory*.

In [123], Zadeh defines a possibility measure Pos:

Given a $[0, 1]$ possibility distribution π that describes an incomplete state of knowledge:

$$Pos(A) = sup\{\pi(x), x \; makes \; A \; true\} \tag{5.9}$$

where A is a Boolean proposition, being either true or false. For two Boolean propositions, A and B it follows that:

$$Pos(A \cup B) = max(Pos(A), Pos(B)) \tag{5.10}$$

and that:

$$Pos(A \cap B) \leq min(Pos(A),\ Pos(B)) \tag{5.11}$$

Possibility theory uses two measures to represent an event: Possibility and Necessity. These are represented as $Pos(A)$ and $Nec(A)$. Possibility and Necessity can be seen as dual set measures. The possibility measure $Pos()$ on a set Ω for an event A is the degree of possibility that A occurs.

$$Pos(\varnothing) = 0 \tag{5.12}$$
$$Pos(\Omega) = 1 \tag{5.13}$$

Fundamental axioms in Possibility Theory differ from the probability ones:

Definition: The disjunction of two events A and B is the maximum of their individual possibilities.

$$Disjunction \quad Pos(A \cup B) = max\{Pos(A),\ Pos(B)\} \tag{5.14}$$

Definition: The necessity of an event A is the negation of the possibility of complement of A:

$$Nec(A) = 1 - Pos(\bar{A}) \tag{5.15}$$

$Pos(A) = 1$ signifies that A is fully possible. The relationships between possibility and necessity are [198]:

$$Nec(A) > 0 \quad \Rightarrow \quad Pos(A) = 1 \tag{5.16}$$
$$Pos(A) < 1 \quad \Rightarrow \quad Nec(A) = 0 \tag{5.17}$$
$$Pos(A) \quad \geq \quad Nec(A) \tag{5.18}$$
$$Nec(A) + Nec(\bar{A}) \quad \leq \quad 1$$
$$Pos(A) + Pos(\bar{A}) \quad \geq \quad 1 \tag{5.19}$$

Eq. 5.18 expresses the fact that something must be possible to some extent before it can begin to be certain. Eq. 5.19 demonstrates the non-deterministic nature of Possibility Theory. While with probability the sum of an event A and its complement must

add up to 1, in possibility it is not required. This allows scenarios where a particular piece of evidence A could be fully possible, and at the same time its complement could be rather possible as well.

Let A and B be two subsets of a sample space Ω:

$$\begin{aligned}
\text{Nec}(A \cap B) &= \min\{\text{Nec}(A), \text{Nec}(B)\} & (5.20) \\
\text{Pos}(A \cup B) &= \max\{\text{Pos}(A), \text{Pos}(B)\} & \\
\text{Pos}(A \cap B) &\leq \min\{\text{Pos}(A), \text{Pos}(B)\} & (5.21) \\
\text{Nec}(A \cup B) &\geq \max\{\text{Nec}(A), \text{Nec}(B)\} & (5.22)
\end{aligned}$$

As can bee seen from Eq. 5.20 a fundamental difference of Possibility Theory from probability is that it is non-compositional. The measures of possibility are decomposable in respect to the union, and the calculation of Necessity is compositional with respect to intersection. For example, if one is completely ignorant about an event A we have $\text{Pos}(A) = \text{Pos}(\bar{A}) = 1$ and $\text{Nec}(A) = \text{Nec}(\bar{A}) = 0$. Also, $\text{Pos}(A \cap \bar{A}) = 0\}$ and $\text{Nec}(A \cup \bar{A}) = 1\}$.

In order to analyse the conceptual difference between possibility and necessity a simple example is presented. Consider the statement P "*The possible interpretations of a discovered pyramid.*" Let the universe of discourse Ω contain these interpretations {Egyptian, Minoan, Roman}. Consequently, consider the following expert opinion from archaeologist X: "*I believe that it is fully possible to be of Egyptian origin. I am not quite sure about Minoan, and even less about Roman.*" Converting this to a possibility distribution, it could be $\text{Pos}_{\text{Egyptian, Minoan, Roman}} = \{1, 0.4, 0.2\}$. In effect, the most plausible situation compatible with A, namely, the interpretation as an Egyptian pyramid, is the one that is chosen, judging from the maxitivity axiom 5.14. Deriving the necessity distribution with Eq. 5.15, gives $\text{Nec}_{\text{Egyptian, Minoan, Roman}} = \{0.6, 0, 0\}$.

These two dual values reflect the concept that while it could most likely be an Egyptian pyramid, there could also be other valid interpretations. As a result, the certainty of our belief in the pyramid being of Egyptian origin, represented by Nec, is 0.6, which is $1 - \max\text{Pos}_{\text{Egyptian, Minoan, Roman}}$.

An extreme, but also possible, scenario, would be an archaeologist saying "*I believe equally that the pyramid could be Egyptian or Minoan, but there is no way it can be Roman.*" This represents a complete state of ignorance since the expert is unable to prioritise between the first two interpretations. The possibility distribution of that statement is: $\text{Pos}_{\text{Egyptian, Minoan Roman,}} = \{1, 1, 0\}$. In that case, the statement is not plausible, and

$\mathrm{Nec}_{\mathrm{Egyptian,\ Minoan,\ Roman}} = \{0,\ 0\ 0\}$. With the evidence in hand, there is no way to be certain about either of these interpretations, however possible, or believable, they may be.

It follows that, the closer Nec is to 1, the more plausible our interpretation is. In order for the plausibility in an interpretation to increase or decrease, supporting or contradicting evidence needs to be added.

5.2   A possibilistic approach to archaeological uncertainty

In Section 4.4 it was demonstrated that it is possible to employ Bayesian probabilities in order to describe archaeological uncertainty. Section 5.2 presents an approach using Possibility theory. To recollect, Possibility Theory does not contradict probability theory; rather, it complements it. The major difference between the two is that Possibility theory uses a pair of dual-set functions, (Possibility and Necessity), instead of the single function used by probability.

In [203, 204] a possibilistic approach to the Bayes theorem is proposed where the additivity normalisation part of Bayes' equation is paralled to the maxitivity axiom of Possibility Theory. The equation, based on subjective possibilities, takes under consideration both possibility and necessity measures, and can be described as:

$$\mathrm{Pos}(H_j \mid E) = \frac{\mathrm{Pos}(H_j)\mathrm{Pr}(E \mid H_j)}{\max\{\mathrm{Pos}(H_1)P(E \mid H_1), \mathrm{Pos}(H_n)P(E \mid H_n)\}} \tag{5.23}$$

There is one restriction; to define the subjective possibilities, one of the possible scenarios must have $\mathrm{Pos} = 1$. In other words, at least one of the suggested interpretations must be fully possible. For reasons of clarity, the hypothesis that is fully possible will be defined as *Dominant Hypothesis* (DH). Dominance does not indicate correctness or validation, but simply indicates which hypothesis is currently considered as fully possible. As evidence is added, the dominance characteristic can shift from one hypothesis to another.

Consider again the example given in 4.4 summarised briefly here:

In an archaeological dig, a column base was discovered followed by pieces of a column top. The archaeologist has the following knowledge about columns *before* unearthing anything: Columns belong to one of two orders, M and N. These columns are distinguished usually by their bases (types A and B). Column bases are always followed by pillars and tops. The related evidence distribution can be seen in Table 5.1.

Until now, the description matches the one for the Bayesian approach. The interpretation again requires a *prior*. In the Bayesian approach, the archaeologist was non-deterministic stating that both column types are as likely to be found, with the sum of that likeliness having to add up to 1. In Possibility Theory, this can be translated as $Pos_{M,N,} = \{1, 1\}$. I and $Nec_{MN} = \{0, 0\}$. This is a unique case where both hypotheses are dominant as they are considered equally possible to be correct.

After evidence of a Base A is found, we can calculate $Pos(M \mid A)$ and $Pos(N \mid A)$:

$$
\begin{aligned}
Pos(M \mid A) &= \frac{Pos(M)(Pos(A \mid M)}{max(Pos(M)Pos(A \mid M), Pos(N)Pos(A \mid N))} \\
&= \frac{1 \cdot 0.95}{max(1 \cdot 0.95, 1 \cdot 0.25)} \\
&= 1 \\
Pos(N \mid A) &= \frac{Pos(N)(Pos(A \mid N)}{max(Pos(M)Pos(A \mid M), Pos(N)Pos(A \mid N))} \\
&= \frac{1 \cdot 0.25}{max(1 \cdot 0.25, 1 \cdot 0.95)} \\
&= 0.26
\end{aligned}
$$

So, the new possibilities are $Pos_{M,N)} = \{1, 0.26\}$. Because $Nec(A) = 1 - Pos(\bar{A})$, the revised $Nec_M = 0.74$ resulting to $Nec_{M,N)} = \{0.74, 0\}$.

The latter result, the outcome of Necessity, is the amount of certainty placed on the specific interpretation. This is the number of interest in the possibilistic representation of uncertainty (Table 5.2) for the hypothesis to be consider fully plausible.

As with the Bayesian example, we can add more evidence, in the form of a recovered Top F:

	Column M	Column N
Base A	0.95	0.25
Base B	0.05	0.75

	Column M	Column N
Pillar C	0.8	0.1
Pillar D	0.2	0.9

	Column M	Column N
Top E	0.3	0.6
Top F	0.7	0.4

	Column M	Column N
Colour Y	0.1	0.7
Colour Z	0.9	0.3

Table 5.1: The knowledge base of two column styles, Mand N.

Prior	M	N	Posterior	M	N
Pos	1	1	Pos	1	0.26
Nec	0	0	Nec	0.74	0

Table 5.2: Subjective possibility prior and posterior

$$
\begin{aligned}
\text{Pos}(M \mid F) &= \frac{\text{Pos}(M)(\text{Pos}(F \mid M)}{\max(\text{Pos}(M)\text{Pos}(F \mid M), \text{Pos}(N)\text{Pos}(F \mid N))} \\
&= \frac{1 \cdot 0.7}{\max(1 \cdot 0.7, 0.26 \cdot 0.4)} \\
&= 1 \\
\text{Pos}(N \mid F) &= \frac{\text{Pos}(N)(\text{Pos}(F \mid N)}{\max(\text{Pos}(M)\text{Pos}(F \mid M), \text{Pos}(N)\text{Pos}(F \mid N))} \\
&= \frac{0.26 \cdot 0.4}{\max(1 \cdot 0.7, 0.26 \cdot 0.4)} \\
&= 0.15
\end{aligned}
$$

The evidence could have been found at the same time, and it is possible to consider them simultaneously as with a Bayesian scenario:

$$
\begin{aligned}
\text{Pos}(M \mid A, F) &= \text{Pos}(M).\text{Pos}(A \mid M).\text{Pos}(F \mid M) \\
&= (1, 1).(0.95, 0.25).(0.7, 0.4) \\
\text{normalisation} &= \frac{1\text{x}0.95\text{x}0.7, \ 1\text{x}0.25\text{x}0.4}{\max(1\text{x}0.95\text{x}0.7, 1\text{x}0.25\text{x}0.4)} \\
&= \frac{0.665}{0.665}, \frac{0.1}{0.665} \\
\text{Possibility of } \{M \mid AF, \ N|AF\} &= 1, 0.15 \\
\text{Necessity of } \{M \mid AF, \ N|AF\} &= 0.85, 0
\end{aligned}
$$

Lastly, following the Bayesian example, discovery of contradicting evidence does change the strong belief in M, as when a Pillar D is found:

$$Pos(M \mid D) = \frac{Pos(M)(Pos(D \mid M)}{max(Pos(M)Pos(D \mid M), Pos(N)Pos(D \mid N))}$$

$$= \frac{1 \cdot 0.2}{max(1 \cdot 0.2, 0.15 \cdot 0.9)}$$

$$= 1$$

$$Pos(N \mid D) = \frac{Pos(N)(Pos(D \mid N)}{max(Pos(M)Pos(D \mid M), Pos(D)Pos(D \mid N))}$$

$$= \frac{0.15 \cdot 0.9}{max(1 \cdot 0.2, 0.15 \cdot 0.9)}$$

$$= 0.675$$

The last result has decreased the $Nec(M|A, F, D = 1 - 0.675 = 0.325$. There is a threshold, when evidence amounts against the interpretation of M where the confidence switches to an N interpretation, and N becomes the DH. For example, consider an extra piece of evidence Y, related to colour.

$$Pos(M \mid Y) = \frac{Pos(M)(Pos(Y \mid M)}{max(Pos(M)Pos(Y \mid M), Pos(N)Pos(Y \mid N))}$$

$$= \frac{1 \cdot 0.1}{max(1 \cdot 0.1, 0.675 \cdot 0.7)}$$

$$= 0.211$$

$$Pos(N \mid Y) = \frac{Pos(N)(Pos(Y \mid N}{max(Pos(M)Pos(Y \mid M), Pos(N)Pos(Y \mid N))})$$

$$= \frac{0.675 \cdot 0.7}{max(1 \cdot 0.1, 0.675 \cdot 0.7)}$$

$$= 1$$

The updated Possibility value indicates that an N interpretation is now fully possible (and somewhat certain since $Nec = 1 - 0.211 = 0.788$) given the new evidence.

As a result, a common evidence pool can be used for both Bayesian and Possibilistic calculations if required.

5.3   The Transferable Belief Model

The Transferable Belief Model (TBM) is a non-probabilistic model of belief representation proposed and extended by Phillippe Smets [187, 120, 205]. It is an interpretation

of the DST model, the main difference being that all connections to a probabilistic model are explicitly rejected. The DST model itself will not be expanded here, the reader is referred to extensive descriptions and comparisons in [206, 207, 208]. Where the TBM deviates from the DST or probabilistic assumptions, this will be made clear in the discussion.

5.3.1   Introduction to the TBM

The TBM works in two levels:

1. The Credal Level (CL) (from *credus* meaning belief in Latin)

2. The Pignistic Level (PL) (from *pignus*, meaning bet, in Latin)

The Credal Level always precedes the Pignistic Level. The CL is where beliefs are entertained and assigned. This is where beliefs are quantified by using belief functions. If any new information arrives, or new beliefs need to be included, it will occur in the CL.

Once all beliefs are gathered, revised, and combined, and if it is necessary (it is not *required*), then the PL assists in making decisions about those beliefs. Beliefs in the PL induce a probabilistic measure; in other words, the beliefs (from CL) are transformed to probabilities in the PL.

In subjective probabilities, as it has been shown, the CL does not exist, and it is embedded within the decision-making process. The transformation to probabilities does not mean that the results will always be the same as in the Bayesian approach.

TBM, apart from its disassociation from the probabilistic model in the construction of belief functions, also differs from DST, possibilistic and probabilistic approaches in other ways, the most important being:

1. It features a dynamic transfer of belief

2. It is not restricted to a closed world assumption.

3. It has two levels (CL/PL) and features the pignistic transformations (otherwise known as Bets)

4. It contains unnormalised belief functions.

The universe of discourse Ω is usually postulated (i.e. in the majority of subjective Probability or Possibility interpretations) as a closed set of hypotheses on which beliefs are entertained. Any proposition outside this domain, is considered improbable (or impossible, depending on what you use). The TBM suggests that this an idealistic viewpoint and that the cognitive process is rarely that simple or so complete that it can include all possible propositions [209]. Three sets of propositions are presented:

- Known as Possible (KP)

- Known as Impossible (KI)

- Unknown (UP)

In a classical Bayesian approach, the UP is definitely empty and one must accept a closed world assumption–that the truth is inside the KP. The TBM works with all those three sets and, according to accummulating evidence, may move a hypothesis from KP to KI via dynamic transfer, for example:

- There is evidence about a KP proposition A which makes it impossible: Conditioning redistributes it to KI.

- There is new evidence about an UP, perhaps a forgotten or unthought-of hypothesis. The UP is transferred to KP.

Each piece of evidence induces a finite (maximum of 1) amount of belief (called *mass of belief*) that must be allocated among different propositions. The mass can be assigned to a proposition or a subset of propositions. For example, supposed a murder scene has three suspects: {Tom, Tim, Sally}. Evidence arises that the perpetrator is male; the belief mass of this evidence goes to hypothesis A ={Tom, Tim}. While with a subjective probability approach we would have to equally split the belief between the two, the TBM does not require this. It represents that the evidence is not clear enough to distinguish between Tom and Tim.

The TBM, like other models for belief quantification (be those possibilistic or probabilistic in nature), has two components of operation [210]. The first component, is the static one, which is where beliefs are allocated. In the TBM this is where the bbms are assigned. The second one is the dynamic component–where beliefs are updated. In the TBM this is where bbms are transferred or updated among the propositions (perhaps due to new evidence, or to conditioning). It is also where beliefs that are induced by different distinct pieces of evidence are combined. It should be noted

that both of those components still happen at the CL. One last difference between the TBM and other models of belief representation is the order of these updating rules. For example, in DST the order below is being followed:

1. The static component (bpms) is assigned

2. Evidence is combined (by using Dempster's rule of combination)

3. Any updating process occurs (by using Dempster's rule of conditioning)

In the TBM:

1. The static component (bbms) is assigned

2. Any updating process (i.e the mass transfer, or a discounting)

3. Evidence is combined by using (unnormalised) Dempster's rule of combination

5.3.2   Belief, combination, discounting and conditioning

The major element of the TBM is the basic belief assignment *(referred to as bba or m).*[1]

For $A \subseteq \Omega$, $m(A)$ is the part of belief that gives support to A. The total belief has a maximum of 1. $[Ev]$ is the particular evidence, of fact, that gives rise to the mass:

$$m^{\Omega}[Ev](A) \in [0, 1]$$

$$\sum_{A \subseteq \Omega} m[Ev](A) = 1$$

Recall from Section 4.3 and Section 5.1.3 that the Probabilities and Possibilities of empty events are 0 ($Pr(\varnothing)=0$; $Poss(\varnothing) = 0$). The same holds for DST, where $m_{bpm}(\varnothing) = 0$. In the TBM this restriction does not hold. A mass given to $\varnothing$ is belief in the UP; an unknown proposition, the concept of the open world. Perhaps, as in the murder example, evidence arises which give support to Sally or someone else other than Tim or Tom.

---

[1]DST users may recognise this notation as basic probability assignment from the DST. Because of the strict absence of probability in the CL, P. Smets does not use the term *probability* and promotes the term *belief* instead.

The degree of *belief* $\text{bel}(A)$ is defined as:$\text{bel} : 2^\Omega \to [0, 1]$ for all $A \subseteq \Omega$. 2^Ω reflects all subsets of the world Ω. Thus the belief function bel is:

$$\text{bel}^\Omega[Ev](A) = \sum_{\varnothing \neq B \subseteq A} m(B) \tag{5.24}$$

Eq. 5.24 represents the *justified specific support [120]* given to A. It is called *justified* because the belief in A gets only support from the bbms to subsets of A. It is called *specific* because it will always exclude $m(\varnothing)$, the mass of the $\varnothing$ subset.

The degree of *plausibility* $\text{pl}(A)$ is defined as: $\text{pl} : 2^\Omega \to [0, 1]$ for all $A \subseteq \Omega$. Thus the plausibility function pl is:

$$\text{pl}^\Omega[Ev](A) = \sum_{B \subseteq \Omega, B \cap A \neq \varnothing} m(B) = \text{bel}(\Omega) - \text{bel}(\bar{A}) \tag{5.25}$$

Eq. 5.25 represents the *potential specific support* given to *A*. What this means, is that some of the bbms contained in $\text{pl}(A)$ may be transferred to other subsets of *A* if new evidence becomes available.

Lastly, the commonality function, q, [211] can be defined as: $\text{pl} : 2^\Omega \to [0, 1]$ for all $A \subseteq \Omega$. Thus:

$$q^\Omega[Ev](A) = \sum_{B \subseteq \Omega, A \subseteq B} m(B) \tag{5.26}$$

The plausibility function *pl* is another way of representing information contained in the belief (bel) function. It is not required, but, along with the commonality function, q, they can be mathematically useful when representing and combining beliefs.

In order to combine evidence, consider $m^\Omega[Ev1]$ and $m^\Omega[Ev2]$, two distinct bbas that lend support to subsets of Ω. To calculate $m^\Omega[Ev1, Ev2] = m^\Omega[Ev1] \cap m^\Omega[Ev2]$:

$$m^\Omega[Ev1] \cap m^\Omega[Ev2](A) = \sum_{B, C \subseteq \Omega, B \cap C = A} m^\Omega[Ev1](B) \cap m^\Omega[Ev2](C) \forall A \subseteq \Omega \tag{5.27}$$

$$q^\Omega[Ev1] \cap m^\Omega[Ev2](A) = q^\Omega[Ev1](A) \cap q^\Omega[Ev2](A) \forall A \subseteq \Omega \tag{5.28}$$

What Eq. 5.27 instructs is the following: Take two pieces of evidence (Ev1,Ev2) giving support (through their masses) to different subsets of A. Multiply each subset of Ev1 with each and every one of Ev2. Where subsets discuss, contain a common hypothesis (symptoms, diseases, murder victims, archaeological interpretations) it is

considered a non-empty intersection. Those that do not intersect their multiplication will be considered as belonging to $m(\varnothing)$. In the end, add the common intersections together. The commonality combination gives the same result, by using the commonality function.

Conditioning is a special case of combination, where evidence arises that completely excludes proposition(s) from the list of KP and thus they are to be transferred to KI. Consider an existing mass m and new information in the form of $m[Ev]$ saying that only $B \in \Omega$. Thus:

$$m[Ev1](B) = \sum_{C \subseteq \bar{A},} m(B \cup C) \quad \text{if } B \subseteq A$$
$$= \qquad\qquad 0 \qquad\qquad otherwise \tag{5.29}$$

The Conditioning equation postulates that one should transfer all mass pertaining to a now-defunct hypothesis to any subsets that used to contain this hypothesis.

Another property of the TBM is its ability to represent total ignorance through the *vacuous belief* function. The *vacuous function* is such as: $m^{\Omega} = 1$ and $bel(A) = 0, A \neq \Omega$ and so $bel(\Omega) = 1$. This is suitable for representing a state of total ignorance about which hypothesis may possibly be true when no information is available. The usefulness is demonstrated in Section 5.4.1 which deals with the Generalised Bayesian Theorem (GBT) .

Lastly, the Pignistic Transform can be calculated as follows:

$$BetP^{\Omega}(A) = \sum_{A:\omega \in A \subseteq \Omega} \frac{m^{\Omega}(A)}{|A|(1 - m^{\Omega}(\varnothing))} \tag{5.30}$$

Eq. 5.30 transforms beliefs held at the CL to a pignistic decision (PL) as a probability function.

In order to illustrate the capabilities of the TBM and explain the terminology, an example will now be presented. A good majority of TBM examples [209, 103, 212] use a fictional murder mystery where the TBM assists an investigator to assemble witness evidence and find out who the murderer is. In this example, the TBM will assist You, a fictional archaeologist, in an archaeological reconstruction.

5.3.3 Archaeological example

A small-scale excavation has just started on a Roman building. The remains of a room have been found. Your task is to discover what the room

was used for, or at least get a good estimate of that. You take the position of the room into account and, combined with the years of experience, you suggest that it must be one of:

$$\Omega = \{\texttt{Bath}, \texttt{Dining}, \texttt{Bedroom}, \texttt{Kitchen}, \texttt{Storage}\}.$$

You found the following evidence:

- **Evid. 1** (biofacts analysis): *"A good quantity of edible seeds has been found. It could be a place where food was stored or consumed."*

    - $m_1(\{\texttt{Dining}, \texttt{Kitchen}, \texttt{Storage}\}) = 0.7 \; m(\Omega) = 0.3$

- **Evid. 2** (room shape): *"Compared to similar buildings, the shape of the room is similar to either a Bath, a Bedroom or a Kitchen."*

    - $m_2(\{\texttt{Bath}, \texttt{Bedroom}, \texttt{Kitchen}\}) = 0.7 \; m(\Omega) = 0.3$

- **Evid. 3** (pottery found): *"A few pieces of broken jars that held wine. I think these belong to a Kitchen or a Storage area."*

    - $m_3(\{\texttt{Kitchen}, \texttt{Storage}\}) = 0.6 \; m(\Omega) = 0.4$

Combining Evidence

To combine Evidence 1 with Evidence 2:

		$Ev_1 \; m_1$	
		Din,Kit,Stor	Ω
Ev_2	m_2	0.7	0.3
Bath, Bed, Kit	0.7	{Kit} 0.49	{Bath, Bed, Kit} 0.21
Ω	0.3	{Din, Kit, Stor} 0.21	{Ω} 0.09

Table 5.3: TBM Combination of Ev_1 and Ev_2

Using Eq. 5.27:

$$
\begin{aligned}
m_{Ev1Ev2}\{\texttt{Kitchen}\} &= 0.49 & m_{Ev1Ev2}\{\texttt{Din}, \texttt{Kit}, \texttt{Stor}\} &= 0.21 \\
m_{Ev1Ev2}\{\Omega\} &= 0.09 & m_{Ev1Ev2}\{\texttt{Bath}, \texttt{Bed}, \texttt{Kit}\} &= 0.21
\end{aligned}
$$

It should be noted that the combinations are associative and the order with which they are combined does not matter. To combine Evidence 12 with Evidence 3:

Ev_{12} m_{12}					
		Kit	Din,Kit,Stor	Bath,Bed,Kit	Ω
Ev_3	m_3	0.49	0.21	0.21	0.09
Kit, Stor	0.6	{Kit} 0.294	{Kit, St} 0.126	{Kit} 0.126	{Kit, St} 0.054
Ω	0.4	{Kit} 0.196	{Din, Kit, Stor} 0.084	{Bath, Bed, Kit} 0.084	{Ω} 0.036

Table 5.4: TBM Combination of Ev_{12} and Ev_3

Again using Eq. 5.27, this time for Ev_{12} and Ev_3:

$$m_{Ev123}\{Kit\} = \quad 0.294 + 0.196 + 0.126 \quad = 0.616$$
$$m_{Ev123}\{Kit, Stor\} = \quad 0.126 + 0.054 \quad = 0.18$$
$$m_{Ev123}\{Din, Kit, Stor\} = \quad 0.084$$
$$m_{Ev123}\{Bath, Bed, Kit\} = \quad 0.084$$
$$m_{Ev123}\{\Omega\} = \quad 0.036$$

To calculate bel and pl for i.e. Kitchen and Storage:

$$bel_{Ev123}\{Kit\} = \quad 0.616$$
$$bel_{Ev123}\{Stor\} = \quad 0$$
$$bel_{Ev123}\{Kit, Stor\} = \quad 0.18 + 0.616 \quad = 0.796$$
$$pl_{Ev123}\{Kit\} = \quad 0.18 + 0.616 + 0.084 + 0.094 + 0.036 \quad = 1$$
$$pl_{Ev123}\{Stor\} = \quad 0.18 + 0.084 + 0.0361 \quad = 0.3$$
$$pl_{Ev123}\{Kit, Stor\} = \quad 0.18 + 0.616 + 0.084 + 0.094 + 0.036 \quad = 1$$

If one wanted to take a decision now, by using the Pignistic Transform (Eq. 5.30) it would give:

$$BetP_{Ev123}\{Kit\} = \quad (\tfrac{0.616}{1} + \tfrac{0.18}{2} + \tfrac{0.084}{3} + \tfrac{0.084}{3} + \tfrac{0.0.036}{5}) \cdot 1 \quad = 0.7692$$

Open World

Some other newly discovered evidence, Ev 4, suggests that:

- **Evid. 4** (lots of glass pieces): "*Most likely from wine /oil holders but it could also be perfumed oils. Or something else completely.*"

$$m_4 \mathtt{Storage, Kitchen} = 0.8; \; m_4 \mathtt{Bath, Bedroom} = 0.1; \; m_4\{\varnothing\} = 0.1$$

Note that the above evidence brings the possibility of a previously unconsidered room type. It is still not known what it could be, but this information must be taken under consideration:

		$\mathrm{Ev}_4 m_4$		
		Kit, Stor	Bath, Bed	$\varnothing$
Ev_{123}	m_{123}	0.8	0.1	0.1
Kit, Stor	0.18	Kit, Stor 0.1444	$\varnothing$ 0.018	$\varnothing$ 0.018
Kit	0.616	Kit 0.4928	$\varnothing$ 0.0616	$\varnothing$ 0.0616
Kit, Stor, Din	0.084	Kit, Stor 0.0672	$\varnothing$ 0.0084	$\varnothing$ 0.0084
Bath, Bed Kit	0.084	Kit 0.0672	Bath, Bed 0.0084	$\varnothing$ 0.0084
Ω	0.036	Kit, Stor 0.0288	Bath, Bed 0.0036	$\varnothing$ 0.0036

Table 5.5: TBM Combination of Ev_{123} and Ev_4

Using Eq. 5.27:

$$m_{\mathrm{Ev}1234}\{\mathtt{Bath, Bed}\} = 0.012 \qquad m_{\mathrm{Ev}1234}\{\mathtt{Kit, Stor}\} = 0.24$$
$$m_{\mathrm{Ev}1234}\{\mathtt{Kit}\} = 0.56 \qquad m_{\mathrm{Ev}1234}\{\varnothing\} = 0.188$$

The mass allocated at $\varnothing$ not only represents any possible Unknown Propositions but also encapsulates any conflict arising from disagreeing evidence. One can deduct that two highly contrasting pieces of evidence, after being merged, will have a high amount of conflict mass and a low amount of mass to their respective supports.

Discounting Evidence

What happens if a piece of evidence is deemed untrustworthy, or for some reason, the expert does not wish it to have a big impact on the total evidence pool? This is solved by applying *discounting*.

- **Ev. 5** "*This marble tile seems to strongly match similar tiles found in Roman baths or bedrooms.*"

$m_5\{Bath, Bed\} = 0.8 \; m_5(\Omega) = 0.2$

However, another expert suggests to You that this tile seems to have arrived from another location. He proposes that its reliability is quite low; he gives it at most 20%. This value, 0.2, is known as the *discounting factor,* or the *reliability coefficient.* Before Ev_5 is to be combined with the rest of the evidence, it will be discounted by the factor. Discounting involves transferring a portion of this mass to the Ω mass, so that:

$$m_\alpha(A) = \alpha m(A)$$
$$m_\alpha(\Omega) = 1 - a + am(\Omega)$$

Thus $m_{5disc}\{Bath, Bed\} = 0.8 * 0.2(reliability) = 0.16$ and thus $m_{5disc}\{\Omega\} = 1 - 0.2 + 0.2 * 0.2 = 0.84$. After discounting, the combination can occur:

		$Ev_{1234}m_{1234}$			
		Bath, Bed	Kit,Stor	Kit	$\varnothing$
Ev_5	m_{5disc}	0.012	0.24	0.56	0.188
Bath,Bed	0.16	Bath, Bed 0.00192	$\varnothing$ 0.0384	$\varnothing$ 0.896	$\varnothing$ 0.03008
Ω	0.84	Bath, Bed 0.01008	Kit, Stor 0.2016	Kit 0.4704	$\varnothing$ 0.15792

Table 5.6: TBM Combination of Ev_{1234} and Ev_{5disc}

Using Eq. 5.27:

$$m_{Ev12345disc}\{Bath, Bed\} = 0.012 \qquad m_{Ev12345disc}\{Kit, Stor\} = 0.2016$$
$$m_{Ev1234disc}\{Kit\} = 0.4704 \qquad m_{Ev12345disc}\{\varnothing\} = 0.316$$

Transferring Belief

Moving to the last piece of evidence, You deduct that it is extremely strong against the hypothesis of the room being a Storage room.

- **Ev. 6** (features analysis): *"The room had a fair amount of constant sunshine, most likely large windows–a storage room would have been dark."*

Thus, it has to be excluded from Ω and moved to the list of all IP. To transfer the mass of evidence beloning to Storage, a *conditioning* should be done:

$m_{6strong}\{Storage] = 0 \; m_{6strong}\{\Omega\} = 1$

This results to:

		$Ev_{1234}m_{1234}$			
		Bath, Bed	Kit,Stor	Kit	$\varnothing$
Ev_6	$m_{6strong}$	0.012	0.2016	0.4704	0.316
Stor	0	$\varnothing$	Stor	$\varnothing$	$\varnothing$
		0	0	0	0
Ω	1	$Bath, Bed$	Kit	Kit	$\varnothing$
		0.012	0.2016	0.4704	0.316

Table 5.7: TBM Combination of $Ev_{12345disc}$ and $Ev_{6strong}$

$$m_{Ev12345disc6str}\{Bath, Bed\} = 0.012 \qquad m_{Ev1234disc6str}\{Kit\} = 0.672$$
$$m_{Ev12345disc6str}\{\varnothing\} = 0.316$$

These masses concluding the evidence combination for the room and describe Your credal state on Ω. Through the TBM–still at the Credal Level–your belief indicates that the room was most likely a Kitchen. There is still the possibility that it is something else ($\varnothing$). That, or the evidence needs to be re-evaluated and perhaps some of it discounted if is appropriate. If You wish to make a decision between the current options (Bed/Bath/Kitchen), the TBM will need to operate on the pignistic level in order to normalise for the conflict. It will also need to operate on the PL in order to separate the subset mass (i.e on Bath, Bed) equally between the two.

Making a Decision

Using the Pignistic Transform (Eq. 5.30) gives:

$$BetP_{Ev123}\{Kit\} = \frac{0.672}{1} \cdot \frac{1}{1-0.316} = 0.982$$
$$BetP_{Ev123}\{Bed\} = \frac{0.012}{2} \cdot \frac{1}{1-0.316} = 0.009$$
$$BetP_{Ev123}\{Bath\} = \frac{0.012}{2} \cdot \frac{1}{1-0.316} = 0.009$$

5.4 The TBM approach to archaeological uncertainty

5.4.1 Generalised Bayesian Theorem

As has been demonstrated in Section 4.3.2, Bayes' rule allows for knowledge update and provides posterior probabilities, based on some prior belief and accummulated information. In the TBM, Bayes' rule has been extended and generalized, resulting into the Generalised Bayesian Theorem. Instead of a probability distribution, one builds a belief function over the hypothesis H given observed evidence e based on

the knowledge of the belief function over E given each $h_i \subseteq H$ and a *vacuous* a priori belief over H. The use of a vacuous prior solves a number of issues that arise, most particularly where no information actually exists on prior values. By using a vacuous prior, one chooses to accept ignorance about any prior values for Ω.

After observing evidence e, the GBT requires the plausibilities for e under all possible hypotheses (ie all $h_i \subseteq H$): $pl^E[h_i](e)$ for all $h_i \subseteq H$. In concept, this is similar to the possibilistic and probabilistic scenario where Poss/Prob(E | H) was the likelihood of the evidence given the hypothesis. For simplicity, the notation followed by [213] is used, and $l(e|h_i)$ representes the plausibility, or likelihood, of E|H. For the GBT, Smets has proved [187, 214]:

Given all likelihoods $l(e|h_i)$ for all $h_i \subseteq H$, then for $e \subseteq E$ and for $A \subseteq H$:

$$m^H[e]\{A\} \;=\; \prod_{h_i \subseteq A} l(e|h_i) \prod_{h_i \subseteq \overline{A}} (1 - l(e|h_i)) \tag{5.31}$$

$$bel^H[e]\{A\} \;=\; \prod_{h_i \subseteq \overline{A}} (1 - l(e|h_i)) - \prod_{h_i \subseteq A} (1 - l(e|h_i)) \tag{5.32}$$

$$pl^H[e]\{A\} \;=\; 1 - \prod_{h_i \subseteq A} (1 - l(e|h_i)) \tag{5.33}$$

$$q^H[e]\{A\} \;=\; \prod_{h_i \subseteq A} l(e|h_i) \tag{5.34}$$

If, in any case, prior beliefs do exist for all $h_i \subseteq H$, represented by $m_{ev0}\{H\}$, then a combination of beliefs would be needed, using the conjuctive rule of combination (Eq.5.27)

A demonstration follows, using the same knowledge base as for the Probabilistic and Possibilistic examples.

GBT archaeological example

As with the previous examples, in an archaeological dig, a column base was discovered followed by pieces of a column top. The archaeologist has the following knowledge about columns *before* unearthing anything: Columns belong to one of two orders, M and N. These columns are distinguished usually by their bases (types A and B). Column bases are always followed by pillars and tops. The related evidence distribution can be seen in Table 5.1. The evidence distribution translates to the plausibility functions for the GBT. The archaeologist has no information (or perhaps does not wish to place belief) on prior distributions of the columns M, N; thus a vacuous

belief will be used. Using Eqs. 5.31-5.34 the table can be extended for the GBT in such a way that, for example:

$$pl^H_{M,N}\{A\} \;=\; 1 - (1 - 0.95)(1 - 0.25)$$

$$m^H_M\{A\} \;=\; 0.95 \cdot (1 - 0.25)(1 - 0)$$

$$q^H_{M,N}\{A\} \;=\; 0.95 \cdot 0.25$$

The results can be seen in Table 5.8.

Bases	{M}	{N}	{M, N}	{∅}
$pl_{h_i}\{A\}$	0.95	0.25	0.9625	0
$m_{h_i}\{A\}$	0.7125	0.0125	0.2375	0.0375
$q_{h_i}\{A\}$	0.95	0.25	0.2375	1
$pl_{h_i}\{B\}$	0.05	0.75	0.7625	0
$m_{h_i}\{B\}$	0.0125	0.7125	0.0375	0.2375
$q_{h_i}\{B\}$	0.05	0.75	0.0375	1

Pillars	{M}	{N}	{M, N}	{∅}
$pl_{h_i}\{C\}$	0.8	0.1	0.82	0
$m_{h_i}\{C\}$	0.72	0.02	0.0144	0.2456
$q_{h_i}\{C\}$	0.8	0.1	0.08	1
$pl_{h_i}\{D\}$	0.2	0.9	0.92	0
$m_{h_i}\{D\}$	0.2	0.72	0.18	0.08
$q_{h_i}\{D\}$	0.02	0.9	0.18	1

Tops	{M}	{N}	{M, N}	{∅}
$pl_{h_i}\{E\}$	0.3	0.6	0.72	0
$m_{h_i}\{E\}$	0.12	0.42	0.18	0.28
$q_{h_i}\{E\}$	0.3	0.6	0.18	1
$pl_{h_i}\{F\}$	0.7	0.4	0.82	0
$m_{h_i}\{F\}$	0.42	0.12	0.28	0.18
$q_{h_i}\{F\}$	0.7	0.4	0.28	1

Colours	{M}	{N}	{M, N}	{∅}
$pl_{h_i}\{Y\}$	0.1	0.7	0.73	0
$m_{h_i}\{Y\}$	0.03	0.63	0.07	0.27
$q_{h_i}\{Y\}$	0.1	0.7	0.07	1
$pl_{h_i}\{Z\}$	0.9	0.3	0.93	0
$m_{h_i}\{Z\}$	0.63	0.03	0.27	0.07
$q_{h_i}\{Z\}$	0.9	0.3	0.27	1

Table 5.8: The knowledge base of two column styles, M and N for the GBT.

Notice the values of 0 and 1 at the plausibility and commonality rows respectively, under the column of conflict. This represents the vacuous belief function over the hypotheses. Consider that, as with the previous examples the following combinations of evidence: ACFZ, ACFY, ADFY. By using Eqs. 5.345.285.27 with a vacuous *prior*, results in Table 5.9.

	{M}	{N}	{M, N}	{∅}
$q^H[ACFZ]\{\}$	0.4788	0.003	0.0014	1
$m^H[ACFZ]\{\}$	0.477	0.0015	0.0014	0.519
$q^H[ACFY]\{\}$	0.0535	0.007	0.0003	1
$m^H[ACFY]\{\}$	0.0528	0.0066	0.0003	0.94
$q^H[ADFY]\{\}$	0.0133	0.063	0.0008	1
$m^H[ADFY]\{\}$	0.0124	0.062	0.0008	0.92

Table 5.9: Combination of evidence with the GBT using a vacuous prior.

Recollect that evidence ACFZ is in favour of Column M, while ACFY and ADFY tend to be contradicting, with ADFY being the most contradicting. This is also represented in the mass of the conflict; ACFZ has much less conflict than the other two thus indicating that the evidence is corroborative rather than contradicting.

This is a very interesting feature of the GBT/TBM as, at any point in time, it is possible to gauge how much conflict, or perhaps support for another interpretation, exists in the evidence itself. Neither the possibilistic, nor the probabilistic approach offer this feature. For completeness, Table 5.10 shows an updated evidence combination with a $\{0.5, 0.5\}$ prior.

	$\{M\}$	$\{N\}$	$\{M, N\}$	$\{\varnothing\}$
$q^H[\text{ACFZ}]\{\}$	0.2394	0.0015	0.0003	1
$m^H[\text{ACFZ}]\{\}$	0.2390	0.0011	0.0003	0.76
$q^H[\text{ACFY}]\{\}$	0.0266	0.003	0.0003	1
$m^H[\text{ACFY}]\{\}$	0.0265	0.0035	0.00009	0.97
$q^H[\text{ADFY}]\{\}$	0.0066	0.00315	0.0002	1
$m^H[\text{ADFY}]\{\}$	0.0064	0.00313	0.0002	0.96

Table 5.10: Combination of evidence with the GBT using a prior of $\{0.5, 0.5\}$.

As with the previous example (room excavation), in order to reach a decision about which of the hypotheses is the most plausible one, the GBT needs to operate on the Pignistic Level. Table 5.11 shows the pignistic calculations for the vacuous and the $\{0.5, 0.5\}$ prior respectively. Notice that the results of the latter are the same as the ones that can be calculated with a Subjective Probability approach.

Vacuous	$\{M\}$	$\{N\}$
BetP(H|ACFZ)	0.9953	0.0047
BetP(H|ACFY)	0.8861	0.1139
BetP(H|ADFY)	0.171	0.829
$\{0.5, 0.5\}$		
BetP(H|ACFZ)	0.9945	0.0055
BetP(H|ACFY)	0.885	0.115
BetP(H|ADFY)	0.1725	0.8275

Table 5.11: Pignistic transformation with the GBT with a vacuous and a non-vacuous prior

5.5   Conclusion

This Chapter has explored non-probabilistic approaches to uncertainty by explaining Possibility Theory and the Transferable Belief Model. It also presented the application [92] of these two models to archaeological examples. Chapter 4, which discussed the application of subjective probability, highlighted a few shortcomings. There are benefits to using a non-probabilistic model as some of those shortcomings can be tackled. These are now explained in more detail.

Prior belief and ignorance

Both Possibility Theory and the TBM offer ways to express total ignorance about prior beliefs. This removes a serious constraint that subjective probability faces.

Strict totality of distribution

Possibility Theory does not inherit the additive axiom of probability. Thus, among two statemens, A and B, if A is quite possible, B could be quite possible as well. One does not negate the other. Balance is achieved through the dual function of necessity.

The effect of 0 likelihood

In the TBM, the assignment of a mass function of 0 to a proposition does not nullify the proposition. It simply states that for the particular evidence, there is no weight to be assigned to that hypothesis, and whatever mass exists will either go to other hypotheses or the world. As a result, 0 means *no evidence at all for*, rather than *it is totally improbable that*.

Forced granularity

The TBM allows for assigning masses to a group of hypotheses without granulating the evidence further (at least at the Credal Level). This is very useful in cases where evidence could support two hypotheses but the expert is not willing to specify to what extent it supports one rather than the other.

Closed world

Both subjective probability and possibility work with the concept of a closed world. The TBM allows for having an open world scenario, where it includes the possibilty of hypotheses that have not been consider or are still unknown. A good analogy to

this is in a medical scenario where mass is assigned to still unknown or undiscovered diseases.

Discounting

Through the feature for discounting of evidence, the TBM allows for representation of evidence weight. This is very important as it could be extremely useful in scenarios such as those encountered through the archaeological questionnaire, in Chapter 3. Claiming one type of evidence as unreliable in interpreting a hypothesis, will not transfer the ratio equally to other hypotheses. Rather it gets included in the mass of the world of possible hypotheses encapsulating the fact that this evidence is not strong enough.

Conflict

The capability of the TBM to measure the conflict between evidence at the Credal Level is considered to be quite useful. There may be situations where the expert needs be aware of it in order to potentially examine the source of it or to visualise it.

Closing, it should be stressed that there is no right or wrong choice between the three models. It all depends on the requirements of the uncertainty representation. If someone uses an evidence table that conforms to a subjective probability model, it will work for both the possibility and GBT approaches. However, an approach that relies heavily on the Credal Level of the TBM would not be reducible to the subjective probability model unless the information is normalised (through the Pignistic Transformation).

The following Chapter explores how the different models behave when using a shared evidence table. Additionally it examines and disputes an uncertainty modelling approach based on fuzzy logic. Lastly, it demonstrates and discusses uncertainty visualisation examples using the three mathematical models.

To what a degree, the same past,
can leave different marks–and
especially admit of different
interpretations?

André Gide

CHAPTER 6

Modelling Uncertainty

6.1 Introduction

The previous two Chapters examined three different ways of uncertainty quantification and presented their application to archaeological scenarios. This Chapter describes innovative contributions to the modelling and visualisation of uncertainty in archaeological reconstructions and is composed of three parts.

The first part critically examines and disputes a previous approach on modelling and representation of uncertainty in archaeological reconstructions involving fuzzy logic [101, 215] and briefly discussed in Section 2.2 [92].

The second part of this Chapter brings together the three mathematical models under an experiment that explores the correlation of their results given a common evidence pool. The experiment also asks questions such as the assignment of values to a hypothesis.

The third part demonstrates a case study visualisation [91, 90]of an existing Romano-British building with had extremely limited available evidence. Different archaeological uncertainty visualisations, using information visualisation schemes and the three uncertainty quantification models, are presented and discussed.

6.2 Previous modelling approaches

In work introduced by Niccolucci and Hermon in 2004 [101, 215], the authors propose a numeric definition of the reliability of an archaeological reconstruction based

on fuzzy logic. The first part of their work discounts the use of a probabilistic approach, while the second one proposes a solution based on fuzzy logic. This section of the thesis demonstrates that a probability solution is actually capable of representing uncertainty, and that a fuzzy logic approach is not suitable for this type of uncertainty modelling. Firstly, the probability section is analysed.

6.2.1 Discounting of probability

The authors define the reliability index of a model M belonging to a universe U as a function $r : U \rightarrow \{0, 1\}$. For every model $M \epsilon U$ a reliability $r_M = r(M)$ is attached. The range between $\{0, 1\}$ is the model's reliability range where 0 indicates total unreliability and 1 absolute reliability. The models are distributed in a vector space X. Within universe U, an aggregation operation A is defined between i.e. two operands M_1 and M_2 in a vector space position X_1 and X_2. This results to a new model M_3 in a position X_3. In other words, $A((M_1, X_1), (M_2 X_2)) = (M_3, X_3)$.

The virtual position X_3 however, does not affect directly the reliability of a model; on the contrary, the reliability is influenced by relative position of details on each model. This is represented as $A(M_1,\ M_2\ , q) = M_3$ where q takes into account the mutual position of M_1 and M_2

Thus, the process of producing an archaeological model is a sequence of models, each obtained from the previous one, as follows $M_{k+1} = A(M_k, m_k, q_k)$ where m_k represents the details added in the last step and q_k the position of M_k and m_k.

The authors suggest that the probability of the reliability of a model M_{k+1} can be calculated as

$$Pr(M_{k+1}) = P(M_k)P(M_{k+1} \mid M_k) \tag{6.1}$$

where both the added details with the previous model as well as their position with regards to it are taken under consideration. By iterating Eq. 6.1 the probability/reliability of the model is:

$$Pr(M_{k+1}) \quad = \quad p_0 p_1 p_2 ... p_n \tag{6.2}$$

In other words, $Pr(M_{k+1}) == Pr(M_{k+1}) \cap Pr(M)$

By following this approach, a building with three parts and 0.9 reliability in each would have approximately 0.65 total reliability. As a result, the authors forego the use of probability for such archaeological scenarios and propose a fuzzy logic approach.

Section 4.4 also acknowledged that conditional probability by itself is indeed not sufficient to handle the impact added by evidence in a model. However, an efficient way to model this would have been by using a Bayesian approach, as has been demonstrated in the same Section. Thus, contrary to the suggestion of [101, 215], probabilistic models are capable of representing archaeological uncertainty. Limitations exist, as highlighted in Section 4.5, but subjective probability can be used, and has been used extensively by a number of other disciplines, to quantify uncertainty.

6.2.2   Fuzzy logic and reliability

The authors also propose a fuzzy logic approach to estimating the overall reliability of an archaeological model. By using the fuzzy logic mathematical foundation introduced in Section 5.1.2, a fuzzy truth-value scheme is proposed for treating the models. More analytically:

A fuzzy truth function f aggregating two model parts, A and B is defined as:

$$f(A \cup B) = \min(f(A), \, f(B)) \tag{6.3}$$

Three reliability indices are proposed: $r^{(a)}$ which stands for absolute reliability, $r^{(r)}$ which represents relative reliability and $r^{(r)}$ which represents positional reliability. The first one takes under consideration the actual reliability of the object itself, the second examines the reliability of the object with the previously chosen details, while the third one represents the reliability of the object's position. In order to calculate a model step reliability:

$$r(M_{k+1}) = \min(r(M_k), \, r^{(a)}(m_k), \, r^{(r)}(m_k), \, r^{(p)}(m_k))$$

The above equation derives the minimum of the reliabilities (relative, absolute and positional) and assigns it as an overall reliability of a specific part. Consider a column's base, with {0.8, 0.6, 0.9} for the relative, absolute and positional reliabilities. Even though the position as well as the context with respect to the other objects are fairly reliable, the absolute reliability is less, perhaps because of fragmented information. This leads to an overall value of 0.6 for the reliability of this specific part.

Calculating the overall reliability of a model, M consists of either applying recursively the previous formula, or being computed step by step for each model part added. As a result, the overall reliability of a model would equal the worst one of its
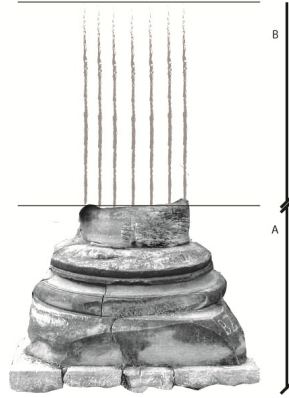
Figure 6.1: Fuzzy logic approach example

parts. A simple example is presented to illustrate the authors' approach, by using a column as a case study:

Consider the reconstruction of a column, M, of which only the base and parts of the head have been found. Figure 6.1 illustrates the first step of the reconstruction. Part A is the recovered part, with a reliability of 1 ($r_0 = 1$), which results to the first step, M_0. As there is no evidence for the pillar, it has to be reconstructed. Thus, a pillar is added, with an unknown height h_1, identified as Part B. The new model, composed of parts A and B is M_1.

In order to compute the reliability of M_1, the different reliabilities are computed for part B. The compatibility with the previous shape and the positional reliability are assigned the number of 1. The absolute reliability, however, depends on the unknown height of B. The authors, then, assign a fuzzy membership function ranging from min. height of B $<$ B $<$ max. height of B. This is illustrated with a fuzzy membership function such as in Figure 6.2.

The next step would be to add the column top, illustrated in Figure 6.3 as part C. Let us consider that the fragment found gives an absolute reliability that the top existed, to an extent of 0.9. Then, the following could also be considered:

- The height of the column top, with a range:

$$\text{min. height of C} < C < \text{max. height of C}$$

This is once more, represented by a membership function, as in Figure 6.2.

- The decorations on the column bust with:

$$r(dec_{none}) = 0.2, \ r(dec_{twirl}) = 0.8$$
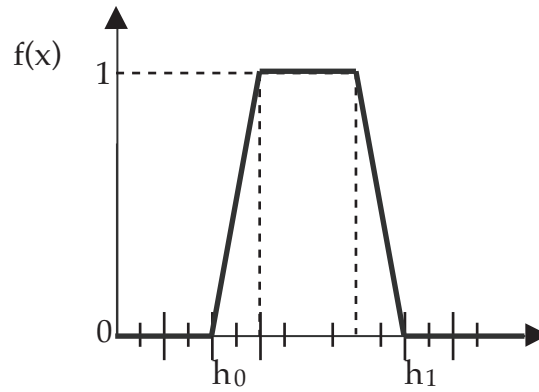
Figure 6.2: Membership function for height of Part B

For the current interpretation, the latter decoration option is chosen, with a reliability of 0.8.
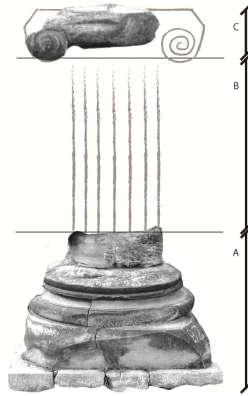
Figure 6.3: Fuzzy logic approach example with full parts.

By examining the reliabilities of the individual parts, it is deduced that: $r_{(A)} = 1$, $r_{(B)} = 1$, $r_{(C)} = 0.8$. Thus, the overall reliability for the current interpretation is:

$$
\begin{aligned}
r_{(total)} &= \min(r_{(A)}, r_{(B)}, r_{(C)}) \\
&= 0.8
\end{aligned}
$$

6.2.3   Limitations of the Fuzzy Logic approach

The fuzzy logic approach to reliability faces a number of limitations. These are analysed below:

Degrees of truth is not uncertainty

As introduced in Sections 5.1.2 and 5.1.3, fuzzy logic examines the degrees of truth in a proposition, such as "John is tall". This requires that all relative information about tallness to be available, for example, the gradual range in between where a man of John's age could be considered tall or short. In the case of a column's possible height, it is not enough to assign a fuzzy distribution of heights and arbitrarily decide all the values that satisfy 1. The simple reason for this is that the exact height of the column is not known beforehand, in order to gradually attach it to a range of interpretations. A high degree of truth assigned to a reliability interpretation, does not make this statement a reality. Rather, one must think of what information is possessed by the expert/agent about this reality, and how these possibly incomplete facts could be described.

In other words, in such scenarios, it is not feasible to describe the degrees of truth of a proposition. Instead, what can be described is the degree of belief that this proposition actually holds–how well the hypothesis conforms with the facts and evidence available. If, and only if, those facts ever become complete, then the degrees of belief can result to degrees of truth.

A possible way with which such fuzzyfied values could be used would be through an information fusion model which would combine subjective beliefs with vague (fuzzy) statements. For example, consider the following expression:

I am quite sure that the wall was tall.

The first part of the statement (I am quite sure that...) carries the belief information for the sub-statement *the wall was tall*. The word tall incorporates a vague value of tallness. In the case of archaeological uncertainty, the first part, which carries the belief in the reconstruction must be there.

Information fusion systems, where different kinds of things can be measured and fused together, would be the next evolutionary step in the representation of uncertainty. A further discussion is provided in Section 8.2.

Reliability depends on the order of addition

Consider the previous example with the column's reconstruction using the fuzzy logic approach. The relative and positional reliabilities rely on the position and the contextual belonging when compared with neighbouring parts. This would mean that the order of addition of new parts in a reconstruction can influence the relative and positional reliability–a fact also acknowledged by the authors themselves. As a result, the reliability of the model as a whole will depend on the order with which parts are added. This is a serious drawback, also encountered in Certainty Factors [181], another approach to quantifying uncertainty. The order with which evidence is added through probability, possibility and the Transferable Belief Model, does not have any effect–the results are the same.

Assignment of reliability values

The reliability values assigned to each part are not sufficiently justified. There is no explanation how the values should be chosen, or if some parts are more significant than others.

Pessimistic assignment

Regardless of how many parts are added, the overall reliability of the reconstruction will always equal the minimum reliability of a single part. There is no possible way to discount the influence that a part might have.

6.3   Comparison of Subjective Probabilities, Possibilities, and Beliefs

This section is dedicated to a comparison between subjective probabilities, beliefs, and possibilities. An experiment conducted that explores correlations between Possibility, GBT and Subjective Probability, is presented.

6.3.1   Experiment design and procedure

The experiment investigates how Possibility theory and Subjective probabilities behave under different priors, given a common evidence pool. As case study, the example regarding columns M and N has been used. The evidence pool used can be found in Table 5.1. The initial prior value for $\{M, N\}$ was $\{0.01, 0.99\}$; for each case

the prior was increased by $\{+0.01, -0.01\}$ respectively, until the 99th case which had a prior of $\{0.99, 0.01\}$. Three evidence scenarios were tested:

1. Evidence fully supports interpretation M; recovered evidence is ACFZ

2. Evidence partially supports interpretation M; recovered evidence is ACFY

3. Evidence does not support either (is contradictory); recovered evidence is ADFY

The following questions are asked:

- How do the results between GBT and Bayesian calculations compare?

- How do the results between Possibility and Bayesian calculations compare?

- Is there any correlation between them? Is it significant or non-significant?

- What happens under different prior values?

- Under Possibility Theory, does it matter which hypothesis is chosen as the DH for prior values? (i.e. the requirement that at least one interpretation must be fully possible).

- What happens to results under: (A) supporting or (B) contradicting evidence scenarios?

The posterior Subjective (Bayesian) probability was calculated for each prior, and then under each evidence scenario, thus resulting to 99 values per scenario. The calculation of possibilities gives rise to a question: Since the major requirement of Possibility Theory demands that at least one of the interpretations has a Possibility of 1, how does this affect the results? In other words, consider a Possibility prior of $\{1, 0.9\}$. Perhaps the expert assigning the prior is undecided and assigns $\text{Poss} = 1$ arbitrarily. If instead, he/she had assigned $\{0.9, 1\}$ what effect would it have to the results?

To extensively explore this, two sets of priors under the three evidence types (ACFZ/ACFY/ADFY) were constructed. The first set always assumes that interpretation M is Dominant, while the second set assumes that interpretation N is. For example, if under a Bayesian scenario the prior is $\{0.1, 0.9\}$ for $\{M, N\}$, the two Possibilistic cases examined are $\{1, 0.9\}$ and $\{0.1, 1\}$ respectively.

6.3.2 Subjective Probabilities and Beliefs

The GBT, presented in section 5.4.1, was modelled in such a way that if priors would exist for the singleton of each hypothesis, it collapses to the theorem of Bayes. This was tested using the same evidence pool and priors as the Subjective Probabilities and Possibilities. At all cases the results were exactly the same as with the Subjective Probability approach. Figure 6.4 illustrates the results of both under ACFZ. To derive those results, Eq. 4.18 and Eqs. 5.31-5.34 were used for Subjective Probability and GBT, respectively. A correlation analysis was also performed (Table 6.1), indicating perfect significance ($\tau_{ACFZ} = 1.000$, $p < 0.01$).
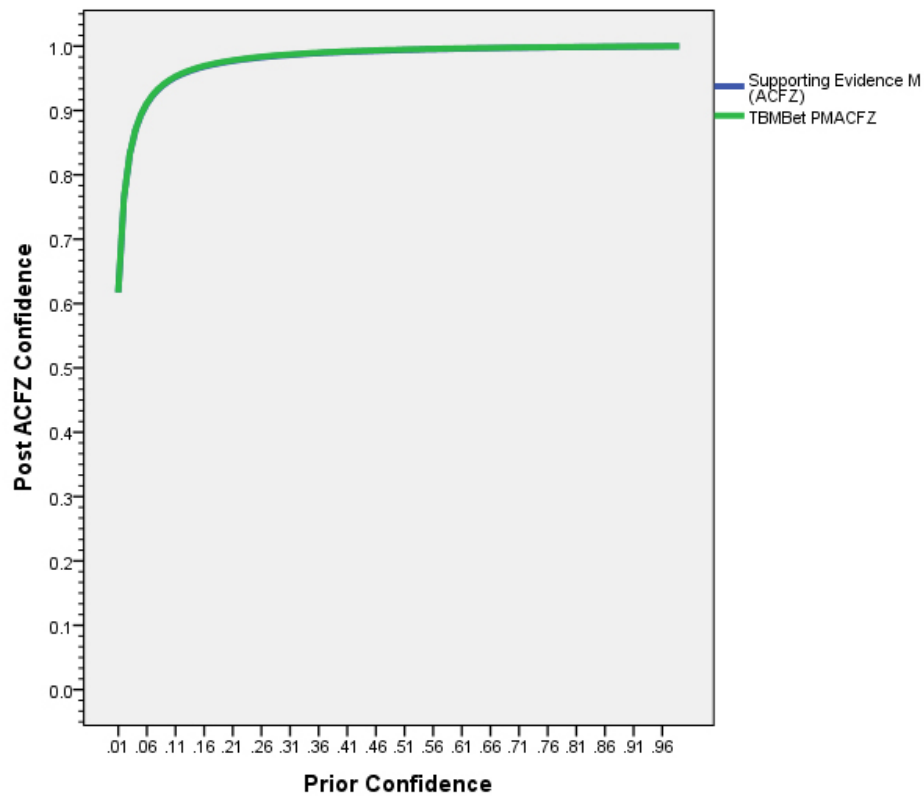
Figure 6.4: Subjective Probabilities and GBT beliefs under ACFZ using Eqs. 4.18, 5.31-5.34

Correlation Tests by Evidence Group

	Probability	GBT/TBM
Probability		
ACFZ		
Pearson	1.000	1.000
Covariance	.003	.000
Kendall's Tau	1.000	$1.000 **$
Sig. (2-tailed)		.000
Spearman's rho	1.000	$1.000 **$
Sig. (2-tailed)		.000
**. Correlation is significant at the 0.01 level (2-tailed).		

Table 6.1: Correlation tests for Subjective Probabilities and TBM/GBT

6.3.3   Analysis

Figure 6.5 illustrates the prior and posterior subjective probabilities under the different evidence. Due to the strong supporting evidence of ACFZ the confidence in an M interpretation grows sharply, even when the prior is against it. Scenario ACFY also highlights the evidence supporting M, especially after a prior of 0.2. As can be seen, it rises less sharply than ACFZ due to the fact that not all evidence supports that interpretation. Lastly, ADFY is very interesting as it shows the indecision of the particular evidence; two of those support M while the other two favour N. and a bit more strongly at that. As a result, depending on the prior, the posterior is quite different. In scenarios such as the latter, one has to accumulate more evidence.

Calculating Possibilities

Figure 6.6 represents the confidence in M under both subjective probabilities and possibilities while considering evidence ACFZ. It can be seen that for a prior of 0.38 and more the values of all three confidences converge strongly. If the prior was less, there are a number of interesting observations. The results were derived by using Eqs. 4.18 and 5.23 for the Subjective Probabilities, and Possibilities, respectively.

Firstly, at low prior levels, the evidence of ACFZ surprises the observer because it is not expected to occur (as the low prior represents that expectancy). Note that if hypothesis M is dominant, even though its certainty would be low (N $<$ 0.38) once the supporting evidence occurs, the post-confidence immediately rises due to the maxitivity axiom. On the other side, when N is dominant, the rise is much less sharp, closely following that of the Subjective Probability.
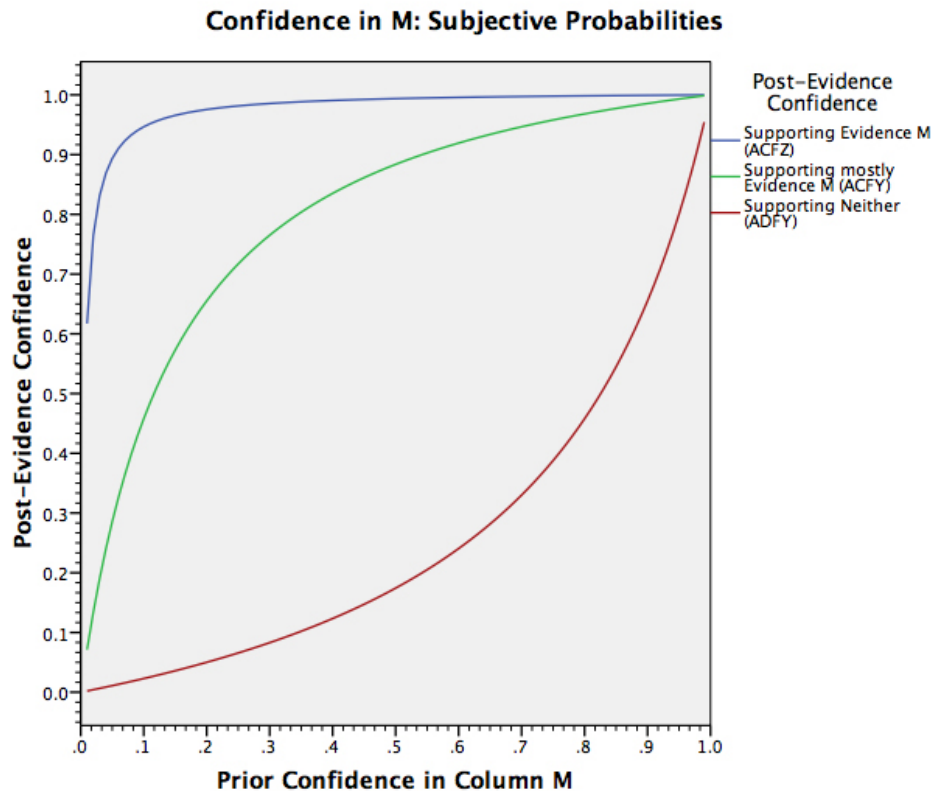
Figure 6.5: Subjective Probabilities under different evidence using Eq. 4.18

The next Figure (6.7) shows the confidence in M under evidence ACFY. If M is the DH, the post-evidence confidence in M rises quite sharply compared to the Subjective probabilities. On the other hand, if N is the DH, the Possibility follows the Subjective Probability at smaller prior values.

It can be noted that a sharp spike occurs at priors around 0.2 This is due to threshold being reached–the confidence actually has changed to an M interpretation, even though we have started with N as dominant(and with a large necessity, close to 0.8). This happens because the evidence suggests that M should now be the the DH as it is fully possible and begins to be certain.

Figure 6.8 illustrates the conflicting evidence ADFY. As with the Subjective Probability results for the conflicting evidence, the post-evidence priors are largely dependent on the pre-evidence priors. Again, a sharp spike can be seen around the prior value of 0.8 when M is the DH. This happens because the DH again shifts, but
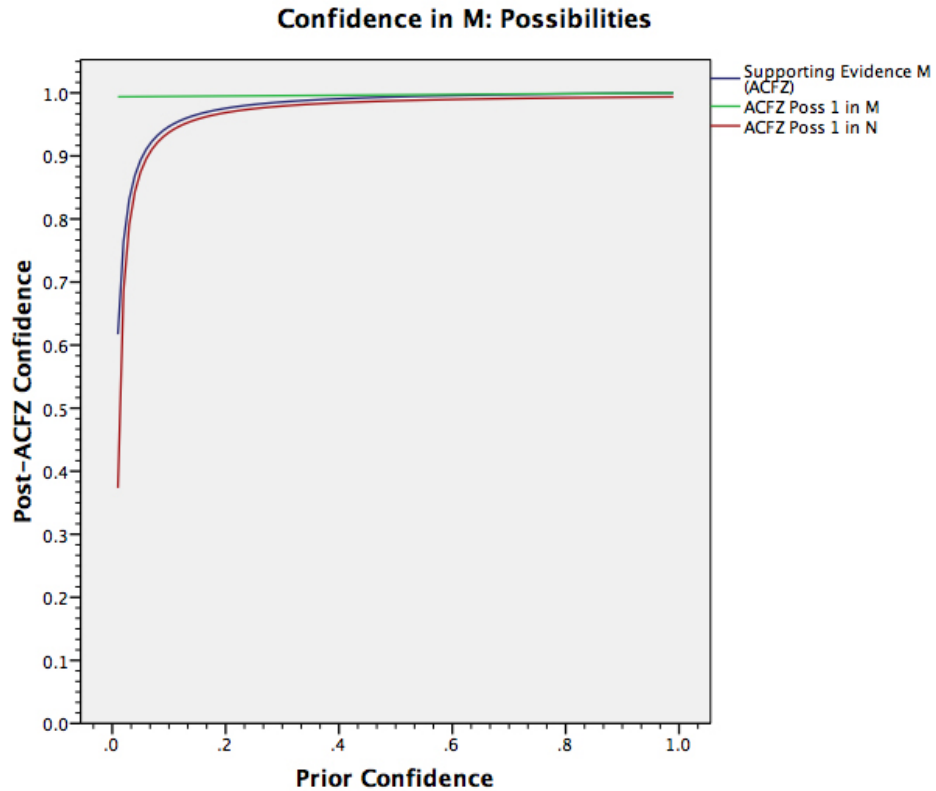
Figure 6.6: Subjective Possibilities and Probabilities under ACFZ using Eqs. 4.18 and 5.23.

that time *back to the original DH which is* M.

Consider the first prior of this scenario, it is $\{1, 0.99\}$; if we apply Eq. 5.23 the post-evidence value would be $\{0.213, 1\}$. This means that support for M under the low prior is very weak to give any certainty to this interpretation. Consecutively, the spike represents the incompleteness of the evidence; if the prior is $\{1, 0.22\}$ the posterior is $\{0.96, 1\}$. If it is $\{1, 0.21\}$ it becomes $\{1, 0.99\}$. As demonstrated in Section 5.2, this represents indecision, both are quite possible and almost nothing is certain. If the Subjective Probabilities are examined around the peak, it is worth noticing that they also lie around 0.5.

The graphs represents the *confidence* in interpretation M, not just the single possibility value of M. To elaborate, when M is the DH, the graph will show its Necessity value. When a shift happens, like from $\{0.96, 1\}$ to $\{1, 0.99\}$ it means that M has to be
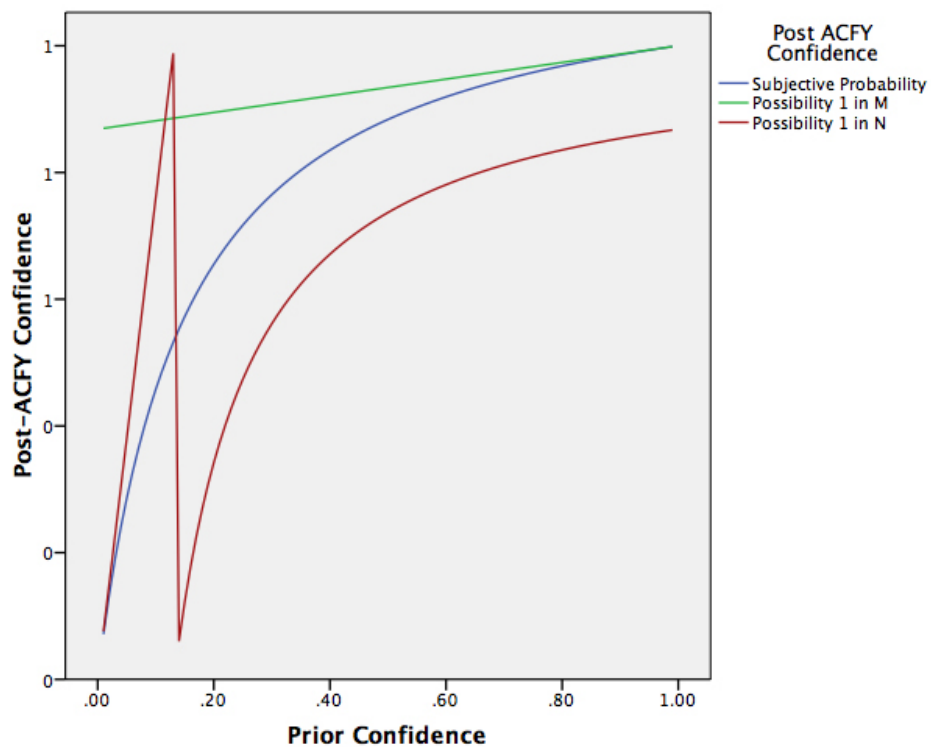
Figure 6.7: Subjective Possibilities and Probabilities under ACFY using Eqs. 4.18 and 5.23

represented by its Necessity again which at the current case is around 0.01.

The occurrence of spike events such as this cannot be predicted. When they happen, the results do deviate sharply from probabilistic and even possibilistic counterparts. Fortunately, a simple measure has been implemented that stabilises the events, further analysed in Section 6.3.4. The following section examines the correlation of the probabilistic and possibilistic values which is fundamental to establishing a relationship between the two.

6.3.4   Correlation between Subjective Probabilities and Possibilities

The data was statistically analysed to determine whether Subjective Probabilities and Possibilities are correlated. The data was tested for normality and the results indicate that data does not conform to a normal distribution and is thus non-parametric. As a result, while certain parametric tests can be used (such as Pearson's correlation coefficient) in order to calculate significance, non-parametric tests were used.
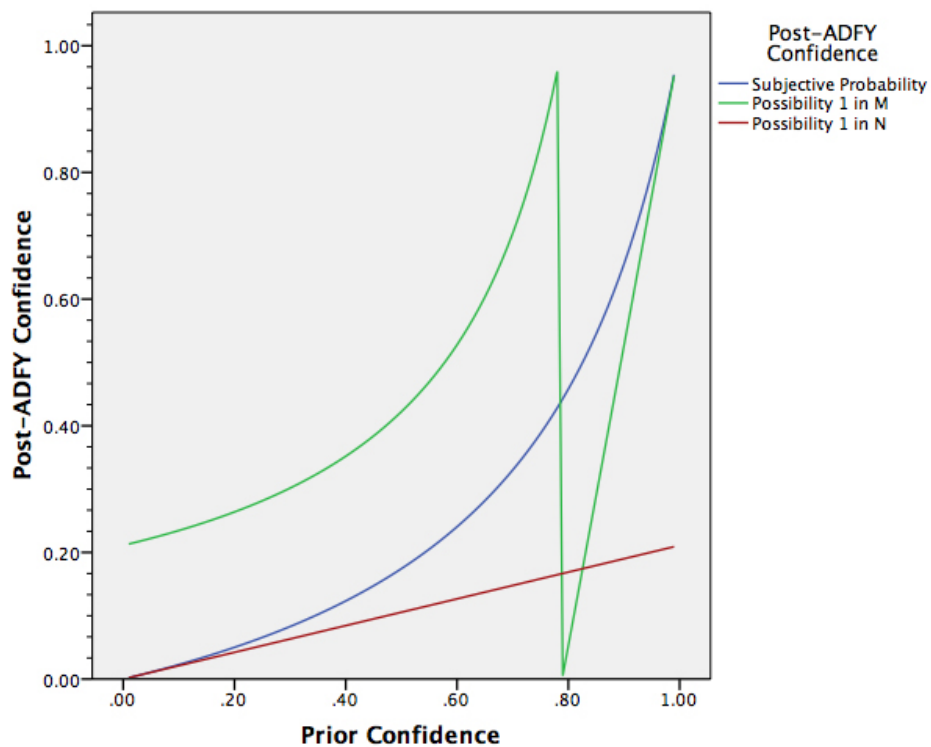
Figure 6.8: Subjective Possibilities and Probabilities under ADFY using Eqs. 4.18 and 5.23

Kolmogorov-Smirnov test (K-S) was used to test for normal distribution adherence. Table 6.2 shows the K-S results. A significant K-S value (K − S < 0.5) v indicates a deviation from normality. As can be seen from the table, all of the results are significant.

Since the data is at the interval value, this satisfies the Pearson coefficient requirement for its use. However, for significance testing Spearman's rho and Kendall's tau are used as they are fitted for non-parametric data. The test is for bivariate correlation examining whether there is a relationship between Subjective Probabilistic and Subjective Possibilistic values and thus bi-directional two-tail tests are performed. Table 6.3 shows the results of the Correlation tests.

The results show that there is a significant relationship between the Subjective Probabilities and Subjective Possibilities, $\tau_{ACFZ} = 0.000$, $p < 0.01$, $\tau_{ACFY} = 0.846$, $p < 0.01$ and $\tau_{ADFY} = 0.706$, $p < 0.01$, Even when considering the conflicting evidence scenario, the correlation still holds. It is interesting to note that on the ACFY/ADFY

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig
Probabilities						
ACFZ	.331	99	.000	.434	99	.000
ACFY	.180	99	.000	.809	99	.000
ADFY	.146	99	.000	.865	99	.000
Possibilities - $\text{Poss}(M) = 1$						
ACFZ	.061	99	.200∗	.955	99	.002
ACFY	.061	99	.200∗	.955	99	.002
ADFY	.126	99	.000	.926	99	.000
Possibilities - $\text{Poss}(N) = 1$						
ACFZ	.360	99	.000	.345	99	.000
ACFY	.164	99	.000	.847	99	.000
ADFY	.061	99	.200∗	.955	99	.002
*this is a lower bound of the true significance						

Table 6.2: K-S Test for Subjective Possibilities and Probabilities

scenarios, the possibilistic data most closely correlated with the probabilistic is the one that inhibited the threshold effect (Figures 6.7 and 6.8).

6.3.5 The case for the Ψ-measure

The previous section has demonstrated that there is a significant correlation between possibility and subjective probability values. However, the analysis of the possibility values, indicated that a series of spikes during threshold occasions can occur. Thus, it is unrealistic to use such an approach where a small shift in a prior could result into a significant difference in belief. As a result, an alternative way to represent possibilistic values had to be proposed. It is suggested that a Ψ-measure is used to represent the Possibility and Necessity of an interpretation during visual or verbal representations.

The Ψ-measure was introduced in [130]. Raufaste and daSilva Neves ran a series of experiments on medical practitioners to investigate if human judgements of uncertainty conform better to Possibility Theory than to Probability Theory. They proposed a new measurement apparatus, the Ψ-scale which allowed them to compare possibilistic vs. probabilistic values for conjunction and disjunction. The Ψ-measure can be derived in a few simple steps.

If a possibilistic confidence judgement from i.e. an expert archaeologist is retrieved, it has to conform to Possibility Theory rules. To its most basic level, that

Correlation Tests by Evidence Group

	Probability	Possibility Poss(M) = 1	Possibility Poss(N) = 1
Probability			
ACFZ			
Pearson	1.000	.562	.988
Covariance	.003	.000	.004
Kendall's Tau	1.000	1.000 ∗ ∗	1.000 ∗ ∗
Sig. (2-tailed)		.000	.000
Spearman's rho	1.000	1.000 ∗ ∗	1.000 ∗ ∗
Sig. (2-tailed)		.000	.000
ACFY			
Pearson	1.000	.879	.745
Covariance	.047	.007	.034
Kendall's Tau	1.000	1.000 ∗ ∗	.846 ∗ ∗
Sig. (2-tailed)		.000	.000
Spearman's rho	1.000	1.000 ∗ ∗	.847 ∗ ∗
Sig. (2-tailed)		.000	.000
ADFY			
Pearson	1.000	.615	.926
Covariance	.059	.032	.014
Kendall's Tau	1.000	.706 ∗ ∗	1.000 ∗ ∗
Sig. (2-tailed)		.000	.000
Spearman's rho	1.000	.690 ∗ ∗	1.000 ∗ ∗
Sig. (2-tailed)		.000	.000

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6.3: Correlation tests for Subjective Probabilities and Possibilities

means that something has to be fully possible before it has to be certain. This is also why when the Bayesian possibilistic approach (Eq. 5.23) is calculated, only Possibility values are requested since the Necessity ones are retrieved by Equations 5.16 and 5.17.

If experts are given a choice between necessity and possibility they would produce a confidence judgement using only one of the two. The suggestion by D. Dubois [130] is to use a scale combining both possibility and necessity measures. This scale would have a range from totally impossible to totally certain. This scale, the Ψ-scale, averages the results into a Ψ-measure. The Ψ-measure is:

$$\Psi(hypothesis) = \frac{1}{2}[Poss(h) + N(h)] \tag{6.4}$$

The values for Possibility and Necessity can be retrieved from a calculated Ψ-measure thus:

$$\text{If } \Psi(h) \leq {}^{1}\!/_{2} \text{ then } \text{Poss}(h) = 2\Psi(h) \quad \text{Nec}(h) = 0 \tag{6.5}$$

$$\text{If } \Psi(h) \geq {}^{1}\!/_{2} \text{ then } \text{Poss}(h) = 1 \quad \text{Nec}(h) = 2\Psi(h) - 1 \tag{6.6}$$

Total ignorance is represented when $\Psi = 1/2$ which is the point where Poss=1 and Nec=0.

In order to establish a prior to be used for the subjective possibilities, only the Possibility values for the different interpretations are required. The calculations do take care of the Necessity values accordingly, and what the expert needs to express is his/her belief in that interpretation.

As already mentioned, spikes can occur under certain priors which reflect the points when the belief shifts from one interpretation to another. Ψ-measures, however, also take the necessity value under consideration and it was deemed that they could balance out those spikes. In order to test this, Ψ-measures were calculated for all three evidence scenarios under $\text{Poss}(M) = 1$ as well as $\text{Poss}(N) = 1$ in exactly similar conditions as the previous experiment.

6.3.6   Correlation between Ψ-measures and Probabilities

The following Figures display the Ψ-measures with the relative subjective probabilities. The results were derived by using Eqs. 4.18, 5.23 and 6.4. As it can be observed, under ADFZ, the graph looks similar to that observed for possibilities/subjective possibilities (Figure 6.9). The next two graphs (6.10 and 6.11) also exhibit similar shape to the related possibilistic ones - with one major difference. There are no spikes near the threshold values, since the Ψ-measures consider both possibility and necessity *at the same time*. The correlation results signify this even further, summarised in Table 6.4.

The correlation results indicate that there is a significant relationship between the Ψ-measures and subjective probabilities. Especially for the cases of ACFY ($\text{Poss}(N) = 1$) and ADFY ($\text{Poss}(M) = 1$) which showed the spikes before ($r = 0.745$ and $r = 0.615$respectively), the correlation is even higher ($r = 0.990$ and $r = 0.993$). Similarly, the significance levels are much stronger for these two scenarios when compared to the case using the subjective possibilities.
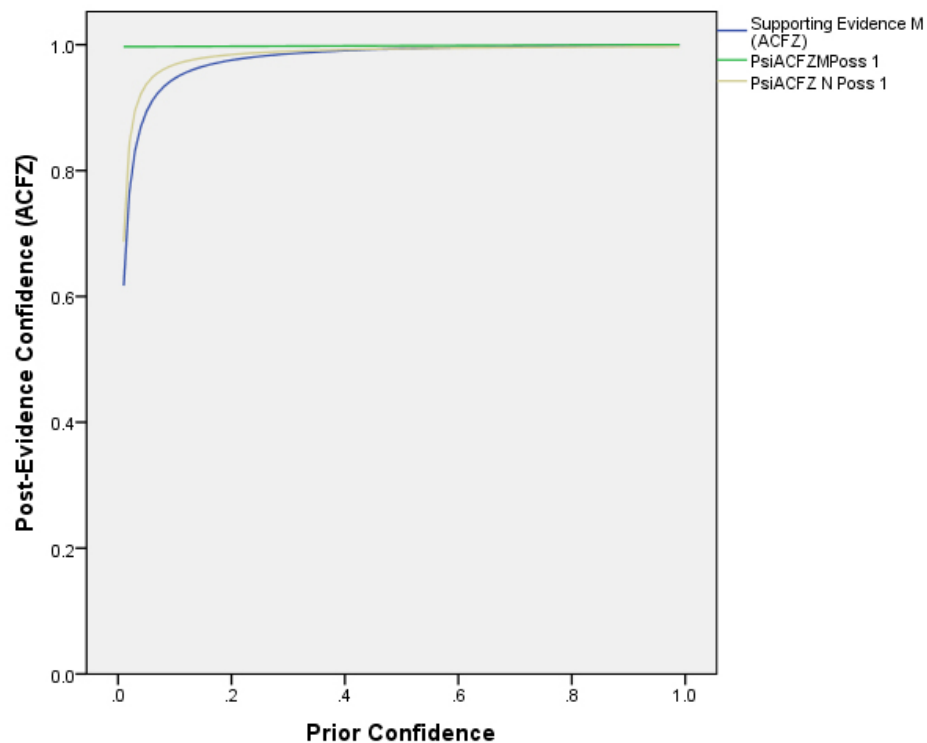
Figure 6.9: Ψ-measures and Probabilities under ACFZ using Eqs. 4.18, 5.23 and 6.4.

As a result, when using a common evidence pool, the TBM behaves exactly like the Subjective Probablity model which also allows it to inherit the similarities and differences observed with the Possibilistic ψ-value.

This section has highlighted a strong correlation between subjective probabilities and possibilities and also introduced the notion of the Ψ-measure as a representation of archaeological uncertainty. An experiment was conducted to test the behaviour of possibilistic and probabilistic values under different priors and evidence combinations. It was discovered that at certain values where the dominant attribute shifts from one hypothesis to the other, sharp spikes occur that can influence the results. The Ψ-measure provides a solution to this issue, by combining possibility and necessity into a single value. It has also been demonstrated that under Ψ-measures spikes do not occur and the relationship with probabilistic values is even more significant.

Sometimes, in both possibilistic cases–when using single values of Possibility/Necessity and when using Ψ-measures, the post-evidence results appear over- or under-confident when compared to probabilistic ones. This is due to the maxitivity axiom
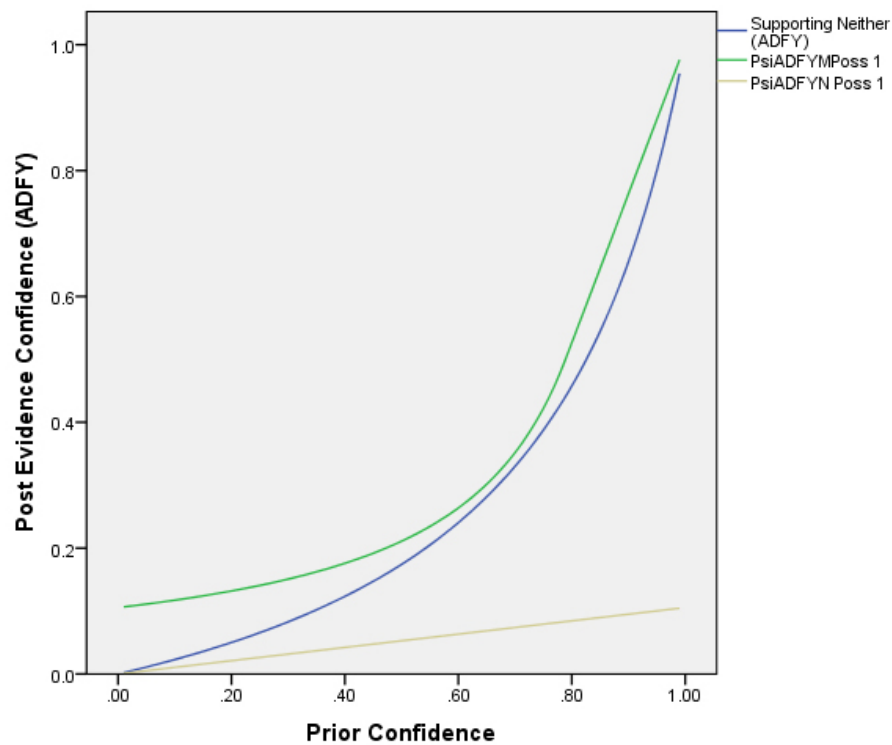
Figure 6.10: Ψ-measures and Probabilities under ADFY using Eqs. 4.18, 5.23 and 6.4.

and the dual measures that characterise Possibility Theory. If the DH at prior is supported by evidence, the posterior result will be even more confident–and vice versa. Since a significant relationship has been established between probabilistic and possibilistic values it does not matter which hypothesis is chosen as the DH–as long as the expert believes it is fully possible.
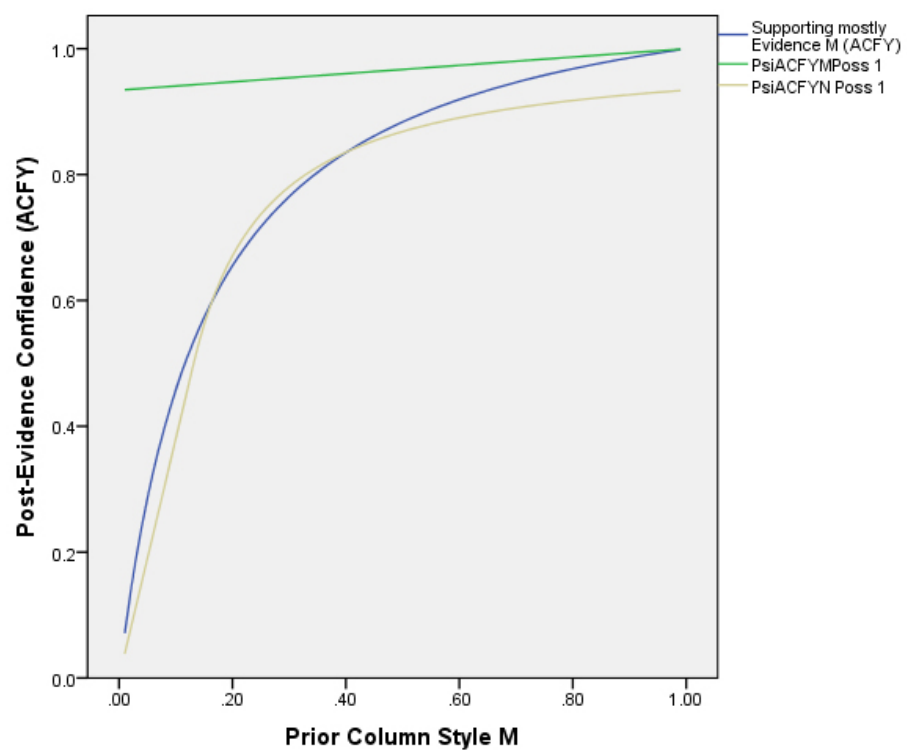
Figure 6.11: Ψ-measures and Probabilities under ACFY using Eqs. 4.18, 5.23 and 6.4.

Correlation Tests by Evidence Group

Probability	Probability	Ψ-measure $\text{Poss}(M) = 1$	Ψ-measure $\text{Poss}(N) = 1$
ACFZ			
Pearson	1.000	.562	.988
Covariance	.003	.000	.002
Kendall's Tau	1.000	1.000 $**$	1.000 $**$
Sig. (2-tailed)		.000	.000
Spearman's rho	1.000	1.000 $**$	1.000 $**$
Sig. (2-tailed)		.000	.000
ACFY			
Pearson	1.000	.879	.990
Covariance	.047	.004	.048
Kendall's Tau	1.000	1.000 $**$	1.000 $**$
Sig. (2-tailed)		.000	.000
Spearman's rho	1.000	1.000 $**$	1.000 $**$
Sig. (2-tailed)		.000	.000
ADFY			
Pearson	1.000	.993	.926
Covariance	.059	.059	.007
Kendall's Tau	1.000	1.000 $**$	1.000 $**$
Sig. (2-tailed)		.000	.000
Spearman's rho	1.000	1.000 $**$	1.000 $**$
Sig. (2-tailed)		.000	.000
**. Correlation is significant at the 0.01 level (2-tailed).			

Table 6.4: Correlation tests for Subjective Probabilities and Ψ-measures

6.4   Visualisations

A Romano-British building excavated in Fishbourne, East Sussex, was chosen as the case study. Building III, is a rectangular structure located just outside the grounds of Fishbourne Roman Palace. Fishbourne Roman Palace is one of the few palaces of that period in Britain which makes it a very significant and important monument.

Building III itself, was unearthed during excavations undertaken by the Sussex Archaeological Society between 1995—1999. The excavation report, published in 2005 [216], highlights both its simple structural form and the absence of substantial evidence which made the building's interpretation all the more difficult. Little more was discovered than the building's stone wall foundations.

Building III (Figure 6.12) represented an ideal candidate to use as a case study for archaeological uncertainty. Despite its simple form and close proximity to a prestigious palace, the evidence (or better the lack of) gave rise to two viable interpretations and possible reconstructions.
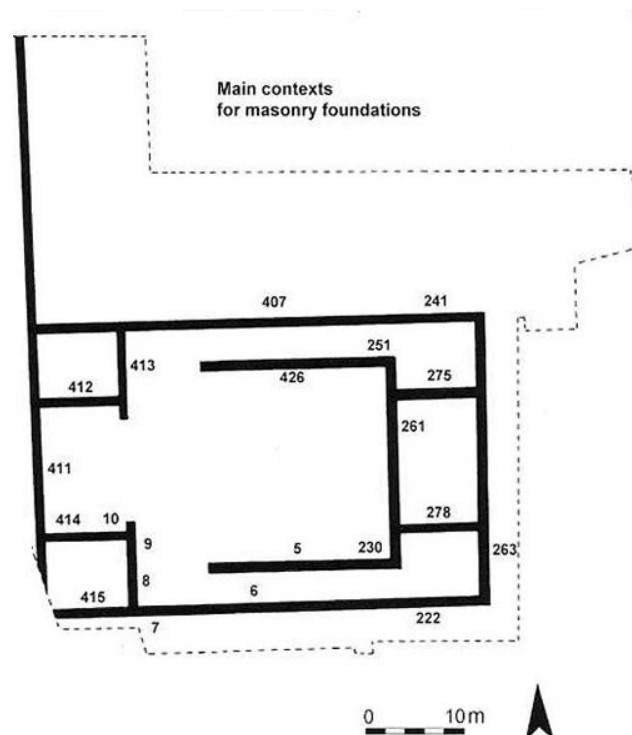
Figure 6.12: Plan of Building III (after 6.12)

A simple model of the military interpretation of the building was created in 3D

Studio Max.[1] The model was then extensively adapted in order to separate it into different structural parts such as roofs, and walls but also recovered foundations from hypothesised information. Finally, the model was exported in X3D format in order to be used by the visualisation system.

6.4.1   The visualisation system

The visualisation system prototype is based on an application (VSAC/VSAM) developed as part of a research thesis by Ben Jackson [217]. VSAC (Visual Simulation Attribute Connector) is an author-friendly, run-time configurable X3D interaction system; it utilises its own X3D/VRML prototype node, VSAM (Visual Simulation Attribute Messenger).

VSAM provides all input and output events for all data types supported by VRML/X3D browsers (Octaga, Cortona, Xj3D among others). All VSAM fields are optional at design time; the fields can be configured dynamically at runtime if required. When used in conjunction with VSAC, the VSAM nodes are self-describing supporting the automated construction of the functional properties of graphical assets during model/scene authoring.

The benefits of using VSAC/VSAM include abstraction from specific simulation scenarios which enables reuse of models and simulation assets. The VSAC/VSAM components and services provide socket-based remote control of the graphical representations, and can be used in single user mode or extended to support multi-user collaboration. The VSAC/VSAM interfaces can control all VRML/X3D attribute values including script functions; this means that the remote controls can add or replace the existing scene content as well as affecting simpler attribute values such as colour or shader values. The services support dynamic construction of control interfaces for scene components and can automate the process of adding controls for changing scene content. VSAM nodes can support all input and output events types included in the X3D specification. VSAM fields are optional at design time and can be configured at runtime if required.

For the purposes of the case study, the use of VSAC/VSAM allowed for changing objects, textures and objects' attributes in real-time without the need to manipulate the model. Figure 6.13 illustrates the prototype design.

The system is divided in two interacting components: control (A) and visualisa-

---

[1]The author would like to gratefully acknowledge assistance from Masters student, George Fassal, who created the first version of the model.
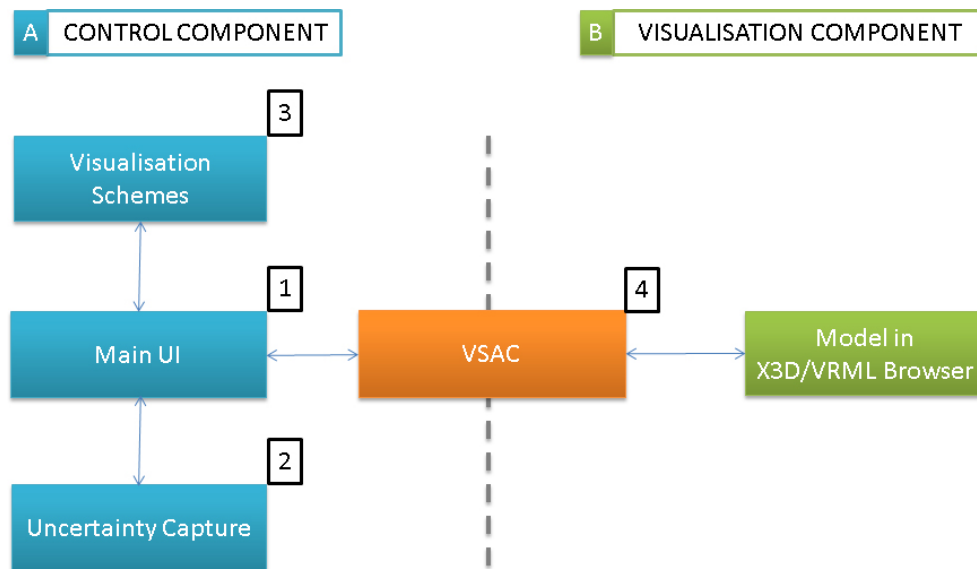
Figure 6.13: Diagram of the system

tion (B). The control component allows the user to select different parts of the model through a user interface (1) and assign uncertainty values (2). The visualisation and communication services are supported by VSAC. Once all the values have been assigned, the user can choose the visualisation scheme to apply (3).

Two different scheme types have been tested on the case study. The first type allows for graphical variable changes on the model, such as hue, transparency, opacity, etc. The second type tested the ability to use X3D shaders [218], which support real-time advanced adjustment of textures, opacity, colour, position and direction of the object.

The capability to use X3D shader nodes allows for embedding GLSL/HLSL shaders in the scene, and as a result swap between different visualisation schemes. The use of shaders opens up possibilities of visualisations such as a combination of texture maps with varying colour overlays.

Finally, the visualisation component displays the model and supports user navigation and interaction with the scene. The system provides bidirectional information transfer between the control component and the scene.

As a result, objects in the scene are identified through the prototype and can be manipulated in real-time. Figure 6.14 shows the building; the visualisation window is to the left and the controller to the right.
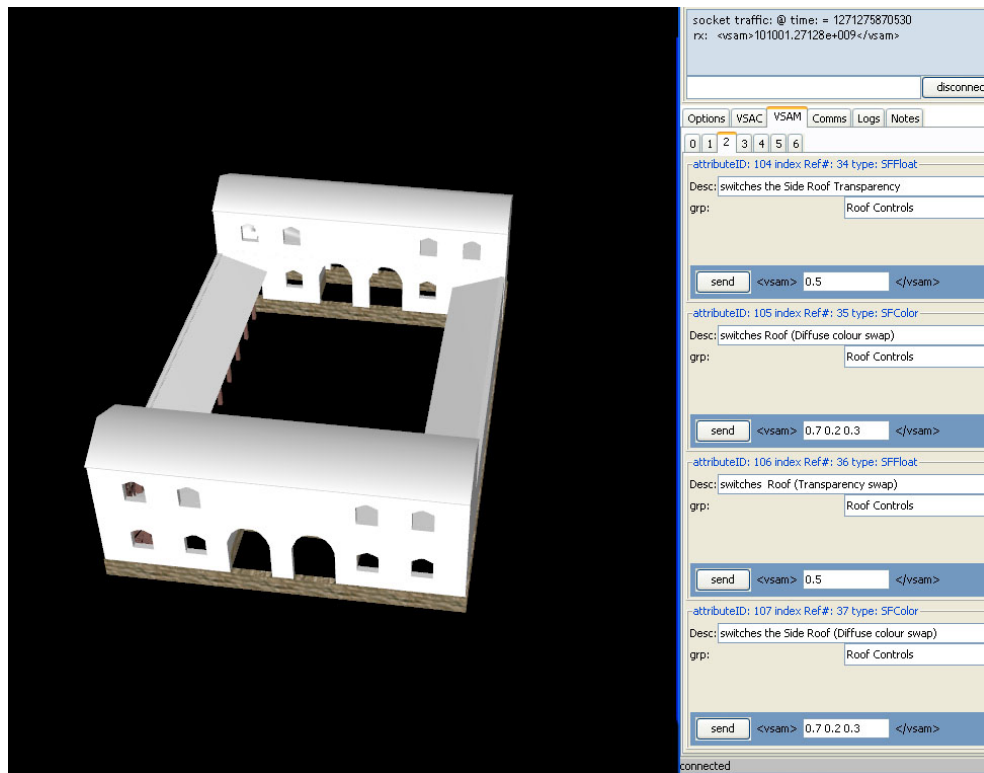
Figure 6.14: System screenshot

6.4.2 Visualising uncertainty in Building III

As introduced in 2.5, colour visualisation schemes are of importance when visually interpreting uncertainty levels, especially ordinal perceptually-ordered pseudocolour sequences. Pseudocolouring is a technique for representing varying values using a sequence of colours.

In archaeology it is mostly used in Geographical Information Systems (GIS) representation, as discussed in Chapter 2. Ordinal perception means that colour grading follows an order (black to white, hue/saturation increase, etc.); the crucial requirement is the change towards opponent colour space. If a data value Y lies between X and Z, in an ordinal pseudocolour sequence the colours should have the same ordering scheme to allow for the visual perception of the ordering of values.

For the purposes of Building III, once the uncertainty is quantified, the results were fed to an uncertainty visualisation scheme. Figure 6.15 shows a pilot run of two columns using a combination of Green-Red and Opaque-Transparent schemes.

Extensive discussions with three expert archaeologists who have worked on Romano-
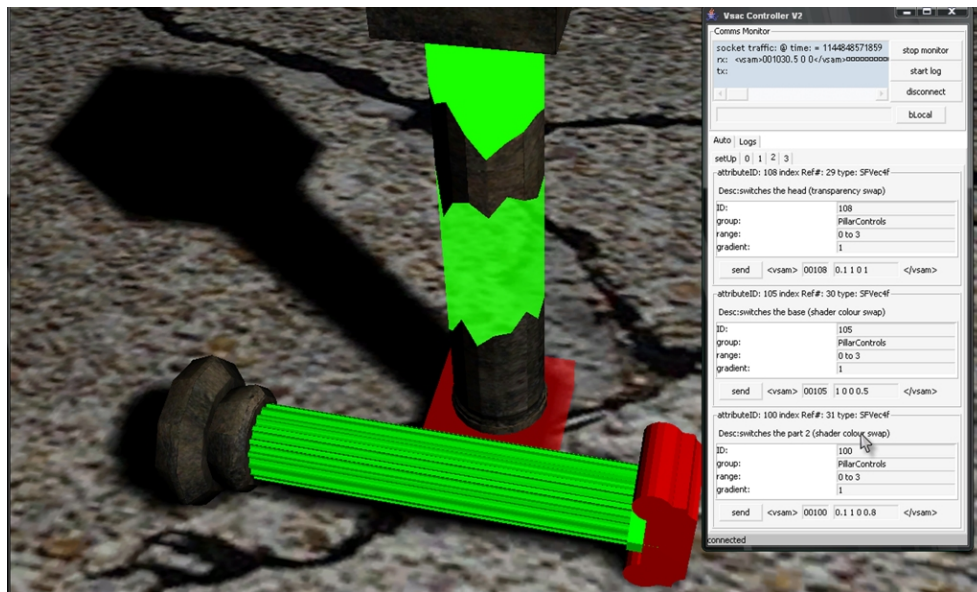
Figure 6.15: Pilot run with column

British structures were held. One of the archaeologists was the lead of the Building III excavations. Discussions were structured and answers were open-ended [89]. During these discussions the archaeologist was asked to list the available evidence in relation to each part of the building (if any), and report how confident he was concerning his subjective associations between the evidence and the suggested hypothesis.

The outcome of these discussions formed a simple knowledge table employed when performing the Bayesian, Possibility, and Belief calculations. Table 6.5 represents the resulting calculations for certain parts of the building and the possibilistic as well as probabilistic uncertainties associated with each part. The hypothesis visualised is the Military one. The higher a result turned to be, the higher the belief resulting in a Military interpretation, derived from the specific part.

The results were derived when two different priors were considered, one favouring neither hypothesis, Prior A{Religious, Military}{0.5, 0.5}and one favouring the Religious scenario Prior B{Religious, Military}{0.75, 0.25}.

	Prob	Poss	Prob	Poss
NS walls	0.636	0.71	0.37	0.43
EW walls	0.76	0.84	0.51	0.8
Roofs	0.30	0.21	0.12	0.11
Stairs	0.5	0.5	0.25	0.25
Trusses	0.72	0.81	0.47	0.66
Prior	A{0.5, 0.5}		B{0.75, 0.25}	

Table 6.5: Subjective probabilities and possibilities for Building 3 Military hypothesis

Examination of the results in Table 6.5 shows that the probabilistic and possibilistic uncertainties do not deviate more than ± 0.09 except in two cases: EW walls (0.75 prior) and Trusses (0.75 prior). In both cases, the evidence highly supports the Military hypothesis. This may suggest that in the face of extremely positive evidence, the possibilistic calculations might get overconfident (or under confident in negative) results while the probabilistic ones will not deviate that much. The following section examines the visualization of the results.

The following figures demonstrate the visualisation of the uncertainty values included in 6.5 as input in our system. Figure 6.16 shows the specific parts of Building III using a transparency visualisation under Prior A. The left figure represents the probabilistic and the right the possibilistic values. More transparency indicates more uncertainty. The difference between the probabilistic and possibilistic values for the roofs is apparent, since the possibilistic figure shows less confidence in the roof reconstruction.
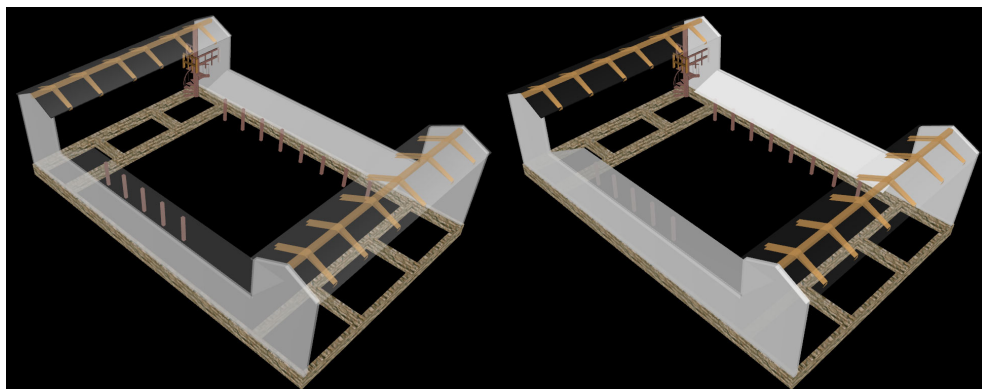
Figure 6.16: Subjective Probabilities and Possibilities with a {0.5, 0.5}prior and a transparency visualization

Figure 6.17 shows the specific parts of Building 3 for which uncertainty values were calculated using a transparency visualization under Prior B. While the uncertainty values do not vary considerably, the overconfidence of the possibilistic results for the east-west walls can be clearly observed.
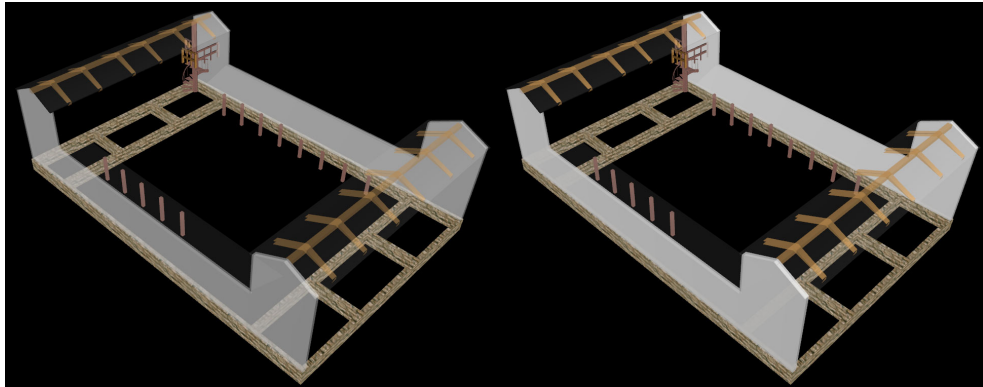
Figure 6.17: Subjective Probabilities and Possibilities with a $\{0.75, 0.25\}$prior and a transparency visualization.

Figure 6.18 shows the specific parts of Building III for which uncertainty values were calculated using a transparency and a Red-Yellow-Green visualization under Prior B where red also represents uncertainty and green certainty. The colour visualization uses a 10-value scale for equal intervals between Red and Green. This time, the difference between the cautious confidence of the probabilistic inference and the over/under confidence of the possibilistic one becomes more apparent, as the east-west wall and trusses appear more yellow, while in the possibilistic version they tend to be greener.
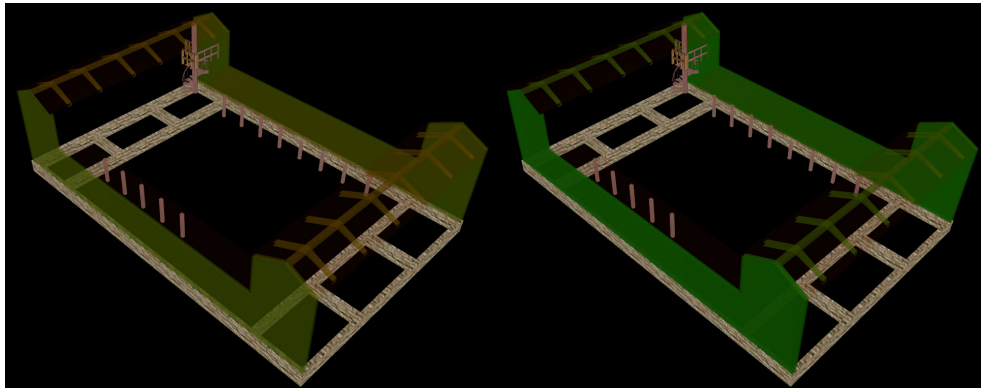
Figure 6.18: Subjective Probabilities and Possibilities with a $\{0.5, 0.5\}$prior and a transparency/colour visualization

Following, the concept of conflict, represented through the TBM, was examined. A Red-Green visualisation was used to represent conflict; red indicates more conflict among contributing evidence, while green indicates less. The uncertainty is represented with transparency; the more transparent, the more uncertain. The result is shown in Figure 6.19.
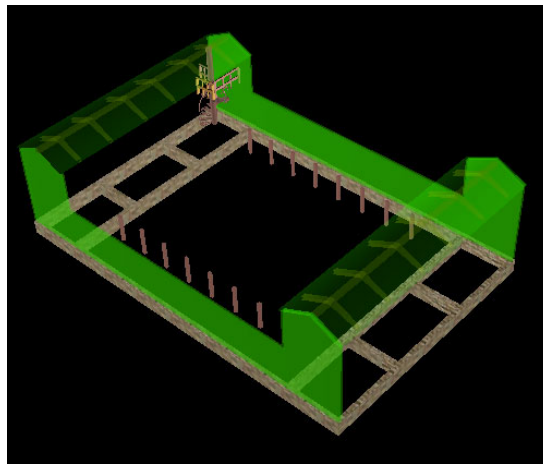
Figure 6.19: The concept of conflict in Building III

The conflict between the evidence among different parts is not very contrasting and it is difficult to show the impact that it has. In other words, because the evidence for the walls, trusses, roofs does not contradict much, the effect for both is towards the Green (low conflict) scale. If there were a highly conflicting part, it would appear like in Figure 6.20.
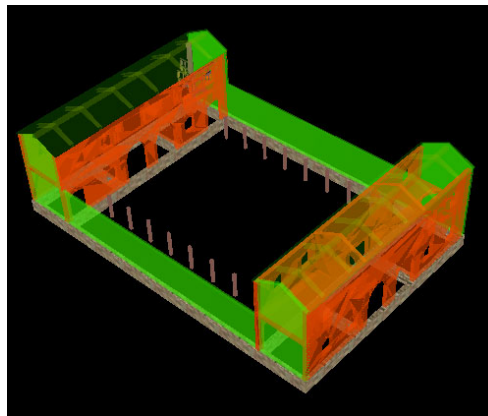
Figure 6.20: Simulated conflict in Building III

The ability to visualise the conflict amongst different evidence that makes up the certainty of an interpretation, is very useful. In the other two models, and in the TBM's pignistic level, the conflict is normalised in the result. However, there may be cases where evidence might be highly conflicting and is in need of re-examination. The option to visualise it instead of examining the numerical values would be desired. Because, as mentioned, Building III does not have highly conflicting evidence, the element of conflict was visualised by using the column example. This is the same example, first introduced in Section 4.4 on page 81, that has been used to illustrate the differences between the three mathematical models and to test for results' similarities in the simulation experiments.

6.4.3 Visualising conflict and uncertainty

Table 6.6 presents the Bet and Conflict calculations for column M and N under evidence combinations ACFZ, ACFY and ADFY. A number of different visualisations are now examined, including pseudocolouring, transparency, side-by-side, texturing and glyphs.

	Bet(M)	Bet(N)	∅
ACFZ	0.99	0.01	0.51
ACFY	0.88	0.12	0.94
ADFY	0.17	0.83	0.91

Table 6.6: Bet and Conflict results for columns M and N

Pseudocolouring

Figure 6.21 shows a visualisation of ACFZ where conflict is represented through a Red-Yellow-Green pseudocolour scheme (Green is less conflict). Belief in evidence is represented with transparency; more transparency is less belief/more uncertainty. The same scheme is followed in the visualisation of ACFY.

The conflicting evidence is extremely apparent in the ACFY interpretation at Figure 6.21; this suggests that the current interpretation is given with highly contradicting information. Similarly, the high opacity of Column M indicates that the evidence lends strong belief to that interpretation, rather than the low opacity of Column N.

Textures

In very complex scenes, or at cases of high uncertainty, the use of transparency can potentially be a serious issue. It might lead to confusion between distinguishing different areas of the model or at ignoring them because of low opacity. Precious information may be lost or obscured by other models.

As a result, visualisations that use texture as well as colour schemes are also examined. The next visualisations use a chequerboard scheme to represent conflict and a Red-Yellow-Green to represent uncertainty.

The higher the opacity of the chequerboard, the more conflict exists between the evidence (Figure 6.22). Notice the very high intensity of the checkerboard in the second set of images (ACFY); this distortion indicates a high level of conflict. Also notice the orange tint at the ACFY visualisation in the same Figure. When compared to the ACFZ, the certainty in the former is slightly stronger.

One of the useful aspects of the chequerboard is its parametrisation. One could opt to change the number of repeated rows x columns, the dimension of the squares, the shape (square, rectangle, circle), and the colours (i.e. black - white, black - transparent). Figure 6.23 illustrates a chequerboard with longer, less frequent, rectangles. Such a representation, especially in the visualisation of ACFY might be preferred in order to minimise the chequerboard's area of influence.

The next visualisation utilises a noise texture to represent conflict. The more conflict that exists between evidence, the higher the noise effect on the object. Similarly to the previous approaches, uncertainty is represented with a Red-Yellow-Green scheme. Figure 6.24 illustrates the outcome. In cases where the existing texture is also highly granulated, a noise representation of conflict would not be efficient. In

the particular case, the difference between the two levels (mid and high) of conflict are apparent. However, the high level of conflict distorts the ACFY so much that it is difficult to recognise that the level of uncertainty in column N is slightly less than ACFZ. As a result, a noise representation can be useful in cases where the conflict is considered of more importance than the evidence certainty; perhaps when identifying where high conflict exists, in order to reduce it.

Glyphs

There may be cases where the expert does not wish to affect the whole model, yet at the same time requires the uncertainty/conflict information to be available. In such scenarios, the use of glyphs is recommended instead of graphical variables. Figures 6.25, 6.26 and 6.27 visualise column M under evidence ACFZ, ACFY and ADFY respectively. Two glyphs are used, an inverted triangle to represent Uncertainty, and $\varnothing$ for conflict. Two different variations are given for the glyph scheme. The first, represented by the figures on the left hand side, fills the glyphs with a colour from a Red-Yellow-Green interval pseudocolour scheme. The closer the colour is to Red, the less the certainty and the higher the conflict. On the figures located at the right hand side, the glyphs are superimposed on a vertical axis illustrating the pseudocolour scheme. This way, the corresponding position of conflict can be compared to the extent of the spectrum and at the same time to the position of the evidence strength.

Figure 6.25 indicates a very high confidence in this interpretation; recollect that evidence ACFZ is highly supportive of column M. At the same time, it is not singularly supportive, as there is still conflict among its evidence represented by the yellowish hue of the conflict symbol. Moving on to Figure 6.26, notice that the certainty is quite less; recollect that evidence ACFY mostly lends support to column M. Similarly, the conflict is quite high here, indicated by both the filled red colour and the position at the pseudocolour bar. Lastly, ADFY is a highly conflicting set of evidence, half of it marginally supports column M and half of it strongly supports column N. This is also reflected in the visualisation in Figure 6.27 where the red triangle represents a high level of uncertainty in this interpretation, followed by an even higher amount of evidence conflict.

(a) Column M
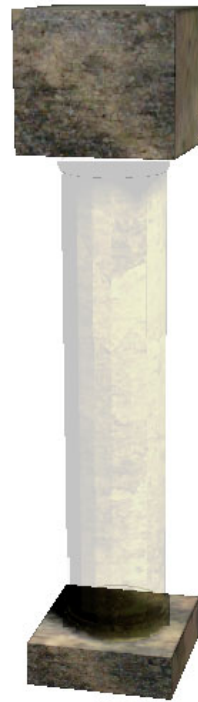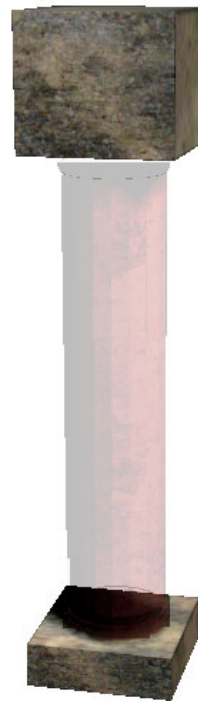
(b) Column N

(c) Column M

(d) Column N

Figure 6.21: Conflict and Evidence visualisation of ACFZ (top) and ACFY (bottom); colour is conflict, transparency is uncertainty.

(a) Column M   (b) Column N

(c) Column M   (d) Column N

Figure 6.22: Conflict and Evidence visualisation of ACFZ (top) and ACFY (bottom); colour is uncertainty, intensity of chequerboard is conflict.
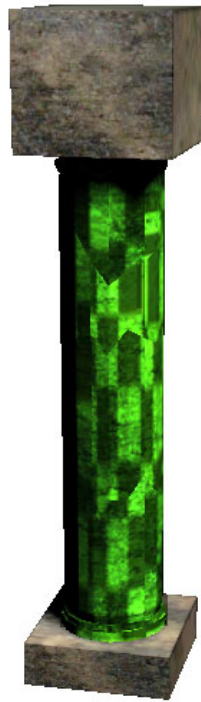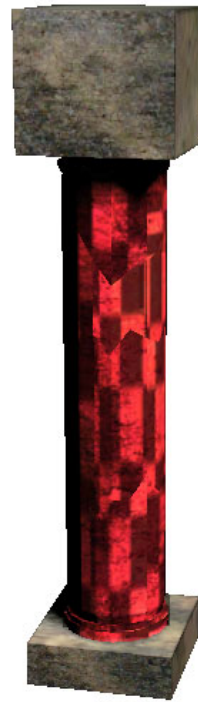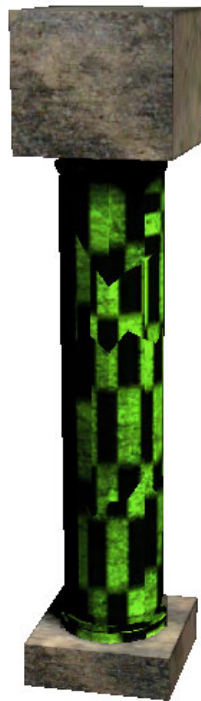
(a) Column M   (b) Column N

(c) Column M   (d) Column N

Figure 6.23: Conflict and Evidence visualisation of ACFZ (top) and ACFY (bottom); colour is uncertainty, intensity of chequerboard is conflict.

(a) Column M   (b) Column N

(c) Column M   (d) Column N

Figure 6.24: Conflict and Evidence visualisation of ACFZ (top) and ACFY (bottom); colour is uncertainty, intensity of noise is conflict.

(a) Uncertainty and conflict as coloured glyphs

(b) Uncertainty and conflict positioned on a range glyph
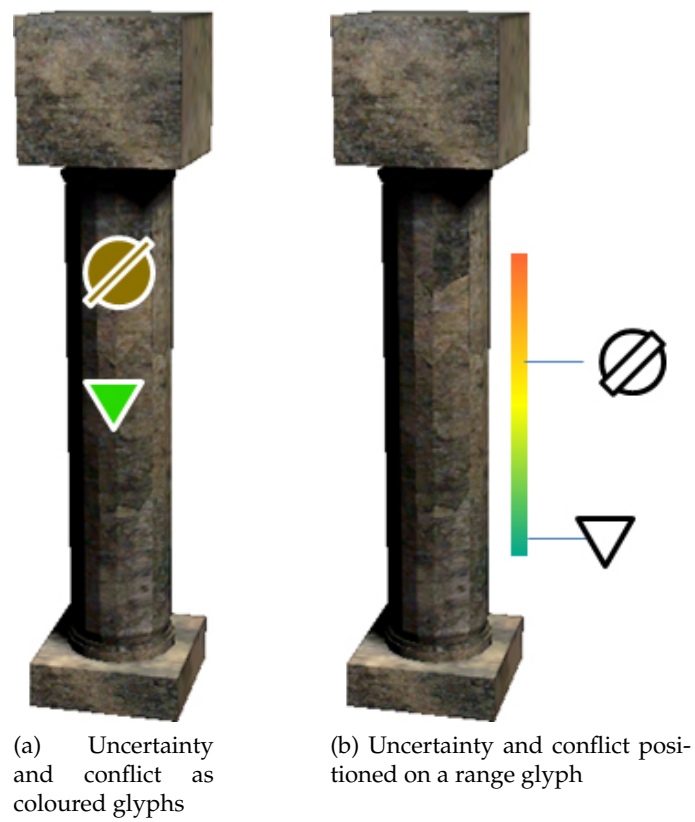
Figure 6.25: Conflict and Evidence visualisation of ACFZ for Column M; inverted triangle is uncertainty, ∅ is the conflict symbol.

(a) Uncertainty and conflict as coloured glyphs

(b) Uncertainty and conflict positioned on a range glyph
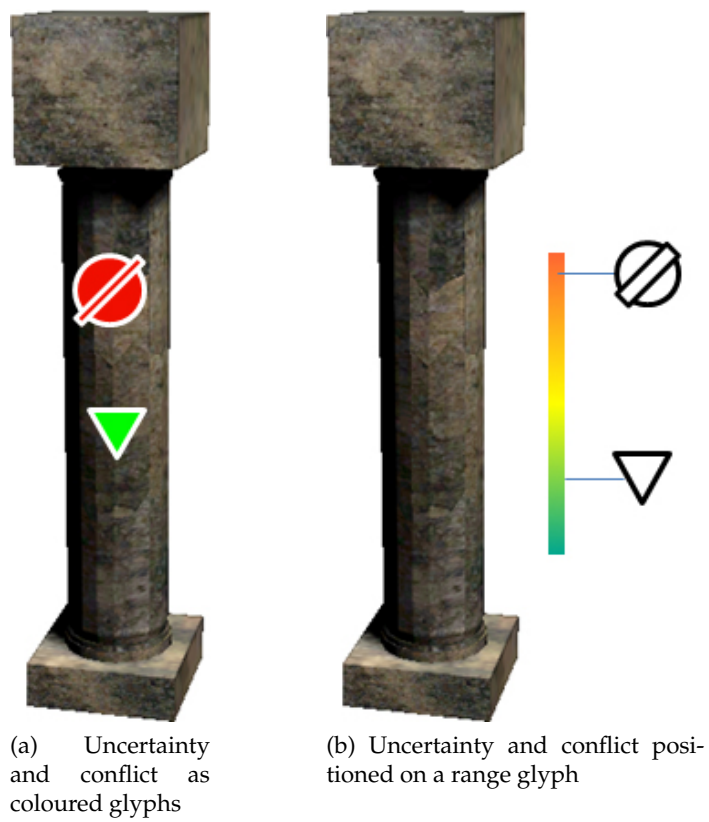
Figure 6.26: Conflict and Evidence visualisation of ACFY for Column M; inverted triangle is uncertainty, ∅ is the conflict symbol.

(a) Uncertainty and conflict as coloured glyphs

(b) Uncertainty and conflict positioned on a range glyph
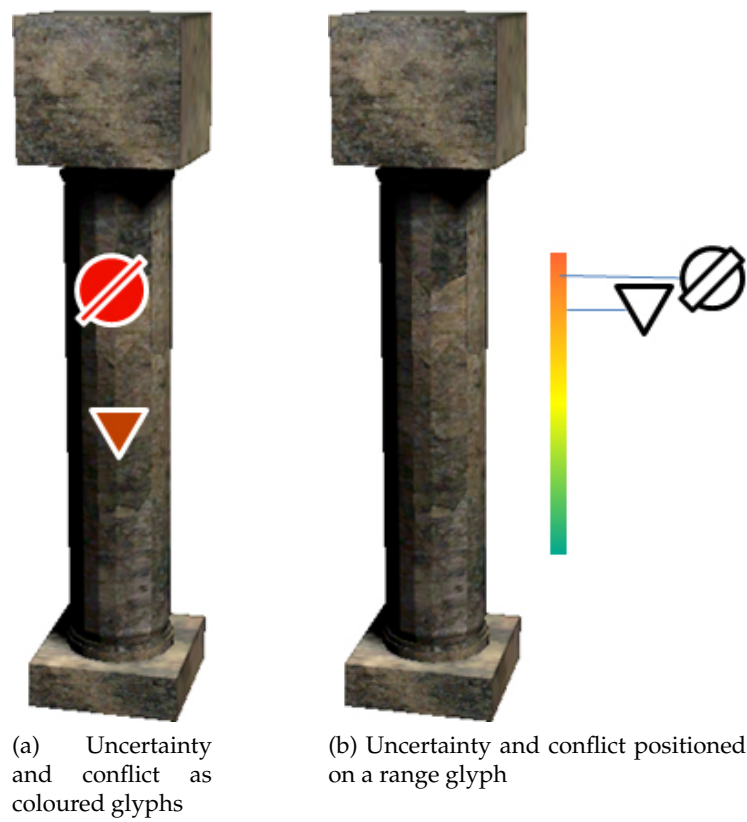
Figure 6.27: Conflict and Evidence visualisation of ADFY for Column M; inverted triangle is uncertainty, ∅ is the conflict symbol.

6.4.4 Visualisation synopsis

A number of points can be drawn from the visualisations of Building III and columns M and N. The simulations presented at Section 6.3 discovered an overconfidence and underconfidence that characterises the Ψ-measure / Possibility theory. Nonetheless, statistical comparison between the results indicated that they are significantly similar.

This over/under confidence was also made manifest in the visualisation of Building III when compared with the conservatism of the Bayesian measurement. As a result, choosing a quantification model not only makes a difference in the ways the expert can express uncertainty but also will have a direct effect in the visualisation of the reconstruction.

Secondly, the use of a TBM/GBT approach allows the expert to represent conflict, alongside with the belief in the reconstruction. The visualisations used the same quantified results to represent conflict and uncertainty through pseudocolouring, transparency, chequerboard and noise texturing and glyphs. A combination of pseudocolouring and transparency might be useful in simple scenarios without a multitude of model. High levels of transparency amongst different models might confuse a viewer.

The majority of transparency and pseudocolouring combinations used at Building III, represented one variable: uncertainty. However, it was felt that for the dual variable representation, of belief and conflict, the importance of either might be lost if represented only via transparency. As a result, different texturing approaches were also implemented, using the column examples that feature low and high values of conflict. The chequerboard texture offers a number of customisation options that make it suitable for a variety of surfaces. On the other hand, the noise texture can be too intense, overlaying uncertainty information, and also confusing when overlaying already granulated textures.

Lastly, the use of glyphs would be a good choice when one wishes for both a realistic-looking model as well as accompanying visual information on uncertainty. The glyphs' positioning at the examples were always next to the columns, but there other ways that they can be generated for example:

- Through proximity: once the viewer comes close to the particular model, the glyphs appear

- Through interaction: By touching/clicking on the specific model/part of the reconstruction

- Through animation: Fades in and out of view at specified intervals.

The choice of a visualisation scheme is also very important. If an investment is made to a transparency-heavy scheme and the model is complex, then it may be confusing to the viewer. Similarly, an influx of noise-based textures might also increase confusion. It depends on the particular model and the amount of uncertainty information that one wishes to visualise. Hopefully in the future it will be extremely easy to swap between different information visualisation schemes in order to effortlessly choose the one most appropriate for the particular model. This is also explored in Section 8.2 of this thesis.

6.5   Conclusion

In the first part of this Chapter, previous approaches to archaeological uncertainty were critically examined. It was suggested that approaches based on fuzzy logic are not appropriate for representing uncertainty based on belief. For such cases, Possibility or Plausibility/Belief models should be used. If expression of vacuous descriptions is required then the belief models can be combined with fuzzyfied expressions in an information fusion system. Additionally, the Bayes Theorem as demonstrated in relation to archaeological uncertainty in Section 4.4 suggests that, contrary to existing literature, a subjective probability approach is capable of dealing with uncertainty.

Following, an experiment conducted to compare the results between Subjective Probability, Beliefs, and Possibility Theory was discussed. Results indicate that due to the max/min axiom of Possibility theory, confidence judgements may appear higher or lower than Subjective Probability especially when conflicting evidence is involved. The differences were found to be insignificant, while a significant relationship was demonstrated between results given by Possibility and Subjective Probability.

As a result, sharp spikes in judgement which were accrued to the duality nature of Possibility Theory were successfully eliminated by using the Ψ-scale. Statistical tests between Ψ-scale and Subjective Probability results suggest a significant relationship between the two, while any deviations observed are even less significant.

Finally, the visualisation implementations were discussed. Building III was visualised in order to illustrate the different approaches to archaeological uncertainty. The visualisation was based on a system (VSAC/VSAM) which utilises X3D models and allows for the real-time altering of objects' properties. The ways to quantify uncertainty are modular and can be adapted to other systems as well.

Limitations with Building III in regards to the visualisation of conflict led to the use of the column example. The conflict and uncertainty was visualised with different visualisation schemes including pseudocolour, texture and glyphs.

The visualisations highlighted some important issues. It goes without saying that the choice of a quantification model should not be taken lightly. Even if the results are statistically similar with limited deviations (as shown in Section 6.3), the visualisation is affected. In other words it can lead to perceptual differences in visualisations, as illustrated by the Possibilistic and Probabilistic representations.

Although the visualisations are not strikingly different, the overconfidence/underconfidence of the Possibilistic model compared to the conservatism of the Bayesian do have an effect in their respective visualisations. This was highlighted further when using colour schemes instead of transparency.

The pseudocolour, texture and glyphs visualisations of uncertainty and conflict showcased different ways of visualising conflict and uncertainty. Glyphs were less invasive to the 3D model itself while texture such as noise might be considered too distracting.

Closing, one of the major difficulties a system like this faces is the discovery of relevant evidence to use as information supporting an interpretation. The visualisation system presented here is a prototype and uses a simple knowledge-base table, however it is envisioned that in the future distributed, shared, knowledge-bases would be more useful and less redundant.

As discussed in Chapter 1, a large part of the archaeological inferential process is making analogies with similar discoveries. This was also evident with the absolute and contextual comparison categories identified by the archaeologists in Chapter 3. Additionally, there may be cases where the user would prefer to display uncertainty, or conflict, above a certain level, or for only specific types of evidence. These characteristics add another element for which search would be useful.

The ability to search and discover relevant associated data is explored in the next Chapter. Chapter 7 presents research results on the design and development of a search engine with a focus on discovering ancient coinage.

His ambition is to be the spider
in the World Wide Web.

_____

Anonymous

CHAPTER 7

Searching for Information

7.1 Introduction

One of the conclusions made in the previous Chapter was that it is extremely difficult
to discover associated information in order to back up an interpretation. Knowledge
base researchers in other disciplines, such as medicine [219], have been aware of this
for a long time. Usually, knowledge in such a system is accumulated by getting in-
formation directly from experts as well as aggregating information from published
research results.

This Chapter analyses contributions and research results achieved through a search
engine tool developed to aggregate and collect information on ancient coinage. The
information was used by a European project which had as its major aim the identific-
ation of illegal electronic trade of antiquities.

The first section provides an introduction to the project and literature review on
electronic illegal trade and web image and text crawling. Section 7.3 describes the
methodology used for the design of the search engine, and the open source software
on which it relies. Section 7.4 extensively analyses the implementation aspects and the
extensions developed for adapting the search engine. Section 7.5 presents test results
of the working engine, while Section 7.6 outlines conclusions, identifies shortcomings
and explores the link between semantic search engines and uncertainty.

7.2   Searching for ancient coins

Electronic illegal trade and selling of antiquities and art works is a very important issue for authorities to handle. The COINS project was designed to provide a substantial contribution to the fight against illegal trade of ancient coins. Such trade appears to be a major part of the illegal antiques market. It was an EU funded project which involved technological, heritage and law enforcement partners.

COINS is not a single piece of software; it provides a set of tools that can assist the user with their search for ancient coins. Three interconnected areas are available: a search engine that crawls the web and retrieves candidate images, an image recognition application for matching candidate coins with stolen ones, and a management tool for archiving, indexing modifying and querying a coin image archive. These tools are able to work independently as well as collaboratively.

Figure 7.1 illustrates the design of the interconnected tools available through COINS. Section A, which involves the COINS spider (search engine) is the focus of this Chapter. The spider has the role of retrieving and saving appropriate content on its database. The content includes the HTML file, images, and any other metadata created on-the-fly in index time. The mapping tool (section B) provides the interaction between the spider and the management tool firstly by translating the retrieved metadata in a semantically meaningful way and secondly by giving the ability to add richer information to the content. Lastly, the Management tool (section C) allows for navigational and semantic search of retrieved data.

The following contributions [93] are made through the design and development of a number of novel plugins attached to the COINS spider:

- Extension to a Free and Open-Source Software (FOSS) search engine allowing it to crawl for images as well as text. It is able to retrieve and store both HTML pages and candidate coin images.

- Further extension to the FOSS search engine by taking under consideration a number of parameters related to the search of ancient coins.

- Implementation of a connector between the FOSS search engine and CBIR software, by extending the FOSS engine to also retain image data.
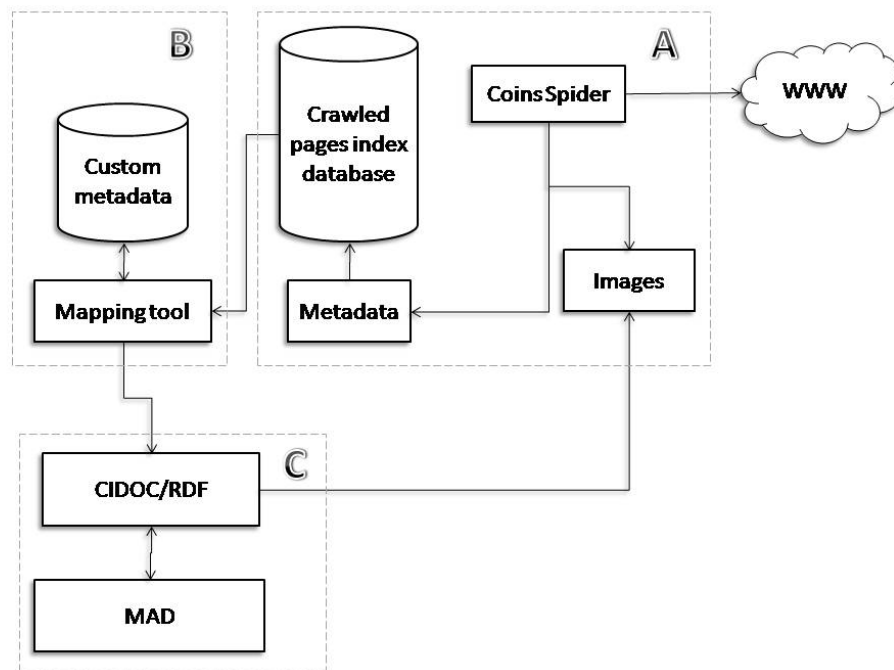
Figure 7.1: Diagram of COINS tools and data flow.

7.2.1 Electronic illegal trade

With the increasing use of the Internet, online trade in stolen goods has reached a large scale. Research by CheckMend, an online stolen property checking service indicates that roughly £5 billion worth of stolen goods are on sale at any time in the UK. Its checking service identifies two items per minute as stolen; a total of 10% of all the second hand items checked with a projected value of č100 million [220]. High profile cases of online illegal trade have surfaced in the news and retailers in the US are putting pressure on online marketplaces to safeguard the origin of the items being sold [221].

In the area of antiquities, illegal trade is also very prominent and the market is tentatively estimated as being worth around $2 billion a year; reliable data is hard to find [222]. Once more, high profile unmaskings of illicit trade have taken place, with notable example the £25million case [223] of 10,000 unprovenanced antiquities recovered by the Italian carabinieri. Illegal electronic sale of antiquities has not been actively documented in terms of volume but its existence is undoubted. In 2006 PAS (Portable Antiquities Scheme), the UK government funded scheme that records ar-

chaeological objects found by the public formed an alliance in 2006 with the online auction house, eBay, in order to curb unreported antiquities trade. In June 2008, the Swiss authorities made a deal with eBay to crack down on the illicit sale of cultural property over the internet. Facilities such as CheckMend are very useful in cases where a serial number of an item exists or at least some sort of uniquely identifying property.

On the contrary, identifying antiquities is considerably more difficult and the majority of successful discoveries until now have occurred because people chanced to look more closely at the items being sold. More appropriate to the case is trace.com, the global database of lost, stolen and seized items, which includes antiquities among others and is regularly updated by the public and law enforcement agencies. However, trace.com does not actively search for content on the Internet but rather relies on the vigilance and feedback of its users. In the case of ancient coins, research by Elkins [224] suggests that about 260,000 and 280,000 coins are sold each year on the eBay-U.S. website, not counting bulk lots.

7.2.2 Web image searching

Section 7.2.1 has illustrated that online trading of illegal antiquities is an active and ongoing problem. The COINS project is following an image-finding oriented approach. This means that it actively searches the web for coin images. Searching for images is still in its infancy stages. Quality of an engine's image results directly depends on the quality of the textual information associated with the images (e.g. filename, nearby text, *alt* tags within etc.). Advanced search abilities can display images according to colour properties (grayscale/colour) or size.

In a few cases (Google, Exalead) the engines apply face recognition technology in order to identify likely images containing faces. Content-Based Image Retrieval (CBIR) search engines are few and mostly in research stages. Examples include Web-Seek and IBM's QBIC. These engines consider the characteristic of the image itself–its shape and colours. Without the ability to examine image content, searches must solely rely on textual metadata.

Queries on a CBIR system, except from the images' characteristics, can be by example (an image is provided by the user as a search criteria), or by semantic retrieval (finding images of a specific subject). Research by Google [225] attempts to blend computer vision techniques with user feedback and preferences. Images are returned by their visual similarities by using local features descriptors. Results are encour-

aging, showing a large decrease in the number of out-of-context images returned. This also follows work done for the National Centre for Missing and Exploited Children (NCMEC). In this scenario, Google assists NCMEC in tracking child exploitation and search for patterns in images of abuse on the web [225]

7.3 Methodology

When designing the search engine, a number of factors influenced the methodology subsequently followed. These factors are better explained through two categories: the behaviour of the search engine, and the target group of users.

7.3.1 Search engine behaviour

- Whole web crawling: In order to compare candidate coin images with a list of known stolen ones, good candidate sites must first be found. It is a tremendous task to crawl the whole web and requires a huge amount of hardware and man-hours.

- Open source approach: One of the COINS project initiatives is its open source approach. This extends also to the search engine.

- Image indexing and caching: The search tool should be able to index and save images as well as text.

- Resource constraints: effort should be placed to limit the image candidates even when indexing the web; the less false candidates, the better.

- Metadata retention: any information related to a specific image should be kept as potential information to be used by the management tool

- Multilingual properties: Not all pages are written in English. The search engine should take under consideration the existence of multilingual websites.

- Index updates: How often should an index be updated? How feasible is it to remove or updated outdated content?

- Image information: How is it possible to archive data information on candidate images so as to be used in an image query?

7.3.2   Users

The COINS system is expected to be used by two different user types: numismatics experts, and law enforcement. As a result, the search engine should be able to cater both for the coin specialist and the hobbyist. The users should be able to query the crawler by keywords, retrieve results from specific sites, or input a candidate image and receive similar results.

7.3.3   Requirements analysis

The factors relating to the users and the search engine behaviour were analysed carefully before any implementation:

Whole web crawling: It was clear from the initiation of the development that deep (extensive) web crawling in the count of millions of pages would be infeasible. Initial research was focused in identifying good candidate web sites containing ancient coinage. After consultation with the heritage experts of the project the web crawling has been tested with coin websites including auction sites. As a result, the crawler begins with a set of starting URLs and fetches relevant pages back. Of course, more starting URLs can be added in the process.

Open source approach: In an effort to be as open about the code as possible, an open source search engine implementation was used. The image searching capabilities would be an extension of the engine. Four open search engines were evaluated for the purposes of the project; their summaries can be found in Table 7.1.

	Web glimpse	Htdig	Swish-e	Nutch
Key feature	suffix arrays	simplicity	metadata	rankings
Licence	nonprofit use	GPL	GPL/LGPL	apache
Active	no	no	yes	yes
Crawling	local FS	intranet	Intranet	all
Caching	no	no	no	yes
Link rank	no	no	no	yes

Table 7.1: Comparision of open source search engines

Evaluation results indicated Nutch[226] as the best candidate for the particular project. It is a complete open source search engine, implemented in Java and can be fully customised by programming plugins for it. Nutch is based on Lucene, another open source project for indexing and searching documents. Research shows that Nutch achieves high performance [227] often as good as commercial search en-

gines [228]. Its use as a search engine is quite wide spread. Notable examples include Krugle, Creative Commons and the Internet Archive. The technical details of Nutch are analysed further in the implementation section.

Image indexing and caching: By default, Nutch is not an image indexer. It works as any text-based search engine. Work such as the one by Zhang [229] introduce an image plugin able to read JPG file formats. However it has been tested on a local filesystem scenario and not by crawling web sites. The solution was to design an image indexing and caching extension to the parser component. The plugin inherits all the properties of the text-based Nutch parser and extends it for the inclusion of images. As a result, this also aids us in the case of web page updates.

Resource constraints: An obvious solution to handling resource constraints is to limit the amount of pages the crawler is allowed to index. However, this approach may erroneously exclude positives. Additionally, it would be important to exclude automatic non-candidates such as banners and logos. The crawler is not aware of what it indexes; it does not know the content of a picture. Taking these under consideration, an approach was used where candidate URLs are fully indexed, but images with size less than a certain threshold (128x128 pixels–customisable) are automatically ignored and not cached. This manages to exclude not only banners and logos but also thumbnail pictures that are too small for image comparison. Lastly, the capability was added to exclude certain types of images (such as GIF) if not wanted.

Metadata retention: By default, Nutch stores the whole textual content of a web page to its database. However, as the text is being retrieved, additional fields can be specified in the database that can dramatically improve the relationship with the management tool. Profiles of sites can be specified during the indexing and thus include coin properties such as dimensions/mint type/etc.

Multilingual properties: To extensively test for multilingual capabilities non - English URLs were included to be cached.

Index updates: Updates to the index can be done as often or as infrequently wished. Different candidate URLs can be re-indexed in varying time intervals. For example, site A can be queried for updates every 5 days while site B every one month. Initial experimentation has shown that a 25-day limit is good enough, since the majority of auction sites keep their listings data for at least one month.

Image information: As mentioned before, Nutch is not capable by default to understand raw image data. For this reason, a number of open source CBIR systems were evaluated, in order to examine which was better suited to the project. The

systems were GIFT, FIRE, img:seek, and LIRE. In the end, LIRE [230] was selected because it supports a wide number of image indexing information, is updated frequently, and shares many common characteristics with Nutch, especially the Lucene indexer.

Figure 7.2 illustrates the desired functioalities of the system through a use-case diagram.
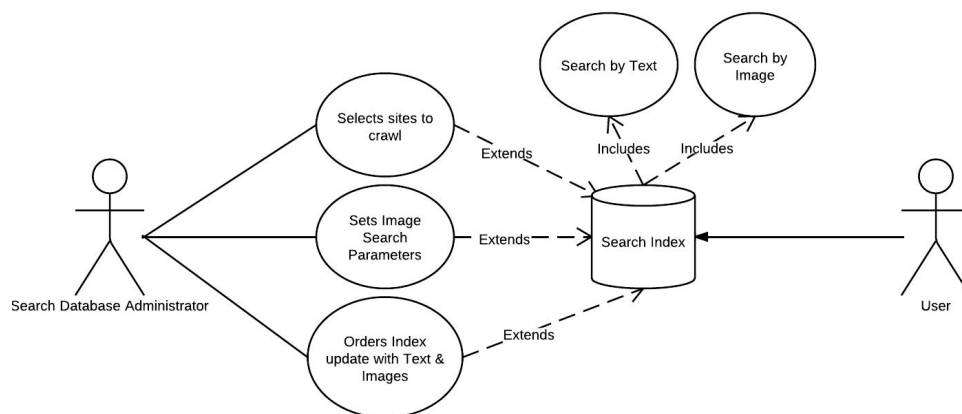
Figure 7.2: Use-case diagram for desired user actions

7.4 Implementation

Before discussing the development of the search engine, it is necessary to analyse further its basic components - Lucene, Nutch and LIRE, and how they relate to each other.

7.4.1 Lucene

Lucene [231] is a high-performance, full-featured text search engine library written entirely in Java. It is able to index and score documents according to their content. There are four fundamental concepts to understanding Lucene: *index, document, field, and term.* An *index* is a collection of documents–these may be webpages, text files, etc. A *document* is a collection of *fields*-fields may be information such as author, date, document content, etc. Fields contain *terms*, these are the values of the fields and may be numbers, signifying a date, an author's name, or an image title. Two terms with the same value, but belonging to different fields are considered differently.

Additionally, an index stores statistics about terms in order to make term-based search more efficient. Lucene is characterised by inverted indexing–this means it can list, for a term, the documents that contain it. Fields may be stored, in which case their text is stored in the index literally, in a non-inverted manner. Fields that are inverted are called indexed. A field may be both stored and indexed. The text of a field may be tokenized into terms to be indexed, or the text of a field may be used literally as a term to be indexed. Most fields are tokenized, but sometimes it is useful for certain identifier fields to be indexed literally.

Lastly, Lucene indexes may be composed of sub-indexes, or *segments*. Each segment is a fully independent index, which could be searched separately. Indexes evolve by creating new segments for newly added documents and merging existing segments.

7.4.2   Nutch

Nutch is an extension of Lucene, for indexing the web. It adds a crawler, a link database, and the ability to parse HTML pages, among others. It also provides support for scaling-up and clustering by using Hadoop, an implementation of Google's map/reduce computing paradigm, and a Distributed File System (DFS). Nutch is roughly composed of three parts (Figure 7.3, after [226]): fetcher, parser and indexer. The fetcher is responsible for retrieving web pages from the web as well as updating any renewed content. The parser navigates through the fetched HTML (or other types) content and identifies tags and content available for indexing. Lastly, the indexer, with all the capabilities of Lucene, assembles indexable parts into the Nutch database.

7.4.3   LIRE

The LIRE (Lucene Image REtrieval) library provides a simple way to retrieve images and photos based on their color and texture characteristics. It creates an index, based on Lucene, but the relevant fields and terms relate to image features for CBIR. The list of features that can be retrieved are presented below:

1. Color histograms in RGB / HSV color space.

2. MPEG-7 descriptors, such as scalable color, color layout and edge histogram

3. The Tamura texture which includes coarseness, contrast and directionality
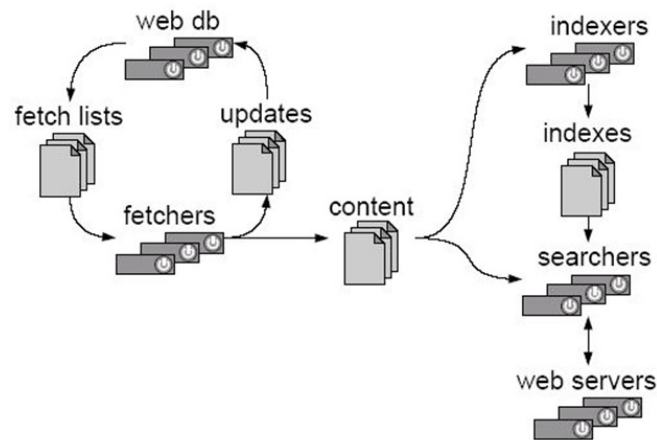
Figure 7.3: The flow of data through Nutch

4. Color and edge directivity descriptor, CEDD,

5. Fuzzy color and texture histogram, FCTH

6. Auto color correlation feature

Each image is saved in a Lucene index as a separate document. Prior to indexing, the desired features that will be saved must be selected. This is implemented by *DocumentBuilders*; for example, the fast Document builder, retrieves only the color layout information. Other builders also include: Extensive (all mpeg-7 descriptors), Auto color correlation, Full (all features). Once an image is indexed, the user is able to query the index with another image. The search functions also relate to the existing index–one cannot search for a Tamura feature if it has not been indexed. Figure 7.4 illustrates the concept.

Figure 7.4: The flow of indexing/searching provided by LIRE

7.4.4 Integration and Implementation

The integration of Nutch and LIRE into a search engine involved extending, modifying and creating a number of plugins for the project. Figure 7.5 shows the state of Nutch and LIRE before the extensions; note that there is no interaction between the two. Nutch is capable of indexing, crawling, and searching just for text and links. On the other hand, LIRE retrieves CBIR information from a repository of images (local or online). Figure 7.6 shows the re-design of the system; the plugin additions are represented with orange arrows.
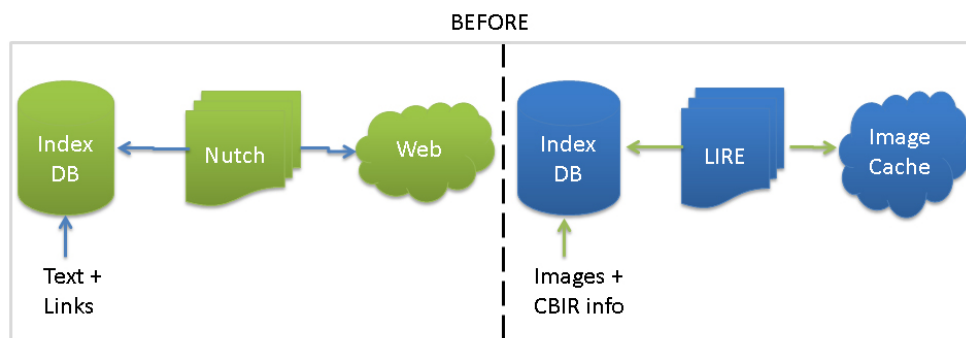
Figure 7.5: Nutch and LIRE basic functionalities before the extensions

Extensions were designed, developed and integrated into Nutch in order to:

1. Save images according to criteria (dimensions/input-output format)

2. Optionally filter out sites by keyword

3. Validate images (discard faulty/invalid links)

4. Cater for duplicates (ignore same images)

5. Be able to use LIRE for:

   a) Open source image information indexing.

   b) Indexing image data such as colour/edge histogram, colour layout etc

Figure 7.7 presents the overall workflow of the search engine in the crawling and indexing phase, with the included modifications.
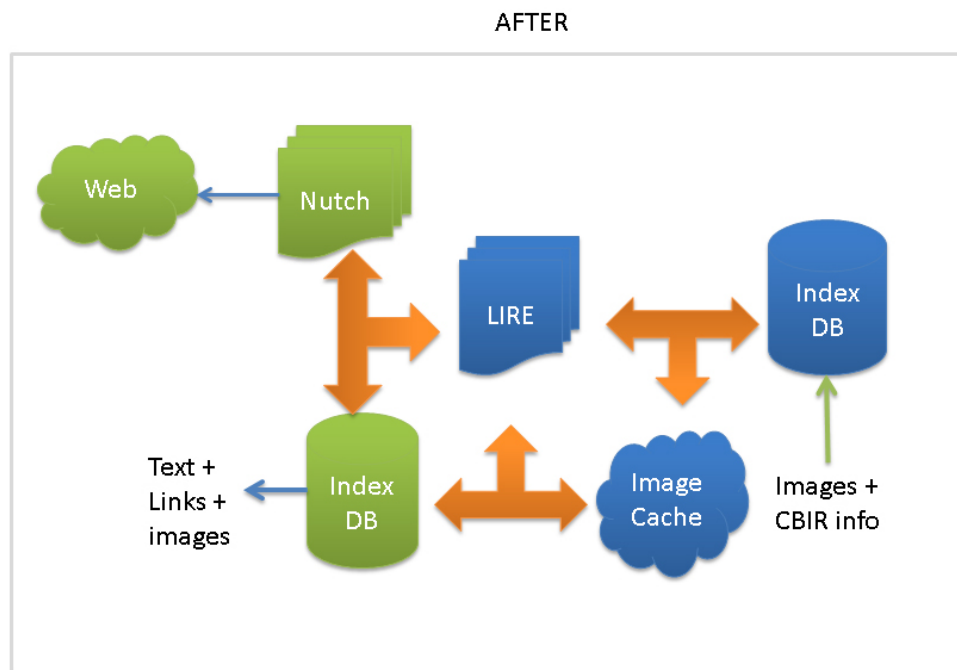
AFTER

Figure 7.6: Nutch and LIRE interactions after the extensions

Figure 7.7: Crawling and indexing text and images from the web

Parsing and retrieval extension

The default HTML parser of Nutch was extended in order to recognise HTML <img> tags. This allowed to get information on the image file, such as its link and dimensions. Testing indicated that the majority of images being discarded were GIFs.

Specifically, testing with approximately 1.4 million URLs returned 6,942,640 link checks for images. Out of those, 88% were GIF checks and 12% were JPG. About 97% of the checks were checks for duplicate images, and only 3% of the images were considered as save-able. An estimated 99% of the images eventually saved were JPEG. Finally, 82% of the time is spent processing GIF images, of which at the best, less than 1% are coins. As a result, the user is provided with the option to exclude parsing and caching of GIF images if so desired.

The flow diagram in Figure 7.8 illustrates how the HTML document is traversed and outgoing image links, belonging to a *src* attribute, are saved. In case that GIF filtering is on, all GIF image links will be omitted.

Indexing extension

After the fetching cycle finishes, the indexing commences, where data from the web sites is actually indexed in the Nutch local database. During this phase Nutch retrieves the crawled web page, indexes its information, and ranks the keywords with its ranking system. The default Nutch indexing filter was extended in order to be capable to index any images within a web page. Additionally, a number of features have been included that make the archiving of images more efficient.

1. Size check: If an image is below a certain threshold provided by the user, it is not archived (e.g. 128x128 pixels).

2. Validity check: If for any reason, the image file is erroneous it is not archived.

3. Image indexing check: If for any reason, the image fails the image indexing check (of LIRE), it is not archived

4. Duplication check: If the image already exists, with the same URL and path, it is not double saved.

5. Unique ID check: For each image to be saved in the database, a unique MD5 hash key is composed by the URL of the image and used as a file name.

6. Image output: The user is able to specify what output the images will be saved as. The default is PNG, but the user is able to choose from well known formats such as JPG/TIFF if so wished.

7. Coin candidate: This includes the capabiilty to identify while indexing an image as probably belonging to a coins-related page. The user is able to provide keywords or site-specific tags, and if located within the text, the particular page is tagged as a probable coins candidate. This gives greater flexibility in the search results.

The above were implemented in a new indexing filter plugin, its flow diagram is presented in Figure 7.9.

Figure 7.8: Parsing and archiving image links in an HTML page

Index HTML text & Images

This class saves suitable images as a Lucene Document and indexes both relevant HTML text as well as image CBIR information by connecting with LIRE. This way both types of content become searchable.

Implements the default Nutch IndexingFilter.

Receive fetched HTML page

imglink>0?

Yes — Create imgLink Hash

No — Ignore Page

imgLink exists in HashSet?

No — Add ImageLink to HashSet

Yes

Attempt to save Image

More Images in Page?

Yes

No

Img dimensions > min?

No

Yes

Send img & MD5 hash to **ImageCBIRIndexer**

Img CBIR info processed and added to LIRE index?

No

Yes

Image saved in DB with MD5 hash

SavedImage ++

SavedImage > 0 ?

No

Yes

Text has coin keywords?

No — Page Info indexed & saved in DB

Yes:: Flag as possible coin

Figure 7.9: Filtering and indexing suitable images from an HTML page.

Image Indexing extension

The Nutch image plugin has been further extended in order to encapsulate LIRE in the search engine architecture. The image CBIR indexer examines the incoming file, and saves the image characteristics in a LIRE/Lucene index, according to the features chosen (as described in 7.4.3). Each image is a Lucene document with a unique ID as the one in the Nutch index, formed by the MD5 hash. In this way, a direct link between the two indexes is established and subsequently allows for querying both or one of them at a time. The image indexer checks for:

- File validity: can the requested features be retrieved from the image? if not, it is not archived.

- Index existence: if an index is not available, it is created.

- Filter check: checks which feature filter is requested and in case of errors it uses the default MPEG-7 filter.

- Optimisation: Optimises and cleans the index for faster retrieval.

Figure 7.10 illustrates the flow.

7.4.5   Searching

In order to demonstrate the effectiveness of the search tool, two methods were deployed for searching the indexed content. The first one is a web-based application while the second is a Java-based standalone application.

Web-based search

The first searching method uses a web-interface and behaves as any well-known search engine. The search runs as a Java servlet on an Apache Tomcat server. The user can search by keywords and results displayed include the images related to the specific page. Other options include: viewing only images/viewing images from certain web sites/viewing content classified as potential coin candidates. Figure 7.11 illustrates the concept.

Image search application

The second search implementation demonstrates a CBIR search on the indexed images by calling on the LIRE API. The user is asked to input the index directory (of

This Class indexes and stores an image's CBIR information in a LIRE Lucene document according to the parameters desired

Receive image & MD5 hash name

Get chosen image filter from configuration.

Add Image to LIRE index

Optimise index
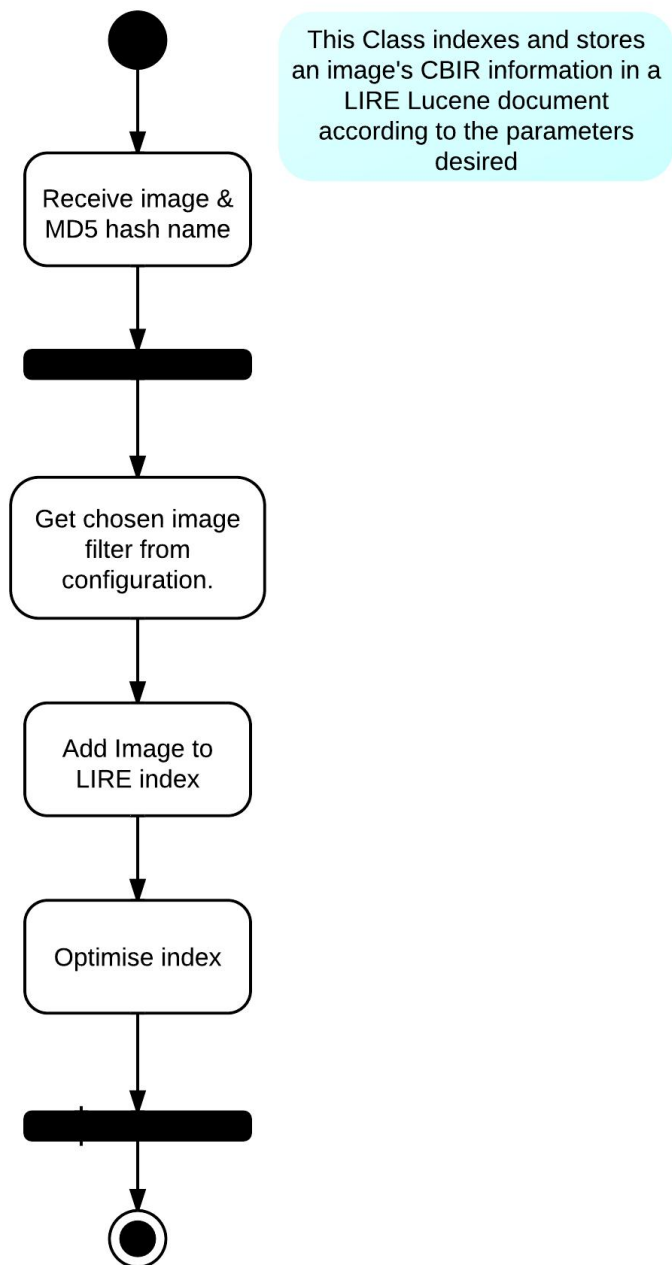
Figure 7.10: Indexing an image's CBIR information.

(a) A simple search that displays images.

(b) A search that displays images, relevant links, and summary description.
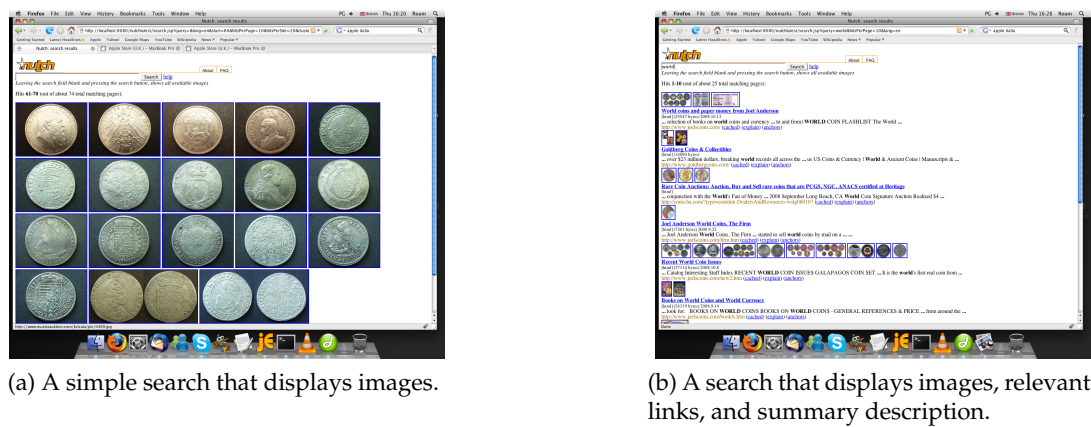
Figure 7.11: Example text searches

(a) Configuring the search application

(b) Results with a low threshold and 30 maximum returned.
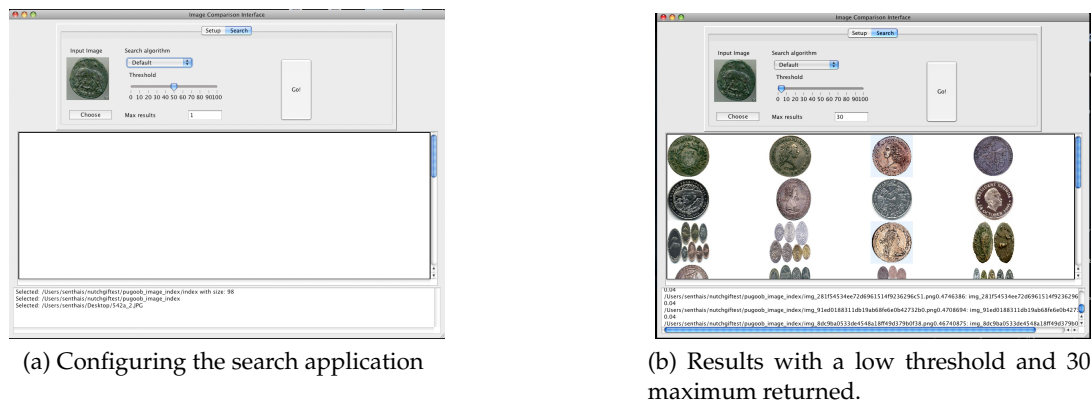
Figure 7.12: Example image searches.

the image features) and the location of the archived images. Then, the user is able to upload an image to use as a comparison. It also incorporates the ability to select a search method according to the features saved, the maximum results returned, and a threshold value for likeness. The threshold value is a number from 0 to 1 that indicates how similar a picture is to another; the closer the number is to 1 the more likely they look similar. By having a higher threshold value only the most likely of results will be shown. Figure 7.12 shows the user interface in action with a random coin image from an auction site chosen as input. The default MPEG 7 descriptor has been chosen, and the results appear as most likely first. The feedback log at the bottom of the interface shows the likeliness of each image, its filename, and its location.

Configuration file

All the configurations and the options mentioned as part of the implementation are integrated as part of an XML configuration file. This is used by default by Nutch in order to configure aspects of the crawling/indexing process, plugin information, etc. It also serves as a configuration basis for customised plugins such as the one implemented. Through this way, the user is able to make any chances simply by interacting with the XML file. For example, to choose a differeent image indexer one can just change a number in the configuration file as shown in the code snippet.

```
<property>
        <name>imgindexer.current.indexer</name>
        <value>1</value>
        <description>
                These are the available indexers
                provided by LIRE. Values:
                1: Default
                2: Extensive
                3: Fast
                4: Full
                5: CEDD
                6: Tamura
                7: AutoColorCorrelation
                8: Fast AutoColorCorrelation
                9: Colour Histogram
        </description>
</property>
```

7.5 Experiments

This section analyses the experiments done when testing the search engine. Because it is composed of different parts, such as the crawler, the indexer, the image indexer, and the searcher, a number of different tests was also used. The tests check for performance of the algorithms as well as the validity of the results returned. All benchmarks were composed on a single PowerMac G5 with 1GB of ram and a 2 Mbit ADSL line. A list of 43 coins-related URLs including eBay was fed to the engine and search was restricted on those sites. The following Section illustrates how a site was chosen for inclusion.

Site name	Local search	Worldwide search
Italy	1686	6495
US	3383	7680
UK	935	6358
France	1388	4227

Table 7.2: Comparison of localised eBay sites

Keyword: Denarius	Category: Coins	Category: Other	%
UK	144	3	2.1
US	308	4	1.3
Italy	427	3	0.7
France	17	0	0

Table 7.3: Comparing results across eBay categories (A)

7.5.1 Focus: eBay

As mentioned in the introduction, a large proportion of illegal online trade occurs on the eBay site. As a result, one of the sites the spider uses is eBay. A study was made of four eBay sites, namely the French, UK, Italy and US in order to estimate:

- Which of the three sees the most traffic.

- Number of relevant coin results in the ancient coins category

- Number of relevant coin results in other categories.

Table 7.2 shows the number of results retrieved for all items in the Ancient Coins category, both for a local and a worldwide search. The US site gives the maximum number of results.

Another point to consider was whether coins were found in other categories than the Ancient Coins one and whether these categories should be included in the crawl. Tables 7.3 and 7.4 show the results of a simple keyword case; in Table 7.3 queries were done in the local language. Results are included from Ancient coins and the other categories. It was found that a very low number of coins appeared in other categories than Ancient Coins, mostly in the Antiquities one.

It would seem that eBay users try to be as accurate as possible when assigning categories for their listings. These observations also include items that may belong in

Keyword: Roman Coin	Category: Coins	Category: Other	%
UK	413	70	16.9
US	1026	74	7.2
Italy	114	5	4.4
France	91	3	3.3

Table 7.4: Comparing results across eBay categories (B)

both categories and would be thus retrieved from a crawl in just the Ancient Coins one. As a result the cost of crawling other categories for this case study was considered unnecessary.

Lastly, the spider adheres to the good practices of web robots and fully respects robots.txt permissions before crawling a web page.

7.5.2   Crawling and indexing

Table 7.5 shows the time taken for crawling and indexing respectively. While the crawling appears to take a large amount of time, this is due to the threaded crawling requests to the sites. They are low in number, for politeness so as not to overbear any site. An additional matter is the bandwitdh, and the use of a single machine; since Nutch supports cluster architecture, the same tests on a cluster would take less time. In both cases, the indexing and saving of the images takes about 21-23% of the time–and again this could be decreased by using a cluster architecture.

N	URLs fetched	t fetching	t indexing imgs	Cache size	Total imgs
10,000	12,195	13 hrs	3 hrs	1 gb	11,161
100,000	146,804	6 days	30 hrs	17 gb	166,501

Table 7.5: Crawling and indexing benchmarks

7.5.3   Image characteristics

In 7.5.2 one of the two tests resulted in 166,501 images. Further tests were conducted on those results in order to identify the types of images returned and examine any po-

tential waste of time/bandwidth. While the end result was 166,501 images, the actual examination for links were 6,942,640. From these examinations, 87% (6,083,730) were checks for GIF files, 12% for JPG and the rest for other types e.g. PNG. These millions of checks (which is actually a first check for duplicates) allowed 195,295 of images to pass through second checks (i.e. size/validity). The end result after size/validity is the original 166,501 where 99% are jpg and 0.7% are GIF. Upon examining the GIF for relevant coin content, 8% (81) were coins.

Following these results, an option to exclude GIF files if desired, was implemented. This extension has been described in Section 7.4.4.

7.5.4   Image features indexing

Tests were also conducted in the indexing phase in order to retrieve the indexing times of the image features. Table 7.6 illustrates the indexing time taken by a simple colour histograpp and the full index which combines all features. The difference is that the full is about 32 times slower, considering that the simple takes 104 ms/image and the other 3400ms/image.

Type of Index	N. of images	Time/Image
Colour histogram	30,000	104 ms
Full index	1,340	3.4 sec

Table 7.6: Image features indexing

7.5.5   Searching

The searching capabilities were tested in a local filesystem, by two different means: through the web application, and the image application index. Both perform quite fast, the former giving an average of 3 sec/search and the second 1.1 sec/search with the lowest threshold (maximum numbers of results returned).

7.6   Conclusion

This Chapter described extensions to an open source search engine developed as part of the COINS project. The extensions integrate and extend a number of open source applications such as NUTCH and LIRE. As a result, it provides an engine capable of querying coins web sites, retrieving candidate images, indexing both HTML information and image features and searching both by keyword and image data. The

flexibility of the systemy was achieved by providing two means of search capabilities, a web-based application and a Java stand-alone system. The transfer from a Java application to another, preferred medium, would involve only the writing of the frontend, as the behaviour remains the same. Further work can include testing scenarios with remote/concurrent users and the deployment on a cluster system.

Due to the modularity of the new plugins developed for this project and also of the underlying applications, this engine can feasibly and easily be extended to other types of searches. The indexed results can also be further enriched with additional metadata in order to be semantically queried. In fact, one of the main aims of COINS was to add capabilities for semantic search and classification to the cached index of text and images. The search engine is tightly integrated with an ontology-based semantic database capable of storing and querying information. The semantic database uses the CIDOC-CRM standard [232]. CIDOC is an ontology mapping for cultural heritage data; a portion of it is devoted specifically to archaeological data.

Crisp semantic descriptions are in development for Nutch in different search areas (Mathematics [233], e-learning [234]). At the same time, ontological descriptions for archaeological scientific publications, one of the main sources of archaeological knowledge repository, have also been explored [235] Additionally, projects such as 3D-COFORM [94], which extend work done with the COINS Semantic database, provide ways to semantically describe excavation data. As with the development of mathematics, from crisp to elastic, fuzzy sets, researchers in semantic descriptions are now exploring the concept of vagueness and uncertainty.

A seminal paper by Lukasiewicz and Straccia [236] detailed significant research in the area of semantic ontologies towards representing and reasoning with uncertainty and vagueness. They describe a set of possibilistic and probabilistic expressive descriptive logics that can aim to extend classical ontology languages. During the same period, the World Wide Web consortium initiated an incubator group to examine uncertainty reasoning for the web [237] which examined fuzzy reasoning, probability theory and evidence theory. One of its aims was to identify case studies where uncertainty reasoning would aid in extracting useful information.

Let us return to Sally, the archaeology graduate, who unearthed an unidentifiable shard of pottery (Chapter 3.3). She uploads a picture of it to a semantic database of archaeological information, together with as much description as possible, however vague, i.e. brown colour, about 30 gr. in weight, rough texture. The search engine, extended as to understand uncertain descriptive logics such as this, crawls through the

database and returns relevant results that can immediately aid in identification. At the same time, the archaeologist can update said database with her own discoveries and identifications.

Writing non-crisp semantic extensions for Nutch, or other, search engines is fully feasible and it is envisioned that this will be an active research area in the next years. Furthermore, non-crisp semantic descriptions could enhance the CIDOC-CRM standard.

As with archaeological uncertainty, semantic search featuring vague and uncertain logics, instead of crisp descriptions, is a long way from being widely used. However, the author cannot help but notice that in the future these two areas will be very much related in order to construct and query archaeological knowledge bases with correct ontologies and adequate descriptive languages that include uncertainty and vague knowledge. Figure 7.13 illustrates the concept diagram of a proposed system for archaeological 3D reconstructions, which acknowledges uncertainty and evidence and can be powered by a semantically searchable Knowledge Base.
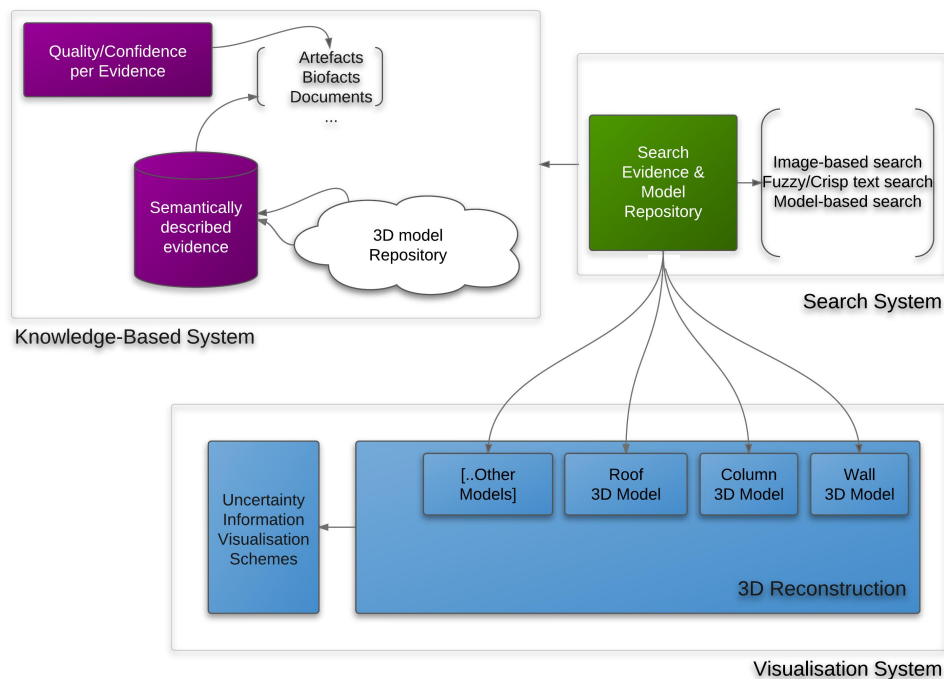
Figure 7.13: Global system diagram

The Knowledge Base will include (fuzzy) semantically-described archaeological

evidence. It will be updateable by experts and evidence can have assigned confidence levels. Any 3D models created, must acknowledge evidence relevant to their creation. This repository is envisioned to be cloud-based and distributed. In the visualisation system segment, the expert wanting to create an archaeological reconstruction has to query the knowledge base for available model parts. Alternatively, she can also query the KB for ambiguous/unidentified evidence or models. Uncertainty visualisation schemes can be changed locally to provide for a variety of visualisations.

The next Chapter, the conclusion to this thesis, provides an overall discussion of the thesis's problem questions and contributions and expands on future research work needed.

The End Is the Beginning Is the
End.

CHAPTER 8

Conclusion and Future Work

This thesis explored, described, quantified and visualised uncertainty in a cultural informatics context, with a focus on archaeological reconstructions. For quite some time now, archaeologists and heritage experts have been rightly criticising the often too-realistic appearance of three-dimensional reconstructions. They have been highlighting one of the unique features of archaeology: the information we have on our past will always be incomplete. This incompleteness should be reflected in digitised reconstructions of the past.

This criticism composed the major motivation behind the thesis. The research started by examining archaeology itself, archaeological theory and archaeological inference, and followed with an insight into computer visualisation and how these two areas have formed a useful but often tumultuous relationship through the years.

An extensive literature review on uncertainty concluded that the quantification and visualisation of archaeological uncertainty itself was at the very first, conceptual, stages when compared to other disciplines. By examining the uncertainty background of disciplines such as GIS, medicine, and law, the thesis postulated that archaeological visualisation, in order to be taken seriously, should move towards archaeological knowledge visualisation. Three steps were identified for the initial exploration of archaeological uncertainty: identification, quantification and modelling. A questionnaire, targeting expert archaeologists, was designed, modelled and distributed in order to gauge perception about uncertainty as well as different types of evidence. On the matter of quantification, three established mathematical models of uncertainty quantification, Probability Theory, Possibility Theory, and the Transfer-

able Belief Model, were presented, explained, innovatively applied to archaeological scenarios and then evaluated. Following, a previous modelling approach to uncertainty using fuzzy logic was critically examined. A series of simulations were designed and run in order to examine differences between results of the mathematical models. The visualisation of Building III aimed to explore the differences and capabilities between the various models. Additional visualisations of using information visualisation schemes evaluated the uses of texture, glyphs and graphical variables as representations of uncertainty and conflict. Lastly, the design and development of a extensions to a search engine led to the suggestion that perhaps all this knowledge, as well as the awareness of its limitations, can be all tied together with vague semantic descriptors. The thesis contributions are now presented in more detail.

8.1 Thesis contributions

The main contributions of this thesis are:

- The innovative design, development and distribution of a questionnaire which examines archaeological uncertainty in interpretation during excavations.

- One of the questionnaire's conclusions was that archaeologists across different disciplines are equally concerned about uncertainty in interpretations and reconstructions.

- Another conclusion was that archaeologists may indicate preference among different categories of evidence. This can influence the choice of a quantification algorithm as well as visualisations.

- Three mathematical uncertainty quantification models were explained, applied and evaluated for the first time, for the purposes of archaeological belief representation and reasoning visualisation.

- The behaviour of Possibility, TBM and Subjective Probability were examined through simulations. Statistical tests indicated that the results converge significantly. However, the unpredictable over/under confidence of Possibility was highlighted and the application of the Ψ-scale was proposed.

- Examination of previous approaches to modelling uncertainty on fuzzy logic were critically examined and disputed.

- Suggestions that a probabilistic approach is not capable of representing archaeological uncertainty were refuted and was proved that it can be done through a Bayesian approach.

- Visualisation of Possibilistic and Probabilistic beliefs in an actual case study highlighted that even though the results of the two are significantly similar, during the visualisation differences are apparent. As a result, the importance of choosing a mathematical model appropriate for the circumstances is even more crucial under the light of differences in visualisation.

- Innovative visualisation of belief and conflict elements on 3D archaeological reconstructions by using an actual archaeological site as a case study as well as simulated models. Application and evaluation of a variety of information visualisation schemes.

- The design and development of extension to a search engine for searching for archaeological data. Results from this research can be extended towards a semantic search engine capable of handling uncertainty. Finally, it is suggested that semantic annotation of uncertainty and representations of it (i.e the models) would be a positive step in the direction of typologies and standards.

8.2 Future work

Representing uncertainty in archaeological reconstructions is currently a novel and niche research area. This thesis has aimed to contribute in examining some foundational aspects. There is vast amount of further work to be done towards this area, the following sections analyse possible directions for such work. Figure 8.1 recaps the predicted stages of uncertainty research.

8.2.1 A typology for archaeological uncertainty in 3D reconstructions

Archaeologists face two tasks when evaluating evidence towards an interpretation. The first one is to understand how accurate the evidence is; measurement quality, granularity, possible errors etc. The second one is to assign their own level of confidence, importance to that evidence. Depending on the type of evidence, or what it is being used for, it may be deemed more, or less important.

One of the contributions of this thesis has been applying mathematical models of uncertainty representation to archaeological scenarios. In order to develop the repres-

Figure 8.1: Stages of uncertainty research

entation further and use features such as information fusion, a typology framework for the different types of uncertainties faced by archaeologists should be established. A first step would be to evaluate similar efforts done for analysts in intelligence gathering [238]. The particular research by Thomson uses classifications by Pang [110] and Gershon [108] to suggest a typology for intelligence gathering. Zuk [140], based on Thomson's suggestions extends the typology to a more general typology for reasoning.

Much work has been done in the cultural heritage sector in CIDOC-CRM [239]. CIDOC-CRM can be used as a bridge between other typologies and one that expresses uncertainty in reasoning and data when archaeological 3D reconstructions are involved. Concise identification of typologies will hopefully, subsequently lead to the establishment of standards.

8.2.2   Advanced quantification: Archaeological information fusion

Longley [111] provides a conceptual view of uncertainty, where the actual, real world has its complexity distorted and reduced through its representation in a GIS. Figure 8.2 presents an adaptation of this model for archaeological uncertainty.

In order to move forward towards the fusion of different kinds of uncertainty information (as identified through the typologies), the actual creation and propagation of uncertainty should also be acknowledged. For example, what kind of simplifications and assumptions are being made during a reconstruction? What information is willingly (or unwillingly) lost? Stages U1-U5 provide building blocks for further consideration.
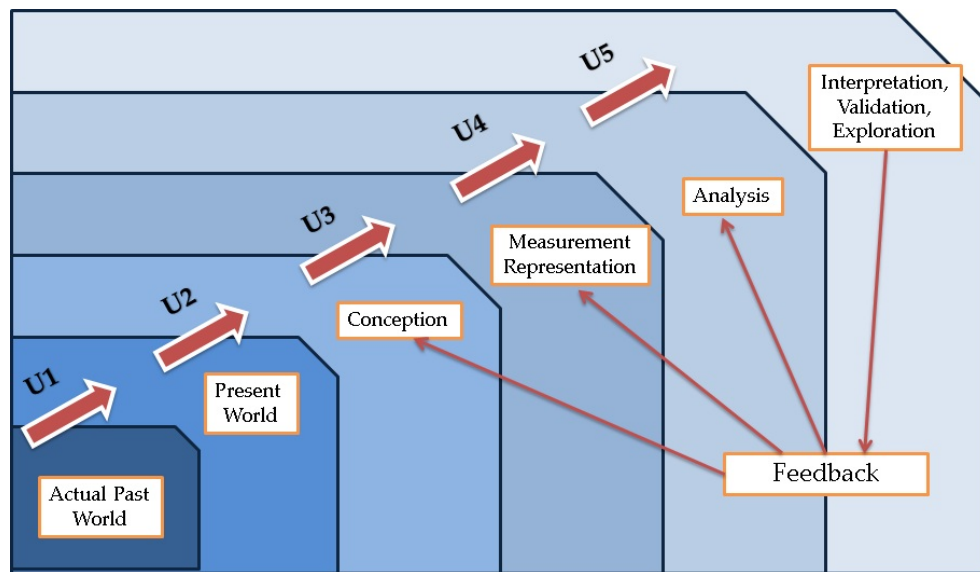
Figure 8.2: Uncertainty propagation

A good reason to acknowledge what kind of uncertainty is involved and where in the system it is most apparent, is that it can affect the choice of quantification and modelling. Taking another example from the GIS universe, IDRISI GIS allows for representation using the Dempster-Shafer, fuzzy as well as subjective probability. Especially if one chooses to use an information combination (fusion) system, perhaps involving fusing fuzzy and probabilistic information through belief functions [240]; the way the uncertainty moves through such a system would be crucial in modelling it.

8.2.3 Advanced modelling and visualisation approaches

A cumbersome issue in wanting to represent uncertainty is that the need for alternative models escalates. It is perfectly reasonable, if an evaluation or comparison is made between two or more hypothesis, that the expert would want to view both. Of course, with the current available tools it is not quite feasible due to cost and time constraints. This was also encountered during the visualisation of Building III, constraints prevented from creating the alternative model of a religious representation.

This also raises another issue: redundancy. So many models are created, somehow visualised and then abandoned. To continue this way with archaeological knowledge visualisation in mind would surely lead to many, many, more redundant models and limited collaboration or use between the experts creating them.

Two crucial areas for further work are the research of alternative ways of creating and sharing models of archaeological data, and exploration of visualisations. Two directions are proposed for the former:

Expert, community-monitored, open-standard model repositories

A Wikipedia for archaeological model components. This concept is inspired by the way computer games populate the game worlds by using building blocks of items, structures, trees, etc. As a result, many professional websites exist which provide such models. The models themselves are often rated by their users in terms of the model's quality and realistic representation.

Such a heritage repository would consist of 3D models or other data that can be used for archaeological representations, along with required information about their creation. For example, consider a texture map for a Roman plaster wall. The creator would have to explain how it was created, based on what sources and evidence. Also, other experts would be able to rate, make alterations and suggestions to the texture or model themselves–this is where the Wikipedia analogy comes in. The involvement of experts is very crucial as the crowdsourcing validation of the models would lend them a higher credibility.

If this repository and its data are also semantically described through CIDOC (and they should be, for completeness), a search engine capable of understanding semantic logic can provide a connecting point between a 3D modelling software and the repository itself.

An approach such as this would greatly reduce redundancy, associate an origin record to each modelised part which in turn would connect with actual archaeological knowledge through articles etc. Furthermore, it would provide a consistency that is currently lacking in archaeological visualisations.

Procedurally generated content

A diametrically different approach would be to invest in describing archaeological models procedurally. Procedural generation of 3D object means that the objects are created from a set of rules which describe not only their structure but relationship between their neighbours. For example, a procedurally-generated door will have sufficient knowledge for its dimensions but also where it can be placed, how many doors may be allowed in a building, its style (interior door, exterior door), its proximity to windows, etc. In essence it is a semantic description of the object. Such modelling

```
attr randomly = 0
attr winW = 2 // fixed , no uncertainty
attr winH =
60%: 1.5 // from "arch.␣source"
20%: 1.6 // from "arch.␣source␣2"
else: 1.7
```

has been used for quite a long time in computer games, especially for procedurally rendering huge cityscapes where individual modelling would be tedious, costly and extremely time consuming.

Such an approach was recently suggested by Haegler [241], who applies procedural modelling concepts to representing alternative archaeological reconstructions of buildings. Procedural models, because of their semantic capabilities, would feature an awareness of relationships between objects. For example, the family of columns belonging to an ancient Greek timeline would be Corinthian, Doric and Ionic. A procedurally generated city, set in the ancient Greek timeline, when needing some columns, it would be from the relevant family. Havemann also acknowledges the existence of uncertainty, and takes a purely crisp-set, probabilistic approach to its representation. The code snippet below (after [241]) represents a window's uncertain height as a probability distribution.

Although this is a step in the right direction, it is a very simplistic representation. In order to better represent the uncertainty involved in such a visualisation the procedural language and rules used must be extended to cater for vague meanings, belief functions, and external association to sources, rather than in a comment. Inspiration can be received from medical rule-based systems which also use procedural rules but also cater for vague, fuzzy information.

The very recent acquisition (late 2011) of City Engine, a graphics engine for procedural generation of buildings, by ESRI means only one thing: GIS is going to be immensely enhanced with procedural rules for generating structures, vegetation and more. It is only a matter of time before the uncertainty plugins are also applicable to procedural modelling.

Further exploration of visualisations

This thesis applied a number of visualisation approaches to archaeological uncertainty. Further work is needed in a number of areas. Firstly, the issue of swapping

between different visualisation schemes has to become as easy as changing, for example, a chart style in Microsoft Excel. The resources and amount of time required to create a 3D model is already quite high and it only escalates with the uncertainty quantification and visualisation. Plus, the issue of redundancy also occurs here; currently there is no way to transfer visualisation styles between models.

New approaches to 3D modelling such as those previously discussed, can lead to a parallel systematisation of visualisation approaches. Another future area of interest is the user experience and perception of such visualisations.

8.2.4 Knowledge bases

Chapter 7 briefly examined the difficulties of discovering expert information and expanded on the aspect of searching for said information in associated repositories. There is a second aspect, which is the elicitation of expert knowledge: in other words, how to make archaeologists, doctors, lawyers express their knowledge in a consistent manner. With regards to archaeological uncertainty, this is another area where further work is definitely needed. Firstly, to create a knowledge base about an archaeological domain, experts in that domain would need to be enrolled to share their knowledge. Secondly, an ontologically concise way is needed in order to express that knowledge. It should be semantically understandable by the system, and this would also put a great limitation on the experts. Thirdly, the elicitation of knowledge must take under consideration factors such as the conjunction fallacy and overconfidence/underconfidence that are often expressed by humans. Again, a first step of examination could be research in elicitation of expert medical knowledge which has a strong, 40-year, background.

8.3 Closing Remarks

The main objective of this research has been to highlight the concept of uncertainty in archaeological reconstructions, research means to quantify it, discover ways to model and visualise it and explore future ways to search information for it. It is extremely encouraging that through the time this thesis was being completed, new research has surfaced that deals with archaeological uncertainty and knowledge representation. The STAR and STELLAR projects [242], which make archaeological data (particularly grey literature and excavation reports) semantically accessible and searchable, through ontologically mapping them to CIDOC-CRM. Research by Moussa and

Fritsch [243], attempts to link 3D models with bibliographic repositories containing relevant evidential information. Verhagen et al [244] devote a book to archaeological site prediction involving the visualisation of its uncertainty, through using models such as Dempster-Shafer and Bayesian analysis.

Those research projects reinforce a number of points also made in this thesis and justify the original research position: That archaeological visualisation has the potential to be transformed to archaeological knowledge visualisation. This thesis suggests that the direction to get there would be through a re-examination of how 3D models are created. In order for the model to carry a representative significance, it should have a semantic awareness of what it is, what it can be associated with, and what information is available that gives it this representation. Semantic annotation of information and lack of information, vagueness, uncertainty, will be crucial for the further development of archaeological knowledge representation.

References

[1]   P. Miller and J. Richards, "The good, the bad, and the downright misleading:
       archaeological adoption of computer visualisation," in *Proceedings of CAA Com-
       puter Applications in Archaeology CAA '04*, J. Hugget and N. Ryan, Eds., 1995, pp.
       19–22.

[2]   C. Renfrew and P. Bahn, *Archaeology: Theories, Methods and Practice*, 5th, Ed.
       Thames & Hudson, 2008.

[3]   B. G. Trigger, *Gordon Childe: Revolutions in Archaeology*.   Londo:Thames and
       Hudson, 1980.

[4]   S. R. Binford and L. Binford, *New Perspectives in Archaeology*.   Chicago, Aldine
       Press, 1968.

[5]   I. Hodder, *Reading the past: Current approaches to interpretation in archaeology*.
       Cambridge: Cambridge University Pres, 1986.

[6]   D. L. Clarke, "Archaeology: The loss of innocence," *Antiquity*, vol. 47, pp. 6–18,
       1986.

[7]   M. Shanks and C. Tilley, *Reconstructing Archaeology: Theory and Practice*.   Cam-
       bridge: Cambridge University Press, 1987.

[8]   B. Trigger, *A History of Archaeological Thought*.   Cambridge University Press:
       New York, 1989.

[9]   L. R. Binford, "Data, relativism and archaeological science," *Man*, vol. 22,
       no. 3, pp. pp. 391–404, 1987. [Online]. Available: http://www.jstor.org/stable/
       2802497

[10]   A. P. Sullivan, "Inference and evidence in archaeology: A discussion of the conceptual problems," *Advances in Archaeological Method and Theory*, vol. 1, pp. pp. 183–222, 1978. [Online]. Available: http://www.jstor.org/stable/20170133

[11]   R. A. Watson, "Inference in archaeology," *American Antiquity*, vol. 41, no. 1, pp. pp. 58–66, 1976. [Online]. Available: http://www.jstor.org/stable/279041

[12]   R. H. Thompson, "The subjective element in archaeological inference," *Southwestern Journal of Anthropology*, vol. 12, no. 3, pp. 327–332, 1956. [Online]. Available: http://www.jstor.org/stable/3629088

[13]   A. Wylie, "The reaction against analogy," *Advances in Archaeological Method and Theory*, vol. 8, pp. pp. 63–111, 1985. [Online]. Available: http://www.jstor.org/stable/20170187

[14]   A. R. Smith, "Alpha and the history of digital compositing," in *Microsoft Technical Memo 7*, 1995.

[15]   E. Catmull and J. Clark, "Recursively generated b-spline surfaces on arbitrary topological meshes," *Computer Aided Design*, vol. 10, no. 6, pp. 350–355, 1978.

[16]   D. Doo and M. A. Sabin, "Behaviour of recursive subdivision surfaces near extraordinary points," *Computer Aided Design*, vol. 10, pp. 356–330, 1978.

[17]   D. Zorin and P. Schröder. (2000, March) Siggraph course notes: Subdivision for modeling and animation. Internet. [Online]. Available: http://mrl.nyu.edu/publications/subdiv-course2000/

[18]   A. G. Requicha, "Representations for rigid solids: Theory, methods, and systems," *ACM Computing Surveys*, vol. 12, no. 4, pp. 437–464, 1980.

[19]   A. Requicha and H. Voelcker, "Boolean operations in solid modeling: boundary evaluation and merging algorithms," *Proceedings of IEEE*, vol. 73, no. 1, pp. 30–44, 1985.

[20]   J. F. Blinn, "Simulation of wrinkled surfaces," *ACM SIGGRAPH Computer Graphics*, vol. 12, no. 3, pp. 286–292, August 1978.

[21]   B. T. Phong, "Illumination for Computer Generated Pictures," *Communications of the ACM*, vol. 18, no. 6, pp. 311–317, June 1975.

[22] S. R. Buss, *3-D Computer Graphics: A Mathematical Introduction with OpenGL.* Cambridge University Press, 2003.

[23] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Transactions on Graphics*, vol. 1, no. 1, pp. 7–24, January 1982.

[24] C. Schlick, "An inexpensive BRDF model for physically-based rendering," *Computer Graphics Forum*, vol. 13, pp. 233–246, 1994.

[25] A. Appel, "Some techniques for shading machine rendering of solids," in *Proceedings of AFIPS 1968 Spring Joint Computer Conference*, ser. 32, 1968, pp. 37–45.

[26] T. Whitted, "An improved illumination model for shaded display," *Commun. ACM*, vol. 23, no. 6, pp. 343–349, 1980.

[27] C. M. Goral, K. E. Torrance, D. P. Greenberg, and B. Battaile, "Modeling the interaction of light between diffuse surfaces." *Computer Graphics (SIGGRAPH '84 Proceedings)*, vol. 18, pp. 213–222, July 1984.

[28] G. Pasko, A. Pasko, T. Ikedo, and C. Vilbrandt, "Virtual shikki and sazaedo: Shape modeling in digital preservation of Japanese lacquer ware and temples," in *Proceedings of SCCG the 17th Spring Conference on Computer Graphics '01*, 2001, p. 147.

[29] J. Stumpfel, C. Tchou, T. Hawkins, P. Debevec, J. Cohen, A. Jones, B. Emerson, P. Martinez, and T. Lochman, "Assembling the sculptures of the parthenon," in *Proceedings of SIGGRAPH 2003 conference on Sketches & applications*. New York, NY, USA: ACM Press, 2003, pp. 1–1.

[30] D. o. A. Brown University. (2005) Internet. [Online]. Available: http://www.brown.edu/Departments/Anthropology/Petra/excavations/history.html

[31] E. Vote, D. A. Feliz, D. H. Laidlaw, and M. S. Joukowsky, "Discovering petra: Archaeological analysis in vr," *IEEE Computer Graphics and Applications*, vol. 22, no. 05, pp. 38–50, 2002.

[32] M. Ashley López, "Real webs and virtual excavations: A role for digital media recording in archaeological site management," in *Proceedings of the Virtual World Congress*, Mexico City, Mexico, 2002.

[33]   M. White, F. Liarokapis, N. Mourkoussis, A. Basu, J. Darcy, P. Petridis, M. Sifniotis, and P. F. Lister, "ARCOLite - an XML based system for building and presenting virtual museum exhibitions using web3d and augmented reality," in *Theory and Practice of Computer Graphics (TPCG'04)*.   Birmingham, UK: IEEE Computer Society, 2004, pp. 94–101.

[34]   M. White, P. Petridis, and M. Sifniotis, "ARCO - Augmented Representation of Cultural Objects," in *Computer Applications and Quantitative Methods in Archaeology, CAA UK 2005, Chapter Meeting*, January 2005 2005.

[35]   V. Vlahakis, N. Ioannidis, J. Karigiannis, M. Tsotros, M. Gounaris, D. Stricker, T. Gleue, P. Daehne, and L. Almeida, "Archeoguide: an augmented reality guide for archaeological sites," *Computer Graphics in Art, History and Archaeology": Special Issue of the IEEE Computer Graphics and Applications*, vol. 22, no. 5, pp. 52–60, 2002.

[36]   V. Vlahakis, A. Demiris, E. Bounos, and N. Ioannidis, "A novel approach to context-sensitive guided e-tours in cultural sites: "light" augmented reality on PDAs," in *Proceedings of VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST '04*, Y. Crysanthou, K. Cain, N. Silberman, and F. Niccolucci, Eds., 2004, pp. 57–66.

[37]   G.Papagiannakis, S. Schertenleib, B. Kennedy, M. Poizat, N.Magnenat-Thalmann, A. Stoddart, and D.Thalmann, "Mixing virtual and real scenes in the site of ancient pompeii," *Computer Animation and Virtual Worlds*, vol. 16, no. 1, pp. 11–24, February 2005.

[38]   U. World Heritage Tour. (2005). [Online]. Available: ~http://www.world-heritage-tour.org/list.html

[39]   S. Kenderdine, "1000 years of the olympic games: treasures of ancient greece. digital reconstruction at the home of the gods," in *Proceedings of WWW '01: Proceedings of the 10th international conference on World Wide Web*.   New York, NY, USA: ACM Press, 2001, pp. 67–75.

[40]   A. Gaitatzes, D. Christopoulos, and M. Roussou, "Reviving the past: Cultural heritage meets virtual reality," in *Proceedings of VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST '01*, 2001.

[41] M. Roussou, "Virtual heritage: From the research lab to the broad public," in *Proceedings of VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST '00*, ser. BAR International Series 1075, F. Niccolucci, Ed. Arezzo, Italy: Oxford, Archaeopress, 24-25 November 2000, pp. 93–100.

[42] D. Pletinckx, D. Callebaut, A. E. Killebrew, and N. A. Silberman, "Virtual-reality heritage presentation at ename," *IEEE MultiMedia*, vol. 7, no. 2, pp. 45–48, 2000.

[43] A. Chalmers and S. Stoddart, *Imaging the past*. British Museum: Occasional Paper number 114, 1996, ch. Photo-Realistic Graphics for Visualising Archaeological Site Reconstructions, pp. 85–94.

[44] K. Devlin and A. Chalmers, "Realistic visualisation of the pompeii frescoes," in *Proceedings of the 1st international conference on Computer graphics, virtual reality and visualisation*. ACM Press, 2001, pp. 43–48.

[45] N. Ryan, *Imaging the past*. British Museum: Occasional Paper number 114, 1996, ch. Computer based visualisation of the past: technical 'realism' and historical credibility., pp. 95–108.

[46] D. Green, J. Cosmas, T. Itagaki, M. Waelkens, R. Degeest, and E. Grabczewski, "A real time 3D stratigraphic visual simulation system for archaeological analysis and hypothesis testing," in *Proceedings of VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST '01*. ACM Press, 2001, pp. 271–278.

[47] J. Woodwark and A. Bowyer, "Better and faster pictures from solid models," *IEE Computer Aided Engineering Journal*, vol. 3, no. 17, pp. 17–24, 1986.

[48] J. Woodwark, "Reconstructing history with computer graphics," *IEEE Computer Graphics and Applications*, vol. 11, no. 01, pp. 18–20, 1991.

[49] P. Reilly, "Data visualization in archaeology," *IBM Systems Journal*, vol. 28, no. 4, pp. 569–579, 1989.

[50] K. Delooze and J. Wood, "Furness abbey survey project: the application of computer graphics and data visualisation to reconstruction modelling of an historic monument," in *Proceedings of CAA Computer Applications in Archaeology*, ser. 565, K. Lockyear and S. Rahtz., Eds. British Archaeological Reports, 1990, pp. 141–148.

[51] B. Collins, "From ruins to reality - the Dresden Fraüenkirche," *IEEE Computer Graphics and Applications*, vol. 13, no. 6, pp. 13–15, 1993.

[52] P. Palamidese, M. Betro, and G. Muccioli, "The virtual restoration of the visir tomb," in *Proceedings of IEEE Visualization '93*, G. M. Nielson and R. D. Bergeron, Eds., 1993, pp. 420–423.

[53] M. P. Gavin Bell, Anthony Parisi, "The vrml 1.0 specification," in *in Proceedings of the Second International Conference on the World Wide Web*, Chicago, October 1994.

[54] M. P. Gavin Bell, Rikk Carey, "The vrml 2.0 specification," in *VRML 97 Proceedings*, 1997.

[55] D. Bearman, "Standards for networked cultural heritage," *Archives and Museum Informatics*, vol. 9, pp. 279–307, 1995, 10.1007/BF02770469. [Online]. Available: http://dx.doi.org/10.1007/BF02770469

[56] M. Gillings and G. Goodrick, "Sensuous and reflexive gis: Exploring visualisation and vrml," *Internet Archaeology*, vol. 1, no. 1, pp. –, 1996. [Online]. Available: http://intarch.ac.uk/journal/issue1/gillings/

[57] J. C. Roberts and N. Ryan, "Alternative archaeological representations within virtual worlds," in *Proceedings of the 4th UK Virtual Reality Specialist Interest Group Conference Brunel University*, R. Bowden, Ed., Uxbridge, Middlesex, November 1997, pp. 179–188.

[58] M. Terras, "A virtual tomb for kelvingrove: Virtual reality, archaeology and education," *Internet Archaeology*, vol. 7, pp. –, 1999. [Online]. Available: http://intarch.ac.uk/journal/issue7/terras_index.html

[59] M. Gaiani, E. Gamberini, and G. Tonelli, "Vr as work tool for architectural amp; archaeological restoration: the ancient appian way 3d web virtual gis," in *Virtual Systems and Multimedia, 2001. Proceedings. Seventh International Conference on*, 2001, pp. 86 –95.

[60] G. Earl and D. Wheatley, *Contemporary Themes in Archaeological Computing*. Oxbow Books, 2002, ch. Virtual reconstruction and the interpretative process: a case-study from Avebury, pp. 5–15.

[61] M. White, N. Mourkoussis, J. Darcy, P. Petridis, F. Liarokapis, P. Lister, K. Walczak, R. Wojciechowski, W. Cellary, J. Chmielewski, M. Stawniak, W. Wiza, M. Patel, J. Stevenson, J. Manley, F. Giorgini, P. Sayd, and F. Gaspard, "Arco " an architecture for digitization, management and presentation of virtual exhibitions," in *Proceedings of the Computer Graphics International*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 622–625. [Online]. Available: http://portal.acm.org/citation.cfm?id=1009379.1009660

[62] X3D. (accessed July 2006) WEB3D website. [Online]. Available: http://www.web3d.org

[63] F. Niccolucci, "Xml and the future of humanities computing," *SIGAPP Appl. Comput. Rev.*, vol. 10, pp. 43–47, April 2002. [Online]. Available: http://doi.acm.org/10.1145/568235.568244

[64] R. Hetherington, B. Farrimond, and S. Presland, "Information rich temporal virtual models using x3d," *Computers & Graphics*, vol. 30, no. 2, pp. 287 – 298, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0097849306000355

[65] M. Cabral, M. Zuffo, S. Ghirotti, O. Belloc, L. Nomura, M. Nagamura, F. Andrade, R. Faria, and L. Ferraz, "An experience using x3d for virtual cultural heritage," in *Proceedings of the twelfth international conference on 3D web technology*, ser. Web3D '07. New York, NY, USA: ACM, 2007, pp. 161–164. [Online]. Available: http://doi.acm.org/10.1145/1229390.1229419

[66] M. Sifniotis, "An overview of 3d visualisation using game platforms." *3dVisA bulletin*, vol. 3, pp. –, 2007.

[67] R. Adobbati, A. N. Marshall, A. Scholer, S. Tejada, G. A. Kaminka, S. Schaffer, and C. Sollitto, "Gamebots: A 3D virtual world test-bed for multi-agent research," in *In Proceedings of the Second International Workshop on Infrastructure for Agents, MAS and Scalable MAS*, 2001.

[68] Talin, "Real interactivity in interactive entertainment," *ACM SIGGRAPH Computer Graphics*, vol. 28, no. 2, pp. 97–99, 1994.

[69] A. Amory, K. Naicker, J. Vincent, and C. Adams, "The use of computer games as an educational tool: identification of appropriate game types and game elements," *British Journal of Educational Technology*, vol. 30, no. 4, pp. 311–321, 1999.

[70] S. Clarke-Wilson, "Applying game design to virtual environments," in *Digital illusion: entertaining the future with high technology*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1998, pp. 229–239.

[71] J. Jacobson, "Using CaveUT to build immersive displays with the Unreal tournament engine and a PC cluster," in *Proceedings of the 2003 symposium on Interactive 3D graphics*. ACM Press, 2003, pp. 221–222.

[72] J. Jacobson and J. Vadnal, "Multimedia in three dimensions for archaeology; information retrieval with interactive models," in *Proceedings of the SCI'99/ISAS'99*, 1999.

[73] V. DeLeon and R. Berry, "Bringing VR to the desktop: are you game?" *IEEE Multimedia*, vol. 7, no. 2, pp. 68–72, Apr-Jun 2000.

[74] E. Champion, "Cultural engagement in virtual heritage environments with in-built interactive evaluation mechanisms," in *Proceedings of Presence 2002*, 2002.

[75] ——, "Cultural learning in virtual environments," in *Proceedings of VSMM International Conference on Virtual Systems and Multimedia VSMM '04*, 2004.

[76] ——, "Evaluating a theory of contextual cultural learning through interaction in an online virtual heritage environment," in *to appear in ETCC Fourth International Conference on Educational Technology in Cultural Context: issues of cultural context in the application of educational technology*, Kaohsiung, Taiwan, 5-8 July 2005.

[77] M. Meister and M. Boss, "On using state of the art computer games engines to visualize archaeological structures in interactive teaching and research," in *Proceedings of CAA Computer Applications in Archaeology. Enter the past: The e-way into the four dimensions ofcultural heritage CAA '03*, W. B. J. E. M. GORIANY, Ed., 2003.

[78] M. Anderson, "Computer games and archaeological reconstruction: The low cost vr," in *Proceedings of CAA Computer Applications in Archaeology. Enter the past: The e-way into the four dimensions ofcultural heritage CAA '03*, 2003.

[79] C. Calef, C. Vilbrandt, J. Goodwin, and J. Goodwin, "Making it realtime: Exploring the use of optimized realtime environments for historical simulation and education," in *Proceedings of the International Conference on Museums and the Web*, D. Bearman, Ed., 2002.

[80] "City engine." [Online]. Available: http://www.procedural.com/cityengine/features.html

[81] G. Guidi, B. Frischer, M. D. Simone, A. Cioci, A. Spinetti, L. Carosso, L. L. Micoli, M. Russo, and T. Grasso, "Virtualizing ancient rome: 3d acquisition and modeling of a large plaster-of-paris model of imperial rome," in *Videometrics VIII*, J. A. Beraldin, S. F. El-Hakim, A. Gruen, and J. S. Walton, Eds., vol. 5665, no. 5665.   San Jose, California, USA,: SPIE, 2005, pp. 119–133.

[82] P. Müller, P. Wonka, S. Haegler, A. Ulmer, and L. Van Gool, "Procedural modeling of buildings," *ACM Transactions on Graphics (TOG),*, vol. 25, pp. 614–623, July 2006. [Online]. Available: http://doi.acm.org/10.1145/1141911.1141931

[83] P. Reilly, "Towards a virtual archaeology," in *Proceedings of CAA Computer Applications in Archaeology*, ser. 565, K. Lockyear and S. Rahtz., Eds.   British Archaeological Reports, 1990, pp. 133–139.

[84] ——, "Visualising the problem: advancing graphic systems in archaeological analysis," in *Computing for archaeologists*, S. Ross, J. Moffett, and J. Henderson, Eds.   Oxford University Committee for Archaeology Monograph Number 18. Oxbow Books, Oxford, 1991, pp. 131–152.

[85] ——, "Three-dimensional modelling and primary archaeological data," in *Archaeology and the Information Age*, P. Reilly and S. Rahtz, Eds.   London: Routledge, 1992, pp. 147–173.

[86] M. Gillings, "Virtual archaeologies and the hyper-real," in *Virtual Reality in Geography*, P. Fisher and D. Unwin, Eds.   London: Taylor and Francis, 2002, pp. 17–34, iSBN 0-7484-0905-X.

[87] G. Goodrick and G. Earl, "A manufactured past: virtual reality in archaeology," *Internet Archaeology: Archaeological Informatics*, vol. 15, p. 21, Winter 2004.

[88] J. Barcelo, "Virtual archaeology and artificial intelligence," in *Proceedings of VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage VAST '00*, F. Nicolucci, Ed.   Arezzo: BAR International Series S1075, 24-25 November 2000, pp. 21–28.

[89] M. Sifniotis, P. Watten, K. Mania, and M. White, "Influencing Factors on the Visualisation of Archaeological Uncertainty," in *VAST07: The 8th*

International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage, D. Arnold, F. Niccolucci, and A. Chalmers, Eds. Brighton, UK: Eurographics Association, 2007, pp. 79–85. [Online]. Available: http://www.eg.org/EG/DL/WS/VAST/VAST07/079-085.pdf

[90] M. Sifniotis, K. Mania, P. Watten, and M. White, "Presenting uncertainty in archaeological reconstructions using possibility theory and information visualisation schemes," in *VAST 2006 7th International Symposium on Virtual Reality, Archaelogy and Cultural Heritage-Short papers.*, M. Ioannides, D. Arnold, F. Niccolucci, and K. Mania, Eds., 2006, pp. 198–203.

[91] M. Sifniotis, B. Jackson, M. White, K. Mania, and P. Watten, "Visualising uncertainty in archaeological reconstructions: a possibilistic approach," in *SIGGRAPH '06: ACM SIGGRAPH 2006 Sketches*. New York, NY, USA: ACM, 2006, p. 160.

[92] M. Sifniotis, B. J. C. Jackson, K. Mania, N. Vlassis, P. L. Watten, and M. White, "3d visualization of archaeological uncertainty," in *APGV*, 2010, p. 162.

[93] M. Sifniotis and A. Fellicetti, "Needle in the haystack: finding, indexing and semantically querying image collections of ancient coins," in *Proceedings of the 14th International Conference on Virtual Systems and Multimedia VSMM2008*, Cyprus, October 2008, pp. 165–171.

[94] A. Felicetti and H. Mara, "Semantic Web, Digital Libraries and the future of Cultural Heritage," in *VAST 2008: The 9th International Symposium on Virtual Reality, Archaeology, and Cultural Heritage*, M. Ashley, S. Hermon, A. Proença, and K. Rodriguez-Echavarria, Eds. Braga, Portugal: Eurographics Association, 2008, pp. 117–123. [Online]. Available: http://diglib.eg.org/EG/DL/WS/VAST/VAST08/117-123.pdf

[95] H. Eiteljorg, "Virtual reality and rendering," *CSA Newsletter*, vol. VII, no. 4, pp. –, 1995.

[96] ——, "Photorealistic visualizations may be too good," *CSA Newsletter*, vol. XI, no. 2, pp. –, 1998.

[97] M. Roussou and G. Drettakis, "Photorealism and non-photorealism in virtual heritage representation," in *Proceedings of VAST International Symposium on Virtual Reality, Archaeology and Cultural Heritage including the First*

Eurographics Workshop VAST '03, A.Chalmers, D.Arnold, and F. Niccolucci, Eds. Eurographics, November 2003. [Online]. Available: http://www-sop.inria.fr/reves/publications/data/2003/RD03c

[98] K. Kensek and L. S. Dodd, "On the nature of representation, the san bernardino convent," in *Proceedings of Hawaii International Conference on Arts and Humanities*, 2005.

[99] T. Strothotte, M. Masuch, and T. Isenberg, "Visualizing knowledge about virtual reconstructions of ancient architecture," in *Proceedings of IEEE Computer Graphics International CGI '99*, The Computer Graphics Society. IEEE Computer Society, 1999, pp. 36–43.

[100] T. Zuk, M. Carpendale, and W. Glanzman, "Visualizing temporal uncertainty in 3d virtual reconstructions," in *Proceedings of the 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST 2005)*. The Eurographics Association, 2005, pp. 99–106.

[101] F. Niccolucci and S. Hermon, "A fuzzy logic approach to reliability in archaeological virtual reconstruction," in *Proceedings of the 2004 Computer Applications in Archaeology (CAA) Conference*, 2004.

[102] P. Smets, "Varieties of ignorance and the need for well-founded theories," *Inf. Sci.*, vol. 57-58, pp. 135–144, 1991.

[103] ——, "Application of the transferable belief model to diagnostic problems," *International Journal of Intelligent Systems*, vol. 13, no. 2-3, pp. 127–157, 1998. [Online]. Available: http://dx.doi.org/10.1002/(SICI)1098-111X(199802/03)13:2/3<127::AID-INT3>3.0.CO;2-1

[104] D. Dubois and H. Prade, "Possibility theory, probability theory and multiple-valued logics: A clarification," *Annals of Mathematics and Artificial Intelligence*, vol. 32, no. 1-4, pp. 35–66, 2001.

[105] P. P. Bonissone and R. M. Tong, "Editorial: Reasoning with uncertainty in expert systems," *International Journal of Man-Machine Studies*, vol. 22, no. 3, pp. 241 – 250, 1985. [Online]. Available: http://www.science-direct.com/science/article/B6WGS-4T5C7JK-1/2/9a2a9ab5a7e9b5173bd6d558bb2aa9ca

[106] P. Smets, *Uncertainty Management in Information Systems*, ser. Uncertainty Management in Information Systems.   Dordrecht: Kluwer, 1997, ch. Imperfect Information: Imprecision and Uncertainty", pp. 225–254.

[107] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic: theory and applications*. Prentice Hall PTR, 1995. [Online]. Available: http://books.google.com/books?id=AOhQAAAAMAAJ

[108] N. Gershon, "Visualization of an imperfect world," *Computer Graphics and Applications, IEEE*, vol. 18, no. 4, pp. 43 –45, 1998.

[109] International Organization for Standardization, Geneva, Switzerland, "Guide to the expression of uncertainty in measurement," 1995.

[110] A. T. Pang, C. M. Wittenbrink, and S. K. Lodha, "Approaches to uncertainty visualization," *The Visual Computer*, vol. 13, pp. 370–390, 1997, 10.1007/s003710050111. [Online]. Available: http://dx.doi.org/10.1007/s003710050111

[111] P. Longley, *Geographic information systems and science*, ser. 2nd ed.   Wiley, 2005. [Online]. Available: http://books.google.com/books?id=-FbVI-2tSuYC

[112] W. Thomson, *Popular Lectures and Addresses: Nature Series*.   Institution of Civil Engineers: MacMillan, May 1889, vol. 1, ch. Electrical Units of Measurement, pp. 73–136.

[113] M. Bayes and M. Price, "An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S." *Philosophical Transactions*, vol. 53, pp. 370–418, 1763. [Online]. Available: http://rstl.royalsocietypublishing.org/content/53/370.short

[114] G. Shafer, *A Mathematical Theory of Evidence*.   Princeton: Princeton University Press, 1976.

[115] A. P. Dempster, "Upper and lower probabilities induced by a multivariate mapping," *Annals of Mathematical Statistics*, vol. 38, pp. 325–339., 1967.

[116] P. Walley, "Belief function representations of statistical evidence," *The Annals of Statistics*, vol. 15, no. 4, pp. pp. 1439–1465, 1987. [Online]. Available: http://www.jstor.org/stable/2241684

[117] J. Kohlas, P. Monney, R. Haenni, and N. Lehmann, "Model-based diagnostics using hints," in *Symbolic and Quantitative Approaches to Uncertainty, European Conference ECSQARUŠ95, Fribourg*. Springer, 1995, pp. 259–266.

[118] J. Kohlas and P. Monney, *A mathematical theory of hints: an approach to the Dempster-Shafer theory of evidence*, ser. Lecture notes in economics and mathematical systems. Springer, 1995. [Online]. Available: http://books.google.com/books?id=W_AmAQAAIAAJ

[119] H. N. Hestir and G. Rogers, "A random set formalism for evidential reasoning," in *Conditional Logic in Expert System*, 1991, pp. 309–344.

[120] P. Smets and R. Kennes, "The transferable belief model," *Artificial Intelligence*, vol. 66, pp. 191–234, 1994.

[121] J. Kohlas and P.-A. Monney, "Theory of evidence - a survey of its mathematical foundations, applications and computational aspects," *ZOR- Mathematical Methods of Operations Research*, vol. 39, pp. 35–68, 1994.

[122] L. Zadeh, "Fuzzy sets," *Information Control*, vol. 8, pp. 338–353, 1965.

[123] ——, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets and Systems*, vol. 1, pp. 3–28, 1978.

[124] A. Tversky and D. Kahneman, "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," *Psychological Review*, vol. 90, no. 4, pp. 293–315, 1983. [Online]. Available: http://psycnet.apa.org/index.cfm?fa=search.displayRecord&uid=1984-03110-001

[125] G. Gigerenzer, "The psychology of good judgement: frequency formats and simple algorithms," *Medical Decision Making*, vol. 16, pp. 273–280, 1996.

[126] D. Griffin and A. Tversky, "The Weighing of Evidence and the Determinants of Confidence." *Cognitive Psychology*, vol. 24, no. 3, pp. 411–35, 1992. [Online]. Available: http://www.eric.ed.gov/ERICWebPortal/detail?accno=EJ450925

[127] D. Eddy, "Probabilistic reasoning in clinical medicine: Problems and opportunities." in *Judgment under Uncertainty: Heuristics and Biases*, D. Kahneman, P. Slovic, , and A. Tversky, Eds. Cambridge University Press, 1982, pp. 249–267.

[128] C. Kraemer and M. Weber, "How do people take into account weight, strength and quality of segregated vs. aggregated data? experimental evidence," *Journal of Risk and Uncertainty*, vol. 29, pp. 113–142, 2004, 10.1023/B:RISK.0000038940.62992.1b. [Online]. Available: http://dx.doi.org/10.1023/B:RISK.0000038940.62992.1b

[129] G. Gigerenzer, U. Hoffrage, and H. KleinbÏEĮlting, "Probabilistic mental models: A brunswikian theory of confidence," *Psychological Review*, vol. 98, pp. 506–528, 1991.

[130] E. Raufaste, R. D. S. Neves, and C. Marine, "Testing the descriptive validity of possibility theory in human judgments of uncertainty." *Artif. Intell.*, vol. 148, no. 1-2, pp. 197–218, 2003.

[131] C. Ware, *Information Visualization: Perception for Design*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2004.

[132] R. P. Botchen, D. Weiskop, and T. Ertl, "Texture-based visualization of uncertainty in flow fields," in *IEEE Visualization 2005*, 2005, pp. 647–654.

[133] P. Rheingans and S. Joshi, "Visualization of molecules with positional uncertainty," in *Data Visualization Š99, Proceedings of the Joint EUROGRAPHICS - IEEE TCVG Symposium on Visualization*. Springer-Verlag, 1999, pp. 299–306.

[134] H. Li, C.-W. Fu, Y. Li, and A. Hanson, "Visualizing large-scale uncertainty in astrophysical data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, pp. 1640–1647, November 2007. [Online]. Available: http://dx.doi.org/10.1109/TVCG.2007.70530

[135] C. R. Ehlschlaeger, A. M. Shortridge, and M. F. Goodchild, "Visualizing spatial data uncertainty using animation," *Computers and Geosciences*, vol. 23, no. 4, pp. 387–395, 1997.

[136] T. Zuk, J. Downton, D. Gray, S. Carpendale, and J. Liang, "Exploration of uncertainty in bidirectional vector fields," in *Society of PhotoOptical Instrumentation Engineers (SPIE) Conference Series*, vol. 6809, 2008, published online.

[137] L. Reznik and B. L. Pham, "Fuzzy models in evaluation of information uncertainty in engineering and technology applications," in *10th IEEE International Conference on Fuzzy Systems*. Melbourne, Austraila: IEEE, 2001, pp. 972–975.

[138] B. Pham and R. Brown, "Analysis of visualisation requirements for fuzzy systems," in *Proceedings of the 1st international conference on Computer graphics and interactive techniques in Australasia and South East Asia*, ser. GRAPHITE '03. New York, NY, USA: ACM, 2003, pp. 181–ff. [Online]. Available: http://doi.acm.org/10.1145/604471.604507

[139] H. Griethe and H. Schumann, "The visualization of uncertain data: Methods and problems," in *SimVis*, 2006, pp. 143–156.

[140] T. Zuk and S. Carpendale, "Visualization of uncertainty and reasoning," in *Proceedings of the 8th international symposium on Smart Graphics*, ser. SG '07. Berlin, Heidelberg: Springer-Verlag, 2007, pp. 164–177. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-73214-3_15

[141] C. Johnson and A. Sanderson, "A next step: Visualizing errors and uncertainty," *IEEE Computer Graphics and Applications*, vol. 23, no. 5, pp. 6–10, September/October 2003. [Online]. Available: http://www.sci.utah.edu/publications/crj03/g5visv-lo.pdf

[142] H. Veregin, *Accuracy of Spatial Databases*. London: Taylor and Francis, 1989, ch. Error modeling for the map overlay operation, pp. 3–18.

[143] P. Fisher, *Geographical Information Systems: Principles, Techniques, Management and Applications*. John Wiley and Sons, 1999, vol. 1, ch. Models of uncertainty in spatial data, p. 191Ű205.

[144] ——, "Sorites paradox and vague geographies," *Fuzzy Sets and Systems*, vol. 113, no. 1, pp. 7 – 18, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165011499000093

[145] Y.-C. CHANG, J. R. WRIGHT, and B. A. ENGEL, "Evidential reasoning for assessing environmental impact," *Civil Engineering Systems*, vol. 14, no. 1, pp. 55–78, 1996. [Online]. Available: http://www.tandfonline.com/doi/abs/10.1080/02630259608970210

[146] G. Ferrier and G. Wadge, "An integrated gis and knowledge-based system as an aid for the geological analysis of sedimentary basins," *International Journal of Geographical Information Science*, vol. 11, no. 3, pp. 281–297, 1997.

[147] H. Couclelis, "The certainty of uncertainty: Gis and the limits of geographic knowledge," *Transactions in GIS*, vol. 7, no. 2, pp. 165–175, 2003. [Online]. Available: http://dx.doi.org/10.1111/1467-9671.00138

[148] D. Wheatley and M. Gillings, *Beyond the map: archaeology and spatial technologies*. IOS Press, 2000, ch. Vision, perception and GIS: developing enriched approaches to the study of archaeological visibility.

[149] M. Gillings and D. Wheatley, "Seeing is not believing: unresolved issues in archaeological visibility analysis," in *On the good use of geographical information systems in archaeological landscape studies*, B. SlapŽak, Ed., 2001.

[150] D. Wheatley, "Cumulative viewshed analysis: a gis-based method for investigating intervisibility, and its archaeological application," in *Archaeology and GIS: A European Perspective*, G.Lock and Z.Stancic, Eds. London: Rout, 1995.

[151] P. Fisher, "First experiments in viewshed uncertainty: Simulating fuzzy viewsheds," *Photogrammetric Engineering and Remote Sensing*, vol. 58, no. 4, pp. 345–352, 1992.

[152] S. Canning, "'belief' in the past: Dempster-shafer theory, gis and archaeological predictive modelling," *Australian Archaeology*, vol. 60, pp. 6–15, 2005.

[153] A. Comber, P. Fisher, and A. Brown, *Quality Aspects in Spatial Data Mining*. CRC Press, 2008, ch. Uncertainty, Vagueness, and Indiscernibility: The Impact of Spatial Scale in Relation to the Landscape Elements, p. 239Ű250.

[154] T. Hengl, "Visualisation of uncertainty using the hsi colour model: computations with colours," in *Proceedings of the 7th International Conference on GeoComputation*, 2003, pp. 8–17.

[155] A. M. MacEachren, A. Robinson, S. Hopper, S. Gardner, R. Murray, M. Gahegan, and E. Hetzler, "Visualizing geospatial information uncertainty: What we know and what we need to know," *Cartography and Geographic Information Science*, vol. 32, no. 3, pp. 139–160, July 2005.

[156] G. Gilbert, *Law of Evidence*, Dublin, 1754.

[157] C. Moore, *A treatise on facts : or, The weight and value of evidence*. Northpoint, Long Island, N.Y. : E. Thompson Co., 1908.

[158] S. D. A. Anderson, T. J. and W. L. Twining, *Analysis of Evidence*. Cambridge: Cambridge University Press, 2005.

[159] J. H. Wigmore, *The Principles of Judicial Proof as given by Law, Psychology and General Experience*, 1913.

[160] R. Eggleston, *Evidence, proof, and probability*, ser. Law in context. Weidenfeld and Nicolson, 1983. [Online]. Available: http://books.google.com/books?id=G56FQgAACAAJ

[161] D. Schum, *The evidential foundations of probabilistic reasoning*. Northwestern University Press, 2001. [Online]. Available: http://books.google.com/books?id=8PqsTNULCPEC

[162] P. Murphy and P. Murphy, *Evidence, proof, and facts: a book of sources*. Oxford University Press, 2003. [Online]. Available: http://books.google.com/books?id=srFqqNvV5k0C

[163] A. P. Dawid, *Aspects of Uncertainty: A Tribute to D. V. Lindley,.* Chichester: John Wiley and Sons, 1993, ch. The island problem: coherent use of identification evidence. In, pp. 159–170.

[164] A. P. Dawid and J. Mortera, "Coherent analysis of forensic identification evidence," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 2, pp. pp. 425–443, 1996. [Online]. Available: http://www.jstor.org/stable/2345986

[165] A. P. Dawid, "Bayes's theorem and weighing evidence by jurie," in *In Proceedings of the British Academy*, R. Swinburne, Ed., vol. 113. Oxford University Press, 2002, pp. 71 – 90.

[166] D. A. Schum, "Knowledge, probability, and credibility," *Journal of Behavioral Decision Making*, vol. 2, no. 1, pp. 39–62, 1989. [Online]. Available: http://dx.doi.org/10.1002/bdm.3960020104

[167] D. A. Schum and J. R. Morris, "Assessing the competence and credibility of human sources of intelligence evidence: contributions from law and probability," *Law, Probability and Risk*, vol. 6, pp. 247–274, 2007.

[168] D. A.F., G. P., and N. E., "Belief revision as applied within a descriptive model of jury deliberations," *Information & Communications Technology Law*, vol. 10, no. 1, pp. 53–65, 2001.

[169] A. Josang and V. Bondi, "Legal reasoning with subjective logic," *Artificial Intelligence and Law*, vol. 8, pp. 289–315, 2000.

[170] C. M. Yablon, "On the allocation of burdens of proof in corporate law: An essay on fairness and fuzzy sets;," *Cardozo Law Review*, vol. 13, no. 517, pp. 497–518, 1991.

[171] D. F. Chavkin, "Fuzzy thinking: A borrowed paradigm for crisper lawyering," *Clinical Law Review*, vol. 4, no. 163, pp. 163–194, 1997.

[172] , "Criminal appeal report 369: R v doheny and adams," 1997.

[173] T. W., "Evidence as a multidisciplinary subject," *Law, Probability and Risk*, vol. 2, pp. 91–107, 2003.

[174] W. Twining, *Rethinking evidence: exploratory essays*, ser. Law in context. Cambridge University Press, 2006. [Online]. Available: http://books.google.com/books?id=A_DRugC_bRwC

[175] J. Light, Donald, "Uncertainty and control in professional training," *Journal of Health and Social Behavior*, vol. 20, no. 4, pp. pp. 310–322, 1979. [Online]. Available: http://www.jstor.org/stable/2955407

[176] R. Fox, *Training for Uncertainty*. Cambrirdge, 1957, ch. The Student-Physician, pp. 207–241.

[177] R. S. Ledley and L. B. Lusted, "Reasoning foundations of medical diagnosis," *Science*, vol. 130, no. 3366, pp. pp. 9–21, 1959. [Online]. Available: http://www.jstor.org/stable/1758070

[178] G. T. O'Connor and H. C. Sox, "Bayesian reasoning in medicine," *Medical Decision Making*, vol. 11, no. 2, pp. 107–111, 1991. [Online]. Available: http://mdm.sagepub.com/content/11/2/107.abstract

[179] R. Miller and A. Geissbuhler, *Clinical decision support systems: theory and practice*, ser. Health informatics. Springer, 1999, ch. Clinical diagnostic decision support systems: an overview, pp. 3–30. [Online]. Available: http://books.google.co.uk/books?id=2iqBv92loHoC

[180] M. A. Musen, Y. Shahar, and E. H. Shortliffe, "Clinical decision-support systems," in *Biomedical Informatics*, ser. Health Informatics, K. J. Hannah, M. J.

Ball, E. H. Shortliffe, and J. J. Cimino, Eds. Springer New York, 2006, pp. 698–736. [Online]. Available: http://dx.doi.org/10.1007/0-387-36278-9_20

[181] E. H. Shortliffe and B. G. Buchanan, "A model of inexact reasoning in medicine," *Mathematical Biosciences*, vol. 23, no. 3-4, pp. 351 – 379, 1975. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0025556475900474

[182] F. T. de Dombal, D. J. Leaper, J. R. Staniland, A. P. McCann, and J. C. Horrocks, "Computer-aided diagnosis of acute abdominal pain." *British medical journal*, vol. 2, no. 5804, pp. 9–13, 1972.

[183] I. D. Adams, M. Chan, P. C. Clifford, W. M. Cooke, V. Dallos, F. T. de Dombal, M. H. Edwards, D. M. Hancock, D. J. Hewett, and N. McIntyre, "Computer aided diagnosis of acute abdominal pain: a multicentre study." *Br Med J (Clin Res Ed)*, vol. 293, no. 6550, pp. 800–804, 1986. [Online]. Available: http://view.ncbi.nlm.nih.gov/pubmed/3094664

[184] R. A. Miller, H. E. Pople, and J. D. Myers, "Internist-i, an experimental computer-based diagnostic consultant for general internal medicine," *New England Journal of Medicine*, vol. 307, no. 8, pp. 468–476, 1982. [Online]. Available: http://www.nejm.org/doi/full/10.1056/NEJM198208193070803

[185] L. Kohout, I. Stabile, W. Bandler, and J. Anderson, *Comparative approaches to medical reasoning*. World Scientific, 1995, ch. Clinaid: Medical knowledge-based system based on fuzzy relational structures, pp. 1–25.

[186] R. W. Jones, A. Lowe, and M. J. Harrison, "A framework for intelligent medical diagnosis using the theory of evidence," *Knowledge-Based Systems*, vol. 15, no. 1-2, pp. 77 – 84, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S095070510100123X

[187] P. Smets, *Belief Functions*. London: Academic Press, 1988, ch. Non Standard Logics for Automated Reasoning., pp. 253–286.

[188] M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence*, vol. 7, pp. 225–240, 1998.

[189] M. A. Hearst and M. Y. Ivory, "The state of the art in automating usability evaluation of user interfaces," *ACM Computing Surveys*, vol. 33, pp. 470–516, 2001.

[190] A. Whitefield, F. Wilson, and J. Dowell, "A framework for human factors evaluation," *Behaviour & Information Technology*, vol. 10, no. 1, pp. 65–79, 1991. [Online]. Available: http://www.informaworld.com/10.1080/01449299108924272

[191] S. Charlton and T. O'Brien, Eds., *Handbook of human factors testing and evaluation*. Mahwah, N.J.: Lawrence Erlbraum Associates, 2002.

[192] B. A. Babbit and C. O. Nystrom, "Questionnaire construction manual," U.S. Army Research Institute for the behavioural and social sciences, Fort Hood, TX, Research Product 89-20, June 1989.

[193] R. Likert, "A technique for the measurement of attitudes," *Archives of Psychology*, vol. 140, no. 22, p. 55, 1932.

[194] OED. (2009) Oxford english dicitonary: Electronic edition. [Online]. Available: http://www.oed.com/

[195] E. Schechter, *Classical and Nonclassical Logics: An Introduction to the Mathematics of Propositions*. Princeton University Press, 2005.

[196] Aristotle, *Metaphysica. Books Gamma, Delta and Epsilon.*, C. Kirwan, Ed. Clarendon, 1971.

[197] B. Gaines and L. Kohout, "Possible automata," in *Proceedings of the Internatonal Symposium on Multiple-Valued Logic*, 1975, pp. 183–196.

[198] D. Dubois and H. Prade, *Possibility Theory: An Approach to Computerized Processing of Uncertainty*. Plenum Press, New York, 1988.

[199] D. Dubois, H. Prade, and P. Smets, "A definition of subjective possibility," *Int. J. Approx. Reasoning*, vol. 48, no. 2, pp. 352–364, 2008.

[200] D. Dubois, H. Nguyen, and H. Prade, *Fundamentals of fuzzy sets: the handbook of Fuzzy Sets Series*. Kluwer, 2000, ch. Possibility theory probability and fuzzy sets: misunderstandings, bridges and gaps, pp. 343–438.

[201] I. Hacking, "All kinds of possibility," *The Philosophical Review*, vol. 84, no. 3, pp. 321–337, 1975. [Online]. Available: http://www.jstor.org/stable/2184116

[202] D. Dubois and H. Prade, Eds., *Fundamentals of Fuzzy Sets* , ser. The Handbooks of Fuzzy Sets Series. Boston, Mass.: Kluwer, 2000.

[203] Y. Y. Chen, "Bernoulli trials: From a fuzzy measure point of view," *Journal of Mathematical Analysis and Applications*, vol. 175, no. 2, pp. 392 – 404, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/ B6WK2-45P1529-88/2/fded75c59fd77d81389c2c2060c97eeb

[204] ——, "Statistical inference based on the possibility and belief measures," *Transactions of the American Mathematical Society*, vol. 347, pp. 1855–1863, 1995.

[205] P. Smets and R. Kruse, *Uncertainty Management in Information Systems*, ser. Uncertainty Management in Information Systems. Dordrecht: Kluwer, 1997, ch. The Transferable Belief Model for Belief Representation, pp. 343–363.

[206] P. Smets, "Non standard probabilistic and non probabilistic representations of uncertainty," in *Advances in Intelligent Computing âĂŤ IPMU '94*, ser. Lecture Notes in Computer Science, B. Bouchon-Meunier, R. Yager, and L. Zadeh, Eds. Springer Berlin / Heidelberg, 1995, vol. 945, pp. 13–38, 10.1007/BFb0035934. [Online]. Available: http://dx.doi.org/10.1007/BFb0035934

[207] G. Shafer, "Belief functions and parametric models," *J.R.Statist. Soc B*, vol. 44, no. 3, pp. 322–352, 1982. [Online]. Available: http://www.glennshafer.com/ assets/downloads/articles/article09_belief3.pdf

[208] P. Smets, *What is Dempster-Shafer's model?* New York, NY, USA: John Wiley & Sons, Inc., 1994, pp. 5–34. [Online]. Available: http://portal.acm.org/citation. cfm?id=186965.186966

[209] ——, "The combination of evidence in the transferable belief model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 447–458, 1990.

[210] ——, "Resolving misunderstandings about belief functions," *Int. J. Approx. Reasoning*, vol. 6, pp. 321–344, May 1992. [Online]. Available: http: //portal.acm.org/citation.cfm?id=149073.149076

[211] ——, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, ser. Quantified Representation of Uncertainty & Imprecision. Doordrecht: Kluwer, 1998, vol. 1, ch. Smets Ph.(1998) Probability, Possibility, Belief: Which and Where ?, pp. 11–24.

[212] F. Delmotte and P. Smets, "Target identification based on the transferable belief model interpretation of dempster-shafer model," *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, vol. 34, no. 4, pp. 457–471, 2004.

[213] F. Delmotte and P. . Smets, "Target identification based on the transferable belief model interpretation of dempster-shafer model. pars i: Methodology." INRIA, Tech. Rep., 2001.

[214] P. Smets, "Belief functions: The disjunctive rule of combination and the generalized bayesian theorem," *International Journal of Approximate Reasoning*, vol. 9, no. 1, pp. 1 – 35, 1993. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0888613X9390005X

[215] S. Hermon, *Beyond Illustration: 2D and 3D Digital Technologies as Tools for Discovery in Archaeology.* London: British Archaeological Reports International Series, 2008, ch. Reasoning in 3D: a critical appraisal of the role of 3D modelling and virtual reconstructions in archaeology, pp. 70–84.

[216] J. Manley and D. Rudkin, *Facing the Palace: excavations just east of Fishbourne Roman Palace (Sussex, UK) 1995-1999,*, J. Manley and D. Rudkin, Eds. Sussex Archaeological Collections (SAC) 141, 2005.

[217] B. Jackson, "High level message transfers for the graphical components of virtual prototyping simulations," Ph.D. dissertation, School of Informatics, University of Sussex, Brighton., 2009.

[218] G. N. M. de Carvalho, T. Gill, and T. Parisi, "X3d programmable shaders," in *Web3D '04: Proceedings of the ninth international conference on 3D Web technology.* New York, NY, USA: ACM Press, 2004, pp. 99–108.

[219] J.-B. Y. Guilan Kong, Dong-Ling Xu, "Clinical decision support systems: a review of knowledge representation and inference under uncertainties," *International Journal of Computational Intelligence Systems*, vol. 1, no. 2, pp. 159 – 167, 2008.

[220] CheckMend. (2008) Checkmend. Checkmend progress report. [Online]. Available: http://www.checkmend.co.uk/pdf/CheckmendPressRelease20080518.pdf

[221] K. Wutkowski. (2007) U.S. retailers want online sellers to fight theft. [Online]. Available: http://www.reuters.com/article/technologyNews/idUSN2560973520071025

[222] J. E. Conklin, *Art crime.*   Praeger, Westport, Conn. :, 1994.

[223] P. W. Todeschini, "The Medici conspiracy: The illicit journey of looted antiquities from italyŠs tomb raiders to the worldŠs greatest museums." *British Journal of Criminology*, vol. 47, no. 2, pp. 367–370, 2007.

[224] N. T. Elkin. (2007) Why coins matter: Trafficking in undocumented and illegally exported ancient coins in the North American marketplace. [Online]. Available: http://www.savingantiquities.org/feature_coins.php

[225] Y. Jing and S. Baluja, "Pagerank for product image search," in *World Wide Web WWW 2008*, Beijing, Chaina, April 2008, pp. 307–315. [Online]. Available: http://www2008.org/papers/fp506.html

[226] Nutch. (2008) Nutch engine homepage. Retrieved November 20, 2008 from http://lucene.apache.org/nutch. [Online]. Available: http://lucene.apache.org/nutch

[227] M. Cafarella and D. Cutting, "Building nutch: Open source search: A case study in writing an open source search engine," *ACM Queue*, vol. 2, no. 2, pp. 54–61, April 2004.

[228] L. Benedict. (2004) Comparison of Nutch and Google Search Engine: Implementations on the Oregon State University Website. [Online]. Available: http://www.misterbot.fr/OSU_Queries.pdf

[229] Z. Zhang, C. Rojas, O. Nasraoui, and H. Frigui, "Show and tell: A seamlessly integrated tool for searching with image content and text abstract," ACM-SIGIR Open Source Information Retrieval workshop., 2006.

[230] M. Lux and S. A. Chatzichristofis, "LIRE: lucene image retrieval: an extensible java cbir library," in *MM '08: Proceeding of the 16th ACM international conference on Multimedia*.   New York, NY, USA: ACM, 2008, pp. 1085–1088.

[231] Lucene. (2008) Apache Lucene home page. [Online]. Available: http://lucene.apache.org/java/docs/index.html

[232] ISO:21127:2006, "Information and documentation – a reference ontology for the interchange of cultural heritage information," 2006.

[233] C. J.-A. G. P. M. Kohlhase, S. Anca and I. A. Sucan., "Mathwebsearch 0.4, a semantic search engine for mathematics." 2008.

[234] L. Zhuhadar and O. Nasraoui, "Semantic information retrieval for personalized e-learning," *Tools with Artificial Intelligence, IEEE International Conference on*, vol. 1, pp. 364–368, 2008.

[235] G. V. A. Bonomi, G. Mantegari, "A framework for ontological description of archaeological scientific publications," in *Semantic Web Applications and Perspectives 2006, Proceedings of SWAP 2006, the Third Italian Semantic Web Workshop*, O. S. G. Tumarello, P. Bouquet, Ed., vol. 201, Pisa, Italy, 2006.

[236] T. Lukasiewicz and U. Straccia, "Managing uncertainty and vagueness in description logics for the semantic web," *Web Semant.*, vol. 6, pp. 291–308, November 2008. [Online]. Available: http://dl.acm.org/citation.cfm?id=1464505.1464603

[237] W3C, "Uncertainty reasoning for the world wide web," Online, March 2008. [Online]. Available: http://www.w3.org/2005/Incubator/urw3/XGR-urw3-20080331/#belieffunctions

[238] J. Thomson, E. Hetzler, A. MacEachren, M. Gahegan, and M. Pavel, "A typology for visualizing uncertainty," in *Visualization and Data Analysis. SPIE 2005.*, ser. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, M. G.-K. B. R. Erbacher, J. Roberts, Ed., vol. 5669, Mar. 2005, pp. 146–157.

[239] M. Doerr, "Ontologies for cultural heritage," in *Handbook on Ontologies*, ser. International Handbooks on Information Systems, S. Staab and D. Rudi Studer, Eds. Springer Berlin Heidelberg, 2009, pp. 463–486.

[240] J. Yang, Y. Wang, D. Xu, and K. Chin, "The evidential reasoning approach for mada under both probabilistic and fuzzy uncertainties," *European Journal of Operational Research*, vol. 171, no. 1, pp. 309 – 343, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0377221704006277

[241] S. Haegler, P. Müller, and L. Van Gool, "Procedural modeling for digital cultural heritage," *J. Image Video Process.*, vol. 1, pp. 1–12, February 2009. [Online]. Available: http://dx.doi.org/10.1155/2009/852392

[242] D. Tudhope, C. Binding, S. Jeffrey, K. May, and A. Vlachidis, "A stellar role for knowledge organization systems in digital archaeology," *Bulletin of the American Society for Information Science and Technology*, vol. 37, no. 4, pp. 15–18, 2011. [Online]. Available: http://dx.doi.org/10.1002/bult.2011.1720370405

[243] W. Moussa and D. Fritsch, "A simple approach to link 3d photorealistic models with content of bibliographic repositories," in *Proceedings of the Third international conference on Digital heritage*, ser. EuroMed'10. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 482–491. [Online]. Available: http://dl.acm.org/citation.cfm?id=1939603.1939651

[244] H. K. Philip Verhagen and M. van Leusen, *Archaeological Prediction and Risk Management: Alternatives to Current Practice*, ser. AUP - Leiden University Press. Amsterdam University Press, 2010, ch. The future of archaeological predictive modelling, pp. 19–26. [Online]. Available: http://books.google.com/books?id=GC53yC4wUGcC

APPENDIX A

Questionnaire

Introduction to the questionnaire

The following questions are related to the reconstruction and interpretation of buildings and structures in general; please keep this in mind when giving your answers. The following questions are applied to cases where you have collected all available data from a site.

Please state your opinion as in the following example:

1. I often consider two or more valid interpretations.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

The circled answer indicates that you seldom encounter the above scenario.

Answer the questions independently in the order that they appear.
Do not skip questions or return to a previous question to change your answer.

The following questions explore the notion of uncertainty through archaeological interpretation

ARCHAEOLOGICAL UNCERTAINTY

1. Archaeology is fraught with uncertainty: You can never be too sure about anything.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

2. There is only one true interpretation of a site.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

3. Other alternatives to my interpretation could be equally viable.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

4. I often consider two or more valid interpretations.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

5. I often consider two or more valid interpretations and present one.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

6. I always end up with one interpretation and no alternatives.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

7. I feel sure about my interpretation given the available evidence.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

The following questions relate to the interpretation of features and structural evidence; these indicate any non-portable remnant of human activity whether a structure, ditch, post-hole, etc.

FEATURES

1. While archaeology is uncertain, structural evidence gives me security.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

2. When faced with structural evidence I usually make up my mind very quickly as to how the building looked.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

3. I try to understand the function of the building (i.e. what it was used for) and make my reconstruction around it.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

4. First I interpret the structural evidence and then I deduce the building's function.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

5. Most times I am more sure about the ground floors than an existence of a second storey.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

6. Interpreting basic structural shapes (e.g. walls) is easy even when you have just the foundations

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

7. I am comfortable with placing exits in a building when the structural evidence suggests it.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

Continues overleaf…

8. I am comfortable with placing exits in a building in the absence of structural evidence.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

The following questions relate to the interpretation of:

- *artefacts: objects made or modified by human culture*
- *biofacts; objects found at an archaeological site and carrying archaeological significance, but not altered by human hands (e.g. wood, bone, etc).*

ARTEFACTS

1. Scarcity of artefacts in a site greatly hinders my interpretation of the site

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

2. Interpreting the function of a room mostly depends on what artefacts I have found.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

3. Artefacts mostly help in the details of a structure (e.g. roof style, room decorations)

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

4. I can interpret the look of a room without any artefacts.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

BIOFACTS

5. Scarcity of biofacts in a site greatly hinders my interpretation

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

6. Interpreting the function of a room mostly depends on what biofacts I have found.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

7. I can interpret the look of a room without any biofacts.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

The following questions relate to Comparative data. This means comparison with other cases and the environment. Comparative data includes:

- *Absolute comparisons: compare with similar buildings*
- *Contextual comparisons: examines the context of the building (e.g. its surrounding) as well as the building itself with similar contexts*
- *Topography: The study of the surrounding environment; formulations of the ground, earth features, etc.*

COMPARATIVE DATA

1. I often base my interpretation on the surrounding topography of the area

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

2. I rely on similar structures to establish a layout of the site.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

3. Documented structures (e.g. in archaeological publications) contemporary to the one I work on greatly influence my interpretation.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

4. I often change my interpretation of the structural features when faced with contradictory contextual comparisons.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

5. The surrounding environment greatly influences the structure and layout of a building.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

6. I often place exits on a building according to the surrounding landscape.

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

7. I will change my interpretation of the structural features if textual evidence suggests it.

Continues overleaf…

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

The following questions relate to peer review and explore the depth of discussing results and interpretations with one's peers.

PEER REVIEW

1. I find it useful to discuss my interpretations with my peers

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

2. I rely upon peer interpretations and comments

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

3. I sometimes find that my peers' opinions change my interpretation

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

4. Peer discussion results in better interpretations and results

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

5. How often do you discuss interpretations with your peers?

Never	Almost never	Seldom	Sometimes	Often	Very often	Always

Consider the following types of evidence. Please circle the following combinations in terms of which you think give stronger evidence for your reconstructions. The grading is from 1 to 9; 1 being the weakest and 9 the strongest, as in the following example:

Artefacts & Peer input   weak   1 2 3 4 5 6 (7) 8 9   strong

This would indicate that the above combination tends to be quite strong.

EVIDENCE COMBINATIONS

Features & Artefacts	weak	1 2 3 4 5 6 7 8 9 strong
Features & Biofacts	weak	1 2 3 4 5 6 7 8 9 strong
Features & topography	weak	1 2 3 4 5 6 7 8 9 strong
Features & absolute comparisons	weak	1 2 3 4 5 6 7 8 9 strong
Features & contextual comparison	weak	1 2 3 4 5 6 7 8 9 strong
Features & textual evidence	weak	1 2 3 4 5 6 7 8 9 strong
Features & peer input	weak	1 2 3 4 5 6 7 8 9 strong

Artefacts & biofacts	weak	1 2 3 4 5 6 7 8 9 strong
Artefacts & topography	weak	1 2 3 4 5 6 7 8 9 strong
Artefacts & absolute comparisons	weak	1 2 3 4 5 6 7 8 9 strong
Artefacts & contextual comparison	weak	1 2 3 4 5 6 7 8 9 strong
Artefacts & textual evidence	weak	1 2 3 4 5 6 7 8 9 strong
Artefacts & peer input	weak	1 2 3 4 5 6 7 8 9 strong

Biofacts & topography	weak	1 2 3 4 5 6 7 8 9 strong
Biofacts & absolute comparisons	weak	1 2 3 4 5 6 7 8 9 strong

Biofacts & contextual comparison	weak	1 2 3 4 5 6 7 8 9	strong
Biofacts & textual evidence	weak	1 2 3 4 5 6 7 8 9	strong
Biofacts & peer input	weak	1 2 3 4 5 6 7 8 9	strong

Topography & absolute comparisons	weak	1 2 3 4 5 6 7 8 9	strong
Topography & contextual comparison	weak	1 2 3 4 5 6 7 8 9	strong
Topography & textual evidence	weak	1 2 3 4 5 6 7 8 9	strong
Topography & peer input	weak	1 2 3 4 5 6 7 8 9	strong

Consider the following types of evidence usually available in an excavation and how important they usually are in your interpretations of a site.

 Please read the complete list first.

You are asked to rank the evidence in order of importance, from 1 to 8; 1 being the lowest (least important) and 8 the highest (most important). Use a number only once.

EVIDENCE TYPE RANKS

Evidence type	Score
Features	
Artefacts	
Biofacts	
Topography	
Contextual comparisons	
Absolute comparisons	
Textual evidence	
Peer input/outside support	

OPEN QUESTIONS

Through this questionnaire you have been asked to think about different categories of evidence that turn up in an excavation. The categories were:

- Features
- Artefacts
- Biofacts
- Topography
- Contextual comparisons
- Absolute comparisons
- Textual evidence
- Peer input/outside support

You will now be asked a few questions regarding the questionnaire and its contents. Please feel free to add your comments.

OPEN QUESTIONS

Please circle the correct answer.

Do you consider the categories of evidence list to be complete? Yes   No

If you have answered *NO*, please use the following space to express a more complete opinion and include any other possible categories:

_____

_____

_____

_____

_____

_____

_____

_____

_____

Was the questionnaire difficult to fill in?   Yes   No

Did it take a lot of time to fill in?   Yes   No

Did you find the questionnaire tiring?   Yes   No

Did you find the questionnaire offensive?   Yes   No

If you have answered yes to any of the above question you may use the following space to express a more complete opinion.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Finally, you may use the space to express any additional comments about the questionnaire.

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

Please Use Overleaf If Necessary

Please fill in the following information:

Name: _____

E-mail: _____

Age Group: less than 25
                        26-33
                        34-41
                        42-49
                        50-57
                        58-65
              more than 65

Gender: Male / Female

Date and Time: _____   _____

Years of excavation expertise:
        less than 10   11-20   21-30   31-40   more than 40

Major area of excavation expertise:

_____

Current Status: Technical   Research Fellow   Academic
                      Other (please specify):_____

To what extend do you use a   Not at all   Very much
computer in your daily activities?   1   2   3   4   5

THANK YOU! ☺