Chapter 6

Empirical evidence of learning through Empirical Modelling

The previous chapters have indicated respects in which Empirical Modelling is intimately linked with learning activity of many different varieties. In this chapter, I will examine further informal evidence in support of this claim that is drawn from various student projects in Empirical Modelling. An assessment exercise attached to the 'Introduction to Empirical Modelling' module is one source of evidence for learning in student projects. This assessment takes the form of an open-ended modelling and paper-writing exercise. Such an exercise is shown to be effective for learning about Empirical Modelling. It also promotes self-motivated exploration in unknown domains that is one of the key skills for experiential and lifelong learning. The extent to which students not only learnt about Empirical Modelling, but also about the domain which they chose to model was unexpected. This leads to the suggestion that Empirical Modelling could be effective in facilitating learning in other domains.

Further evidence of learning through EM can be found in third year projects. These projects are significantly larger, taking up most of a student's academic year and contributing to a quarter of their mark for the year. I shall examine two specific projects in detail in an attempt to understand on a deeper level the ingredients that give EM potential for learning in other domains.

6.1 A module for learning about EM[†]

'Introduction to Empirical Modelling' is a module that has been run in the last 4 years for final year undergraduates on the 4-year MEng Computer Science course at the University

[†]This section takes the joint paper presented at ICALT 2006 [BHB06] and builds on the findings with more recent data.

of Warwick. The module aims to teach the basics of EM as briefly described in Chapter 2 of this thesis, and covers the application of EM in a wide range of areas such as artificial intelligence, computer graphics, concurrent systems, human computing, and educational technology. In 2004-5, a new form of assessment for the module was introduced. This involved the informal publication of an online journal 'the first Warwick Electronic Bulletin on EM (WEB-EM-1)' [WEBEM1] to which the students were required to submit papers and associated EM artefacts.

6.1.1 The module assessment

The module ran for 10 weeks, with 2 hours of lectures and 1 hour of computer laboratory sessions per week (with extra laboratory time available). Students were introduced to the concepts of observation, dependency, and agency that they were expected to use when analysing their problem domain. The laboratory gave students experience of the principal EM tool, the `tkeden` interpreter, with its associated family of built-in notations for framing dependencies between scalars, strings, lists, geometric entities and screen display elements. The module also introduced the LSD notation for accounting for inter-agent communication (as discussed in §5.4), and other features of `tkeden` such as agent-oriented parsing [Har03] and models to depict networks of dependencies [EMP:dmtHarfield2006].

For the module assessment, we issued a Call for Papers requiring two submissions. Students first submitted a paper title and abstract. These submissions were reviewed and feedback was given. Students subsequently submitted their full paper and accompanying model.

The coursework had two objectives. The first was to assess the students' understanding of Empirical Modelling through written and modelling exercises based on a common theme of the students' own choice. The second was to equip the students with basic research skills that would be useful in further education. In the Call for Papers, we requested that students submit original and high quality papers relating to Empirical Modelling and its applications supported by a relevant documented modelling study.

6.1.2 Previous assessments

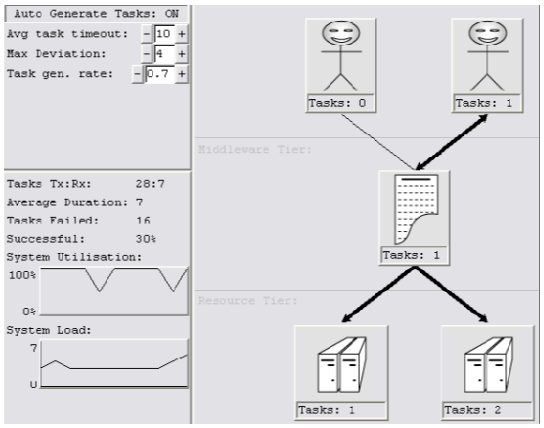
In the academic years 2002-03 and 2003-04, the coursework assessment required students to build a model using the EM tools. The students were all given the same task in 2002-03 to implement a board game, and in 2003-04 to make a model of a heating system. Al-

though many good submissions reflected the hard work of the students it was felt that the scope was not wide enough for capable students. On the evidence of their submissions, many students were keen to put effort into developing their submissions beyond the original specification even though their knowledge of board games and the workings of heating systems was limited. On that basis, it seemed natural to give students the opportunity to apply the EM tools and principles covered in the module to a subject of their choosing. Since many fourth year students are likely to proceed to research the open-ended assessment was also seen as a good way to promote research skills that would assist their future studies.

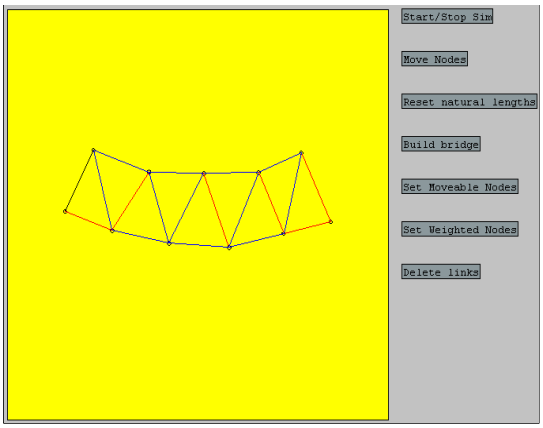
6.1.3 Submissions

The selection of the submissions to which I shall refer in the following section is listed below. A complete list of submissions is available from the first Warwick Electronic Bulletin on Empirical Modelling [WEBEM1]. Screen-shots from the models of some of the following submissions are shown in Figure 6.1.

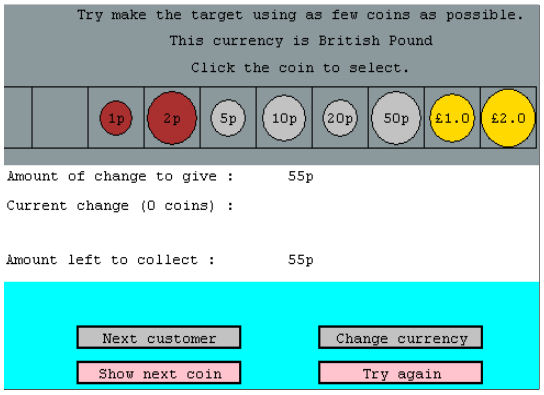
- Tournament. A notation and model for the organisation of knock-out tournaments.
- IceCube. An exploration of IceCube, a technology developed by Microsoft that deals with reconciling divergent replicas of some shared system state.
- Grid computing. A simulation to allow exploration of the efficiency of a computing grid (Figure 6.1(a)).
- Bridges. A model exploring basic engineering principles behind bridge building (Figure 6.1(b)).
- Greedy algorithms. A learning artefact to demonstrate a greedy algorithm for the 'making change' problem with different currencies (Figure 6.1(c)).
- Non-decimal bases. A learning artefact to aid the understanding of non-decimal bases (Figure 6.1(d)).
- Wumpus. A model that illustrates the game of 'Hunt the Wumpus' first introduced in AI research. (Figure 6.1(e)).
- Poker. A model studying the communication of information in a game of poker.
- Frisbee. A model exploring the interaction in a game of frisbee.



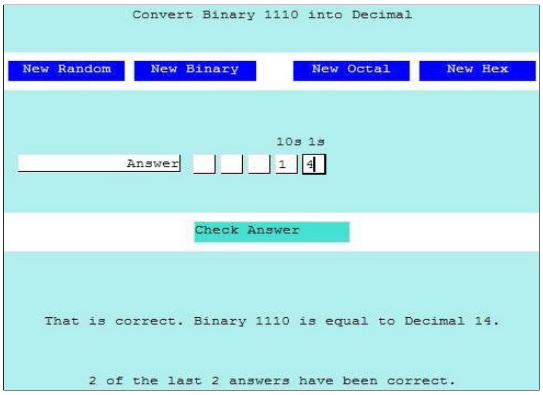
(a) Simulating grid computing



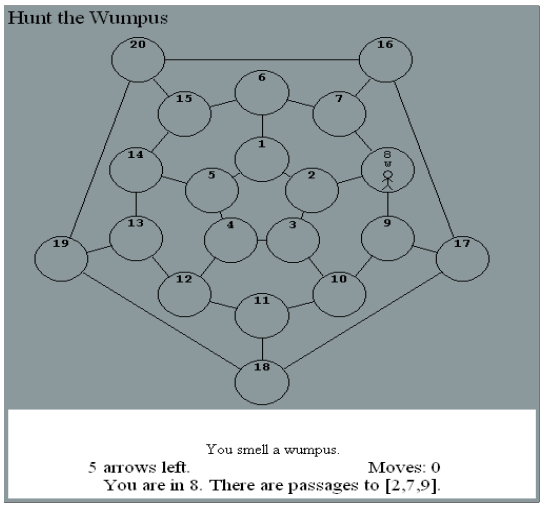
(b) Building bridges



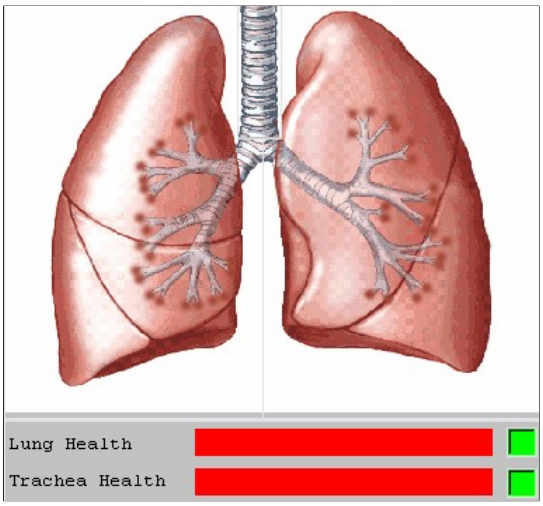
(c) Greedy algorithms



(d) Teaching non-decimal bases



(e) Hunt the wumpus AI game



(f) Learning about human biology

Figure 6.1: A selection of models from WEB-EM-1.

- Human Biology. A dependency-based simulation to illustrate how the lungs function (Figure 6.1(f)).

6.1.4 Marking

Out of the 25 abstracts initially submitted and approved, all but six led on to final submissions to WEB-EM-1. The analysis is based on the final submissions. Each submission comprised of a model and a paper; in the assessment process, these were marked together. The marks served as a good discriminate of skill and understanding in EM, lying in the range 45-80%, with an overall average of 63%.

6.1.5 Analysis

This section describes aspects of learning that were highlighted by the assessment:

- Learning can occur and skills can be developed without a preconceived objective
- Learning is stimulated by personal interest.
- Learning is reinforced when practice and principles are combined.
- Learning is aided by exploration.

In the following section I shall consider respects in which Empirical Modelling is well-suited to supporting learning that exhibits these characteristics.

6.1.5.1 Learning can occur and skills can be developed without a preconceived objective

As in previous years, the coursework helped to develop practical skills with the EM tools. However, the potential for emphasising different aspects of Empirical Modelling was apparent with this new style of assessment. Some students stuck to the basic tools whilst others made use of other, often more technical, tools and notations. The Frisbee model made use of only of the basic notations for data manipulation and line drawing, introduced at the beginning of the course, and the student was clearly proficient in building models with these notations. Other models, such as Tournament, involved the development of special-purpose notations which exercised a different skill-set associated with agent-oriented parsing. Another student modelled the game of poker from different viewpoints using the distributed EM tool. Others emphasised the incremental aspect of model-building in their model. For example, the Making Change model used incremental development to show

how a learning artefact might be adapted to situations that arise and evolve as the understanding of the learner develops. Each student developed the same basic skills but some also demonstrated extra skills. The fact that students could choose what skills to develop within certain broad constraints contributed to the diversity and richness of the submissions.

The journal-style of assessment demands a different skill set from the typical computer science coursework. Coursework is usually a specific task which the student should tackle in a preconceived way and hence often results in similar submissions. In our assessment, the students were given a set of tools and asked to develop their own theme within a general framework of possible applications of EM. This required the student to be self-motivated and to think for themselves about how they should approach the coursework. As can be seen from Figure 6.1, a wide range of topics and interests was represented in the submissions.

We found that students were able to direct their own learning without being given a specific coursework objective. The student who submitted the Wumpus model initially set out to reconstruct the original Wumpus game using the EM tools provided. This proved successful but furthermore the interactive, open-ended nature of the environment allowed the student to model different scenarios they had not originally considered, e.g. by changing the rules of the game and/or manipulating the information presented to the player. In his submitted paper, he explained how through this interaction with his model he had begun to appreciate how his ability to win the game depended upon the rules of the artificial Wumpus world and how, outside such a constrained environment, pure reasoning was not always sufficient.

6.1.5.2 Learning is stimulated by personal interest

The account of learning in Chapter 1 cites James [Jam25] and Dewey [Dwo59] stressing the importance of personal interest as a motivating factor for active and experiential learning. In this spirit, the students were encouraged to think about issues of which they had particularly rich experience or were particularly interested in learning about. One student explored applications of greedy algorithms by carrying out empirical research into how her younger siblings learnt about giving the correct change. Coursework often forces students to study situations with which they are unfamiliar or topics that do not interest them. By choosing their topic, students were able to draw and build upon a wide range of

prior knowledge, interests and experiences. EM actively encourages this type of learning [Roe03]. Because students worked on topics of their own choice, the focus of their effort could be on EM principles and tools and not on an arbitrary topic prescribed for them.

All of the submissions showed evidence of an interest in domains other than Empirical Modelling. These domains ranged from personal hobby interests to aspects of the computer science curriculum. The Poker and the Frisbee models were inspired by recreational interests. One student made use of his Grid Computing model to complement his coursework for another computer science module. Another developed some research by Microsoft into the IceCube framework. From the depth and quality of his submission, it is apparent that this student spent as much time learning about IceCube as they did about Empirical Modelling. This contributed significantly to the quality and ingenuity of his final model; it also demonstrated how the model-building could stimulate learning in other domains. Yet other students chose to model phenomena in other academic subject areas. One submission relating to human biology was a model of the lungs that incorporated a primitive simulation to expose the effects of damage to organs or of cigarette smoking. The simple but effective use of dependency in this model highlighted the extent to which naive medical knowledge of bodily functions is knowledge of basic inter-relationships between physical conditions and parameters. This underpinned the educational purpose behind the model—just one of many references to education in the written submissions.

6.1.5.3 Learning is reinforced when practice and principles are combined

The practical element of a subject can often become divorced from reflection on principles. Although model-building is a useful tool for developing basic EM skills, it should be guided by higher-level motivations and interpretative activities. In previous years, the written component of the coursework had been primarily oriented towards the technical documentation of models. Introducing the paper-writing exercise into the written component of the coursework helped to promote a broader awareness of the thinking behind EM and its implications.

Several models successfully illustrated deeper concepts of direct relevance to EM. In the Wumpus model, for instance, the environment can be configured so as to expose the limitations of logic outside a context of stable expectations and reliable knowledge. In his extension to the traditional AI game, the student was able to expose problematic aspects of a purely logicist outlook on intelligence in a manner that had not been preconceived.

The Grid Computing model was a good example of a student using a model to convey concepts. In this case the model served to illustrate the basic principles of grid computing by generating animations of the kind of diagram that would typically be found in an introductory textbook.

6.1.5.4 Learning is aided by exploration

The quality of the submissions was such that most of the students were able to grasp the use of the EM tools and, in some cases, exploit their more advanced aspects. This is one reason why exploration into other domains occurred so naturally in the coursework. Once the tools had become familiar, the student no longer had to focus on the technicalities of modelling, but could make use of the tools to communicate or develop their domain understanding. As is to be expected, the better the student's EM skills, the more they were able to explore their problem domain.

To demonstrate this, we have categorized the submissions by the extent to which they explored their problem domain: little/no exploration, controlled exploration, and free exploration. Submissions that showed little/no exploration were generally based on the style of implementation that is quite familiar in computer science. A typical example is a model of a game that concentrates more on the implementation than exploring the observations, rules and interfaces that shape the interactions within the game. In the 'controlled exploration' category, the submissions often related to a problem domain with which the student was familiar, possibly drawn from the academic field. An example is the Human Biology model that enabled the user to explore the effect of smoking on the oxygen intake via the lungs. Models in the 'free exploration' category typically resulted from a student's engagement in unfamiliar problem domains. For example, the Bridges model originated in a basic study of the strength of bridges and ended up modelling complex issues in suspension bridges. Applying a statistical T-test at a 99% confidence level shows that students who engaged in exploration achieved higher marks than the students who did not show signs of exploration in their coursework.

These observations suggest that students who had a good grasp of EM tools were able to engage fully with the problem domain and produce coursework of a higher standard. This is what you would expect as we were evaluating Empirical Modelling ability rather than expertise in the problem domain.

6.1.6 The Second and Third Warwick Electronic Bulletins on EM

The submissions from the First Warwick Electronic Bulletin on EM (WEB-EM) demonstrate that there is some potential for EM to support learning that exhibits the above characteristics. These findings were published in [BHB06]. Since then, the module has continued to run in the same format and, so far, a 2nd WEB-EM and a 3rd WEB-EM have been published online [WEBEM2] [WEBEM3].

One additional aspect to the data is that I requested the students to fill in a short questionnaire at the end of the course (before the final examination). The questionnaire was in two parts. The first part contained 17 statements which the students were asked if they agreed with, rating their level of agreement on a scale of 1 to 5. For example, one statement on the questionnaire was "The labs were easy", and students were asked to tick box 1 if it they felt it was always true (i.e. 'I completely agree'), 2 if it was mostly true, 3 sometimes, 4 rarely, and 5 if it was never true (i.e. 'I completely disagree'). The second part of the questionnaire asked the students to write a short (up to a page) biography of their experience of EM. They were asked to reflect on their first contact with the subject, their personal experiences during the labs and coursework, the difficulties they encountered and the problems they overcame. The students were only given 10-15 minutes to fill in the questionnaire. The questionnaire was not an official part of the course, it was not compulsory, but most of the students were happy to fill in the questionnaire, and from the results it seems that many of them put in considerable effort to relate their experiences of EM in the second part of the questionnaire.

In the first three years of WEB-EM there were 43 papers (along with 43 models) published. (A complete list of WEB-EM submissions are detailed in the appendices.) The extra 2nd and 3rd year data has strengthened the original findings. The 4 characteristics highlighted in WEB-EM-1 are also relevant to the 2nd and 3rd year submissions. Reflecting on the overall experiences of the three years of submissions, as well as the feedback in the last two years, there are a number of other points that are worth highlighting.

6.1.6.1 The value of the coursework part of the module

In the first part of the questionnaire, students were asked to rate how much they agreed with 17 statements. Only 3 of these statements were related to the coursework. In analysing the results, I discovered that the three coursework statements were in the top five 'most agreed' statements (see Figure 6.2). (The other statements that were in the

Statement	Agreement	Mean	S.D.
The lab tutors were helpful	88%	1.47	1.07
The coursework was relevant to the course	87%	1.53	0.80
I enjoyed the lab sessions	82%	1.71	0.99
The coursework helped me understand EM	78%	1.88	0.78
The coursework helped me understand more about my chosen topic through the modelling exercise	76%	1.94	0.83
The labs helped me understand the principles of EM	75%	2.00	1.00
The lab sheets were clear and instructive	72%	2.12	1.22
There are advantages to using EM tools over traditional programming tools	72%	2.12	0.78
I understand the principles of EM	68%	2.29	0.69
The labs were more useful than the lectures	63%	2.47	0.80
The EM tools (tkeden) are easy for beginners	62%	2.53	0.87
The tools are unpredictable	62%	2.53	1.07
The tools are easier to learn than Java	54%	2.82	1.01
EM has helped me understand other aspects of CS	50%	3.00	0.87
The labs were easy	49%	3.06	0.75
EM has no relevance to the rest of the CS course	41%	3.35	1.06
The tools are easier to learn than spreadsheets	37%	3.53	1.37

Figure 6.2: Part 1 results from the 17 questionnaires.

top five in terms of agreement were: "the lab tutors were helpful" and "I enjoyed the lab sessions".) Figure 6.2 shows the results from the first part of the questionnaire. Of the 24 students that took the module in the years 2005 and 2006, I was able to collect 17 completed questionnaires.

The statements in Figure 6.2 are sorted according to the level of agreement with each statement. The 'Agreement' column represents the mean as a percentage, in such a way that if the average answer to a statement was 1 (i.e. every person agreed with the statement completely) then the truth percentage would be 100% (and if mean was 5 implying every person disagreed with the statement completely, then the truth percentage would be 0%). The top statement is therefore the one that was felt to be most agreeable on average.

The three coursework related statements are highly ranked in second, fourth and fifth positions in the table in Figure 6.2. The mean for the statement "the coursework was relevant to the course" was 1.53, representing an 87% level of agreement. This was the considered the second most agreeable statement among the students out of the 17 statements. The two other coursework related questions offer more insight in terms of analysing the effectiveness of EM for teaching and learning. The statement "The coursework helped me understand EM" achieved on average 78% agreement. The high rank of this statement helps to confirm the constructionist idea that practical building activities are a benefit to learning. The statement "The coursework helped me understand more about my chosen topic" provided the most interest as it achieved almost as much agreement as the previous

statement on understanding EM. This is surprising because there was no intention for the students to learn about another domain—they were simply asked to engage in a model-building exercise using EM. However, there is a strong agreement that by engaging in an EM model-building exercise the students have learning more about their chosen topic. Such results echo Jonassen's view that if you want to understand a topic then it is useful to build a model of it [Jon06].

The relatively low standard deviation (S.D.) on the coursework statements implies that there was more consistency between students' opinions of the coursework than other aspects of the module (the students agreed with the statement and also with one another!). This shows that there is some indication that these three statements have an element of truth for the majority of students.

The weakness of this analysis is that it is based on 17 results. Due to this, the first part of the questionnaire cannot offer any conclusive evidence. I was not really expecting to find any significant results in the first part when I designed the questionnaire, although I was surprised after analysing the data that the only strongly held opinions among the students were regarding the coursework. My interest was really in the second part of the questionnaire where I asked to students to write a small biography of their EM experience and to give comments. Given that quantitative data would suffer from the low numbers of students taking 4th year modules, I felt that the qualitative evidence might prove more relevant to this thesis. The idea of using biography as a method for analysing EM activity in the 'Introduction to EM' module was inspired by Knobelsdorf & Schulte, whose talk I attended at Koli 2005 [KS05]. Their research is concerned with understanding how students' learn computer science, not what factors affect the learning: "Research has already revealed many influence factors, like e.g. gender, math grade, role model, prior programming experience, self-confidence, and so on. But addressing one (or some) of these factors might not be sufficient to change students general understanding of computer science or to improve the effectiveness of teaching. Instead of revealing more (singular) influence factors we aim to understand students' preconceptions, their conceptual framework of the subject matter and how it evolved." It is in this spirit that I wish to evaluate learning about and learning through EM.

Biography as a method (as used by Knobelsdorf & Schulte) is discussed in detail by Denzin in his book entitled *Interpretive Biography* [Den89]. Denzin recognises various forms of biography, all of which are concerned with the use and collection of "personal-

life documents, stories, accounts and narratives which describe turning-point moments in individuals' lives" [Den89:p13]. To account for the wide range of possible biographies that Denzin describes, I left it to the students to decide what they would like to write about in the second part of the questionnaire. As a guide I suggested that they write a short autobiography of their experience of EM, possibly reflecting on their first contact with the subject, their personal experiences during the labs and coursework, difficulties that were encountered and problems that were overcome. Many of the students completed their questionnaire including their autobiography during a revision class, and so only spent a small amount of time on the exercise. Also, only a page was given for the students to write their autobiographies including any comments. These factors meant that the biographies were not as comprehensive as those gathered by Knobelsdorf & Schulte [KS05].

Many of the biographies mentioned the coursework in a positive manner. In a significant number of biographies there were indications that the coursework reflected the most important part of the module: "I think the labs were the 'frustrating' period of learning [...] whereas the coursework allowed you to get deep into a model and actually understand the power of EM." This complements the results from part 1. There are a number of comments that other aspects of the course, such as the labs, were not as useful as the coursework. One student explained that the coursework clarified the ideas of the module: "The first lectures were obscure to me and I didn't get the idea. Later, facing the first draft of my coursework, I realised aspects of the lecture and began to understand." This last sentence builds on the sentiments of section §1.2.6 that learning occurs best when principles and practice are combined. In referring to the EFL (Figure 2.7) discussed in Chapter 2, learning inevitably must involve a practical/tacit knowing and a more explicit knowledge [Roe03].

6.1.6.2 EM offers something different to the standard CS curriculum

In the biographies given by students there is a common attitude that EM offers a complementary perspective on computing that is not acknowledged or covered in the standard components of a computer science course. One student commented: "The main thing that EM has encouraged me to think about is the relationship between computing and thinking, which as a more abstract concept had not really been touched upon in the rest of the CS course." Similar feelings were echoed by another student: "The first lectures seemed quite philosophical and were a welcome new take of CS. Nothing else on the [computer

science] course has really considered things the way EM does."

Furthermore, evidence from one student suggests that EM can be beneficial in other areas of computer science: "I found EM to be a completely different module to any I have studied." ... "It took me a while to gain an appreciation for the module, but I have learned to look at all aspects of my course from another perspective." Although the student is not specific about what aspects of the course he was able look at differently, the statement shows that there is some potential for using EM for widening a computer scientists view of their subject in a positive manner.

Another student was more specific in describing the benefits that EM brings to a computer science course: "Its great and refreshing to see things up on a screen and how we can alter it. I enjoyed programming in it, something I thought had lost through finding other languages less 'exciting'." The immediacy of experimentation is highlighted here as a refreshing aspect of EM. The comparison with other programming languages is hinting that EM offers something special that other languages are unable to offer.

6.1.6.3 Difficulties in understanding the theory and using the tools

Despite the positive comments in biographies substantiating the potential for EM as a unique approach to computing, there were also a number of comments describing the difficulties of EM. The first issue is that students develop misconceptions initially and it is not until later in the module, usually during the coursework, that these misconceptions begin to be rectified to some degree. The collection of biographies contained many comments similar to this: "My first contact with EM in the first few lectures led me to mistakenly believe that it is another programming course like the java or functional programming module I took in year 1." There was one student who still held some misconceptions at the end of the module: "Still don't get how EM differs from simulation other than being more flexible."

Another issue was that students were seriously concerned by the 'what is EM?' question. Computer science students it seems are keen to have definitions, and when a simple definition is not easy to find or representative of the activity then it presents a difficulty for many students. One student put it very simply in their biography: "Understanding exactly what EM is was quite difficult." Another student observed a general feeling that there was not a simple link between the philosophy and the practice: "I got the feeling from some people that they couldn't understand the link between their application & the

EM principles." King's work is a recent attempt to define EM as a philosophy clearly connected to practice [Kin07], resulting from his own reflections upon the question 'what is EM?'.

The final issue relates to the quality of the EM tools and has been discussed in detail elsewhere [Kin07]. The `tkeden` environment is an experimental piece of software developed by students, over 18 years, and suffers from many idiosyncrasies. A number of student biographies highlighted problems with the tool, for example: "The tools (tkeden) are functional but still could benefit from small usability improvements—some times you're wrestling against the tools instead of doing the actual modelling." It is clear from this that the quality of the tool is lagging somewhat behind the quality of the principles. A number of improvements have been made to the tools for educational applications, as used in this thesis (i.e. AOP, GEL, & EMPE, as discussed in Chapter 4), but further work is needed.

6.1.7 The significance of EM for learning

The characteristics of the learning exhibited in the WEB-EM submissions accord well with the characteristics of EM. Where conventional programmers are encouraged to assemble a secure base of knowledge prior to writing the first line of code, EM practitioners are encouraged to initiate their exploration of the application domain and their construction of a computer-based model simultaneously. The fundamental reason for this distinction in outlook has to do with the perception of knowledge that underlies thinking about conventional programming and EM. The conventional programmer targets sophisticated knowledge that is sufficient to provide a robust logical framework ('knowing that certain relationships hold') and complementary precise recipes for action ('knowing how to achieve specified goals'). By contrast, EM is primarily concerned with a much more primitive conception of knowing (cf. [Bey05a])—with conjunctions between experiences as personally encountered by the modeller. The qualities of EM artefacts stem from this grounding in experienceable connections that pervades the context within which all 'knowing that' and 'knowing how' is subsequently rooted.

The principal technical contribution of EM to moulding one experience so that it 'knows' another is to be found in the notions of observable, dependency and agency. The diversity of the domains represented in WEB-EM-1 as depicted in Figure 6.1 is further evidence of the pervasive relevance of these primitive notions, which are viewed as

conceptually prior to the identification of formal objects and structures. The integrity, functionality and interpretation of an EM artefact is framed by the meaningful interactions that the modeller can carry out with it, as guided and constrained by past experience of interaction. In this respect, EM artefacts are ontologically quite unlike computational objects and structures, even though in practical settings they may resemble them closely. The character of EM artefacts is consistent with the features of the learning activity: not necessarily being associated with a specific learning objective; connecting closely with personal experience; synthesising empirical and theoretical elements; drawing upon extensive exploratory activity. These characteristics are also relevant to whether students chose to pursue the 'Introduction to EM' module to its completion, since it epitomises an orientation towards knowledge that is congenial to some but alien to others. Unsurprisingly, there is only a loose correlation between good overall performance in computer science and aptitude for EM.

As the module was an introduction to Empirical Modelling, it was intended and expected that the students would learn about the basic concepts of EM. Students demonstrated their proficiency with EM tools and techniques through the models they were able to construct. What is more surprising is the extent to which some students learnt about the topic area chosen for their modelling exercise in carrying out the assessment. Learning evidently occurred in both domains—in EM and in their area of interest. Given the strong agreement among the students, it is likely that this will be the case in other domains beyond those covered by student submissions to date.

Furthermore, the assessment exercise promotes those meta-skills relating to self-motivated, self-directed exploration of a problem domain that are most needed by life-long learners (see §5.3). It is characteristic of life-long learners that they are not necessarily following a formal path of education. They are much more likely to have personal goals and individual learning objectives. The above review indicates that EM tools have the potential to address the needs of life-long learners, as discussed in more detail in Chapter 5.

EM and the assessment strategy seem to work well together, and to offer prospects for learning in other disciplines. Whether the style of assessment would be so effective with traditional programming tools is unclear, since it is hard to interpret embodying observation, dependency and agency in models in that context.

When all students built models from the same specification, assessment was more straightforward because it was easier to identify the relative merits of each piece of work.

The new style of assessment has meant that there is a greater focus on the use of EM principles and how they have been applied to the specific topic area chosen by the student. This presents challenges for the marking procedures as it requires additional time to familiarize with each individual model and its topic. However, we have found that the students who focused on unusual subject areas for their model building tended to build models that in themselves promoted our own understanding of that area.

6.2 Third year projects

There is further empirical evidence of EM's potential for supporting experiential learning in a number of undergraduate third year projects. Every computer science undergraduate must undertake a project in their third year which counts for a quarter of their marks for the year (the project is worth 30 CATS and the normal load for a year is 120 CATS). Students may undertake a project of their choice (within the bounds of the course and assuming they find a supervisor who deems the project suitable). For a number of years there have been students undertaking projects relating to Empirical Modelling, supervised by Meurig Beynon and Steve Russ. Each year there have been up to 15 students, and so over a number of years the number of projects has been substantial. Some of these projects are available and freely downloadable from the EM Archive [EMP], which as of May 2007 contains over 161 models. These projects have contributed significantly to the EM tools and to students' modelling activities, as well as the thinking and philosophy behind EM. Many (if not all) of these projects have a learning element that is significant to this thesis. A complete account of all the third year projects is beyond the scope of this thesis. However, these projects perhaps stand as the best source of empirical evidence to date of deep significant learning through Empirical Modelling. This section only touches upon a few selected aspects of two selected projects in which I feel there is a strong sense in which EM has led to significant learning. The aim is to show on a micro-level the contribution that EM can make to the learning of an individual exploring a particular domain.

6.2.1 Learning about planimeters

The first third year project I would like to highlight is that of Charles Care who developed several models of planimeters for his final year undergraduate project [Car04b]. A planimeter is a mechanical device, invented in 1814 for land surveying, that can be used

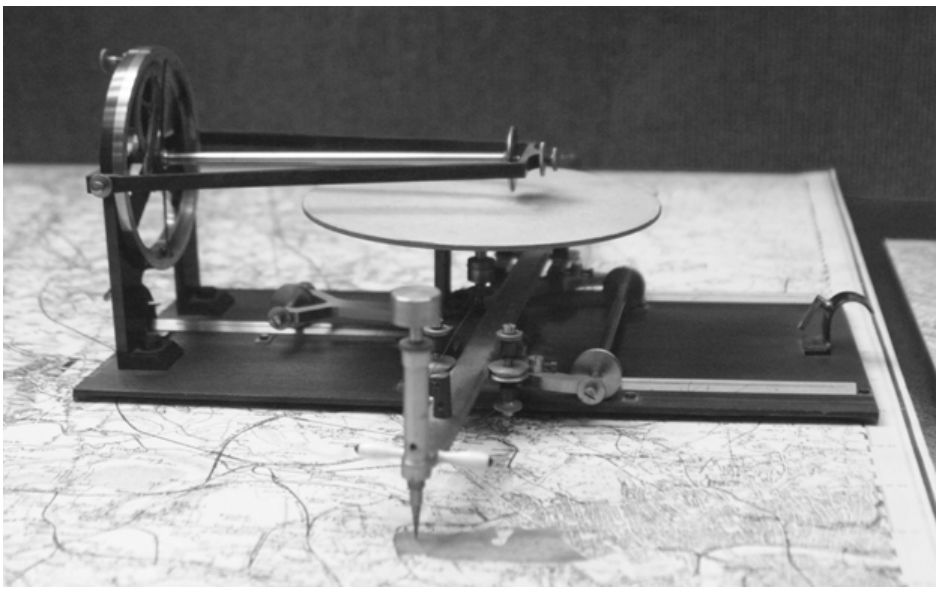


Figure 6.3: A Wetli planimeter as displayed in the London Science Museum (replicated from [War04:p13], photo by Mark Lloyd).

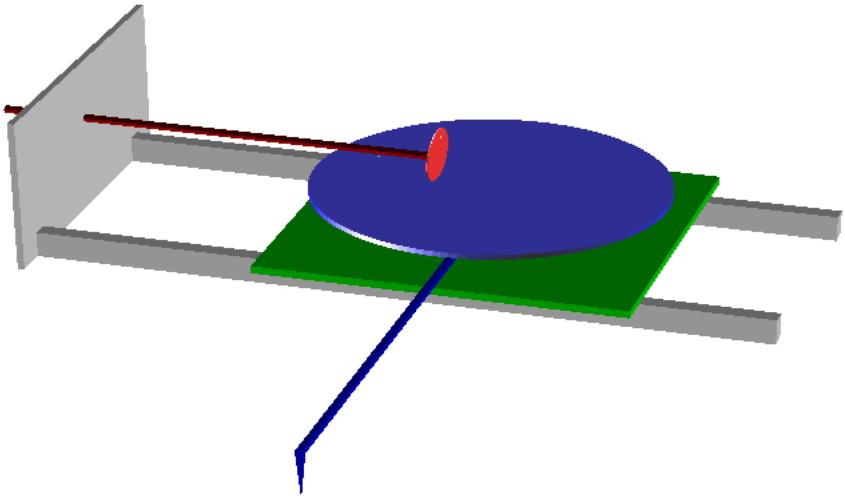


Figure 6.4: Care's model of a Wetli planimeter [EMP:planimeterCare2005].

to measure area [Asp90]. It is considered an important device in computing as its integration capability signifies the beginning of analogue computing [Car04b]. An early example of a planimeter is the Wetli planimeter that uses a disc and wheel mechanism as shown in Figure 6.3. Care developed a number of models using EM techniques including a 2D model, a distributed model, and 3D models of different planimeters—Figure 6.4 shows a 3D model of the Wetli planimeter (developed by Care [Car04b]).

The potential of the planimeter models for learning about planimeters is what strikes me as most significant about Care's project. During a recent talk about planimeters given by Care, he mentioned the effect of his explorations building models of planimeters as leading to a better understanding of their workings than historians might usually obtain. Care

described how on a trip to the Science Museum he met a curator with whom he discussed planimeters. To his surprise, he appeared to have a deeper understanding of the workings of the planimeters on display simply from this experience of constructing and playing with models of planimeters. This suggests that there is a special power in constructing models that brings about a deeper understanding than can be grasped from language alone. The curator, in this case, had available to her at least as much information—in the form of books, historical accounts and records relating to the planimeter—as Care had available to him. Care was surprised because he expected a curator interested in planimeters to have developed a deeper knowledge of the subject. However, Care used the information about planimeters to construct an EM model and it is this process that might have led to a deeper understanding of the workings of planimeters and the important issues around planimeter design. In a paper written after his project, Care shares a similar conviction:

> "Descriptions of various planimeters, their construction and how they work are documented in a number of sources. However, these usually take the form of annotated diagrams with written explanations that are difficult to relate to. Such descriptions fail to give the reader anything like the experience of holding a real planimeter in their hands. My suggestion is that the models described do provide this kind of experience and that this property is not simply associated with the model but also with the modelling process employed." [Car04a]

The deeper understanding that comes from the modelling process is perhaps not as expressible as the objective information about a planimeter that is generally found in books. This deeper understanding might be referred to as 'tacit knowledge' as described by Polyani in his work on the nature of knowing [Pol62]. If it is tacit knowledge then I suspect it is more like the understanding that comes about from the experience of designing and the repeated experiences of using a planimeter. Care did consult technical documents on the design and construction of planimeters, which would fit Lave & Wenger's idea that learning involves a deepening process of participation in a community of practice [LW91]—Care participating in a community of planimeter engineers that no longer existed.

In discussions with Care, I have come to believe that his investigations into planimeters show a deep understanding that is quite different from the explanations given in books, such as the mathematical description in Aspray's book *Computing before computing* [Asp90]. There is something special about Care's knowledge of the planimeter. I can, at best, try to give a glimpse of a small element of the learning that has taken place during his investigations. The following explains one small 'eureka moment' that occurred within a few days early during a 6 month period of modelling. My exploration of the modelling

activity is reconstructed from a history of scripts from 24th to 27th November 2003 that were carefully recorded by Care. Each script represents a point (or state) of significant change in his modelling activity, although I have identified from logs that in between these saved scripts there were a significant number of interactions. A complete analysis of these scripts is included in a separate document [Har07b].

6.2.1.1 The incident

Care's 'learning about planimeters' began after he became interested in planimeters during a module on the history of computing. His third year project was originally aimed at demonstrating that computer models of historical artefacts are beneficial for learning, although after a while the project became more of an investigation into the history of planimeters and their workings. In the initial stages of his project, Care was concerned with constructing a computer model (using EM) of a planimeter. He started considering a small part of the planimeter device, the wheel and the disc, which are essential to the integration function. In the Wetli planimeter (Figure 6.3), the small wheel sits on top of and at right angles to the disc. When the disc spins, friction causes the wheel to turn. The amount the wheel turns relative to the disc depends upon how close the wheel is to the centre of the disc. If the wheel is at the edge of the disc then it will turn more (when the disc spins) compared to if the wheel is nearer the centre. A typical diagram used to teach this principle (and the actual example in the History of Computing module) is shown in Figure 6.5. It is from this mathematical description that Care began his disc and wheel model.

At 12.59pm on the 24th November 2003, Care sat down at a workstation in the Computer Science building, loaded tkeden and started experimenting with the model he had been working on the previous week. At this early stage in his project, his model consisted of simply a 3D disc and wheel. A screen-shot of his model and a selection of the relevant underlying observables and dependencies are shown in Figure 6.6(a). These observables have a direct correspondence to the variables explained in Figure 6.5. The observable 'wheelPos' is equivalent to y, 'discSpeed' is Δx (speed corresponds to the displacement in one clock tick), 'wheelSpeed' is Δz, and 'wheelRotation' is z. It is from here that Care began to explore the model further and eventually realise a subtle problem with the dependencies in his model's current state as of that day.

In the next hour he began to construct a mouse sensitive window with which he used

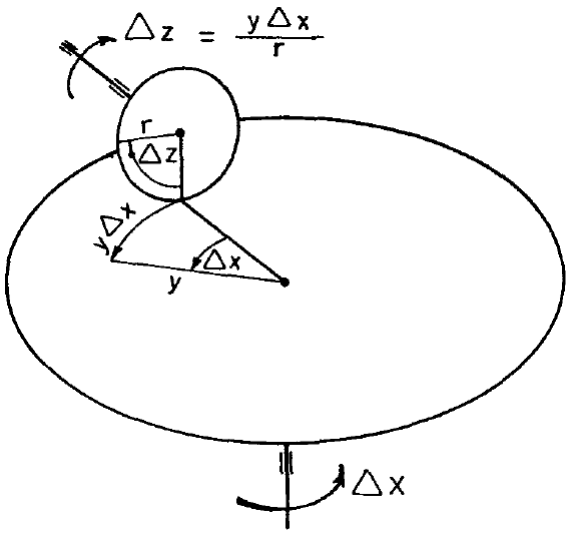


Figure 6.5: "The principle of operation of the Wetli disk-and-wheel integrating mechanism. The dependent variable, y, determines the distance of the integrating wheel from the center of the disk. If the disk is rotated through an angle, Δx, then the integrating wheel is turned by $\Delta z = [y\Delta x]/r$. Hence, after a period of action, $z = 1/r \int y\, dx$." (text and diagram taken from [Asp90:p167]).

to make the position of the wheel dependent on the position of the mouse. He took a break at 1.58pm, but returned at 4.05pm and continued where he left off. Care started adding marks to the disc and wheel (a blue square on the disc, and a red square on the wheel) which helped to indicate the relative speed and amount of rotation of both the disc and the wheel. Care stopped again at 4.46pm, and did not resume modelling until 9.22am the following day (25/11/2003). In the next script that was recorded, we can see that Care had been altering the disc speed and in particular slowing it down. I can only assume that this was to better observe its behaviour, and perhaps even at this stage there was a hint of suspicion in Care's mind that his model was not operating in the same way as the physical device. By 10am, Care had added another mouse sensitive scout box to control the speed of the wheel. The screen-shot in Figure 6.6(b) shows the state of the model at this time. This addition enabled Care to experiment with the position of the wheel and the speed of the disc at the same time. It was the combination of marks on the wheel for observing the speed, the sensitive box for controlling the speed of the disc, and the sensitive box for positioning the wheel that led to an important discovery. Care had worked until 10.04am when he left his workstation, and then he came back at 1.03pm. At 2.04pm he saved his current state once again and it shows that he had made a significant change to the way rotation was being calculated. During that time, Care had removed the discRotation and wheelRotation dependencies and replaced them with a triggered procedure (as defined in Figure 6.6(c)). The reason for this change is because Care noticed that when the disc was

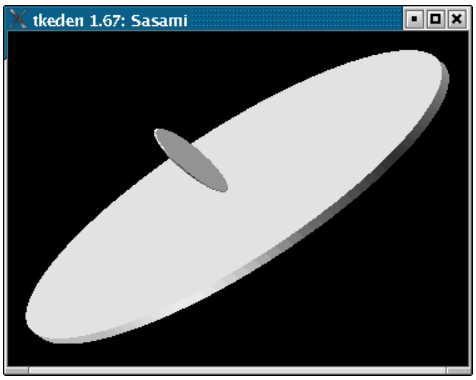
moving very slowly, dragging the wheel across the surface of the disc caused an unexpected rotation of the wheel. If the disc was stationary, then dragging the wheel across it should result in no rotation of the wheel, but this was not the case in the model. The change that Care made to his model was due to the realisation that the rotation of the wheel must be calculated cumulatively, by adding the current wheel speed (Δz) to the wheel rotation at each increment of unit time (hence the use of a triggered procedure running on every clock tick). This change produced a noticeably different behaviour (two videos demonstrating the before and after behaviour can be found in [Har07b]).

Care remembers this moment well and describes it as when he realised the flaw in his conception of the way rotation is calculated on the wheel. However, I do not think that it was particularly a flaw in his model because at that time he had been following mathematical explanations based on those described in Figure 6.5. His model did accurately model the mathematical description from the book. The problem was that the book did not fully represent the subtleties of the physical mechanism. That is, there was no mention in the mathematical description that there was a temporal element to disc speed, that the disc speed would vary over time, and that this would clearly effect the cumulative wheel rotation. Perhaps it was not possible to represent accurately in mathematics the precise nature of interaction between the disc and the wheel—there was a necessary 'agency' element. It was only through building a model, and reflecting on the dependency and agency together, that a deeper understanding of the mechanical device was able to arise.

The modelling activity continued after the described incident, and Care went on to progressively build on the interface during the 26th & 27th November. Figure 6.6(d) represents the model on the 27th by which time Care had improved the mouse sensitive window and added two dimensional views of the disc and the wheel to closely observe their state. This model is used to describe the integration mechanism in his project report [Car04b]. The modelling process did not end here though, the disc and wheel model was used to construct the larger model of the Wetli planimeter as shown in Figure 6.4.

6.2.1.2 The significance of the incident

The incident described above constitutes a very small piece of learning when considering that Care's project was spread across a period of at least 6 months in which he explored many different types of planimeters. It was a particularly important piece of learning for Care though as all of his planimeter models relied on the basic principle of the disc and

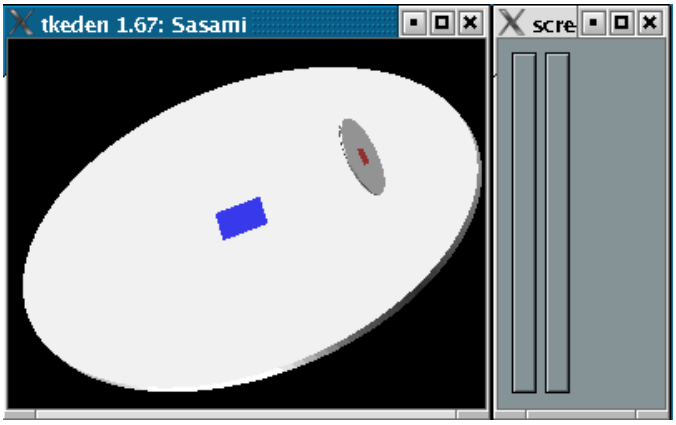



```
 ## The disc is rotating a constant speed
discSpeed = 1.0;

## The speed of the wheel is determined
## by the position of the wheel
wheelPos is discRadius*wheelRatio;
wheelSpeed is wheelPos*discSpeed;

## The amount of cumulative rotation on
## the disc and wheel can be calculated
## over time
discRotation is timeCount*discSpeed;
wheelRotation is timeCount*wheelSpeed;
```

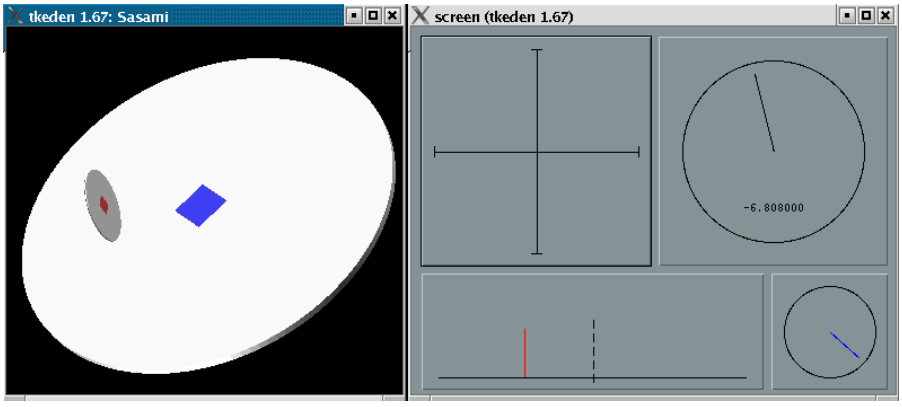
(a): The disc and wheel model on 24/11/2003 with relevant dependencies.



(b): The disc and wheel model on 25/11/2003 with two mouse sensitive bars to control the position of the wheel (wheelPos) and the speed of the disc (discSpeed).

```
proc clock: timeCount {
  discRotation = discRotation+discSpeed;
  wheelRotation = wheelRotation+wheelSpeed;
   todo("timeCount=timeCount+1;");
};
```

(c): A part of the script on 25/11/2003 after replacing the rotation dependencies in (a) with a triggered procedure for calculating the cumulative rotation.



(d): The disc and wheel model on 27/11/2003 with an interface for observing each part of the disc and wheel behaviour.

Figure 6.6: Stages in the disc & wheel model from 24/11/2003–27/11/2003.

wheel for integration.

It is difficult to say precisely what caused the learning in this small incident. The evidence shows that Care's knowledge of the planimeter goes deeper than the information given in any book or other sources that he had encountered. His overall knowledge of planimeters was clearly heavily affected by the information given in explicit sources, but in this particular incident, it is difficult to link it directly to a source of information. Neither would it be correct to link the learning solely to the computer. The computer, or the software, did not make the discovery (the computer was completely unaware that what was being constructed had a relationship to something else in the world). The learning, or the knowledge, arose in the person, but it was in response to the configuration of the observables and dependencies in the model. It is appropriate at this point to raise the question of whether there was anything particularly special in the environment or whether the EM tool provided anything exceptional over any other environment. It is true, that the learning is not singularly the result of EM. However, Care's incident demonstrates that observables and dependencies, including the open environment in which different configurations of *observables* and *dependencies* can be explored, can support the development of understanding that goes beyond explicit knowledge as digested information. The open-ended nature of the environment allows *agency*—making and attributing state changes—to play an important role in developing understanding. There are limitations in the mathematical descriptions of planimeters from textbooks, such as [Asp90]. Polyani views this as explicit knowledge and acknowledges that 'tacit knowing' is a deeper understanding that is difficult to describe in language [Pol62]. One solution to developing a deeper understanding is to construct an external model, as in constructionism. Papert's constructionism [PH91] is not concerned how the model or program is developed. EM specifically offers principles for constructing models (with observables, dependency and agency) that accord with experiential learning and tacit knowing in a way that differs from traditional model-building and programming techniques.

Care's interactions with the disc and wheel model show on a micro-level the way in which combined model construction and use can promote the development of a knowledge of the workings of planimeters. Furthermore, it has been observed that the knowledge of planimeters gained by Care is quite different from the book knowledge that is taught on history of computing courses. I have claimed that Care's experience is probably better described as similar to an engineer engaged in practical use and experimental construction

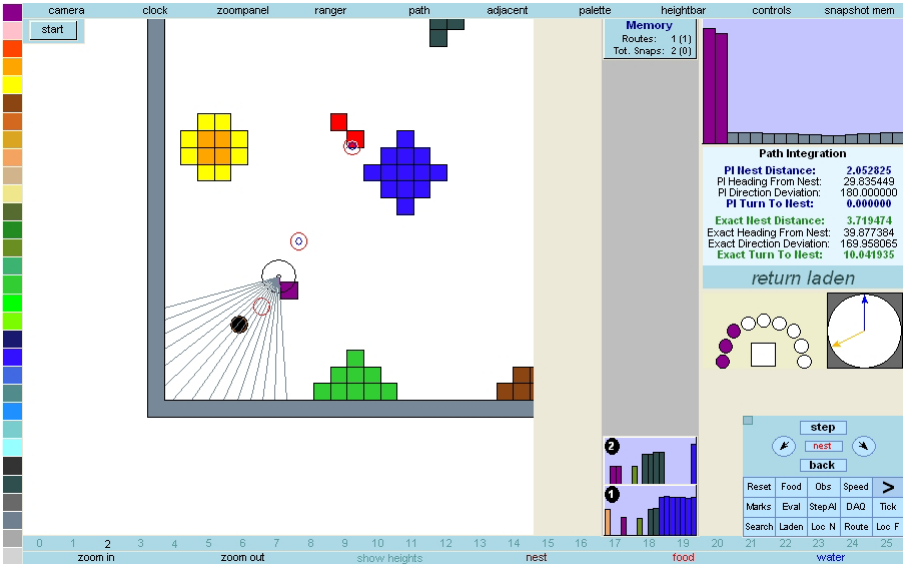


Figure 6.7: The ant navigation model by Keer.

of planimeters. This is supported by the idea that significant learning takes place when engaged in activities linked to a concrete situation.

6.2.2 Learning about ant navigation

The second project to be considered is about ant navigation and was undertaken by Daniel Keer as a final year project [Kee05]. The model, as shown in Figure 6.7, can be found in the EM archive [EMP:antnavigationKeer2005]. The aim of this project was to understand the method which a particular species of desert ants (genus cataglyphis) use to find food. Previous entymology research by Collett suggests that these ants, compared to other species, have relatively accurate colour vision and are able to recall and match images that they have seen before [CDGW92]. Ants, as well as other insects and animals, have a 'path integration' sense (a construal of the position of their nest) that enables them to relocate their nest after a long meandering forage for food. Collett has found that desert ants are able to maintain an accurate construal of their position on food runs of up to 200 metres [CDGW92]. He also suggests that when an ant has located a food source, it is able to remember a list of snapshot views on returning to its nest that allows it to find the food a second time [CDGW92].

6.2.2.1 Modelling process

Whereas the planimeter model was constructed from scratch, Keer's project was based around an existing ant model. Previously, K.C. Tan developed a simple model of an ant moving around a square space containing blocks [Tan99] (first in the Maintainer of Dy-

namic Dependencies (MoDD) tool, but later in `tkeden`). Tan's model [EMP:antsTan2004] formed the basis of Keer's ant navigation model, allowing Keer to begin experimenting with ant navigation at an early stage. This demonstrates the benefit of model reuse, that experimentation in the domain can begin at early stage if an existing model fits the needs of the model-builder.

Although Keer was able to utilise Tan's ant model, it was heavily extended during the modelling process to create a suitable environment for exploring ant navigation. Keer considers the initial part of the project to prepare the environment as the most straightforward: "For the environment/interface, I had a clear plan of what was necessary, and developed [Tan's] model accordingly." Contrast this with the planimeter model where the modelling process was relatively open at the beginning in terms of how to develop the model. The possible reason for Keer working with an initial model that was already fairly well specified was likely to be because his main concern was for the experimental work. Keer's interest was in how to develop AI that mimics the behaviour of desert ants. This was further complicated by a lack of deep understanding in the literature as to the behaviour of desert ants. Thus, Keer's project was an investigation into desert ants and AI that imitates desert ants. It is clear that Keer was a lot less sure about what he was looking to do to develop the AI: "When it came to the AI, I had some ideas about how I might implement the snapshot matching, but was unsure how successful these ideas would actually be in practice." [Kee05].

6.2.2.2 An ant environment for experimental learning

To understand EM's potential for supporting the type of experimentation that Keer undertook, it is worth considering deeper issues regarding this type of learning. Beynon & Russ describe the usual type of experimentation supported by computers as being associated with "a stable objective context of observation in which parameters can be changed and the outcomes observed" [BR07]. Whereas, traditionally, usual scientific experimentation is more typically concerned with "identifying appropriate contexts for reliable observation, distinguishing between essential and accidental features of interaction, deciding what is deemed to be an outcome and what is deemed to have significant implications for this outcome" [RB07]. It is this type of exploratory sense-making activity that they call 'pre-theory' experimentation [RB07]. Such thinking is aligned with Gooding, a philosopher of science, who explains, in his book on 'Experiment and the making of meaning', that Fara-

day's knowledge of electromagnetism evolved through pre-theory experimentation [Goo90]. Gooding interprets that Faraday did not perform experiments to explain some aspect of reality but to find out what the reality was. Faraday developed experiments with his own artefacts and using his own procedures, and it is these 'construals' that contributed most to the science of electromagnetism [Goo90].

While other examples in this thesis may evoke an image of post-theory experimentation (i.e. interactions with jugs), the ant navigation project undertaken by Keer has a distinct element of pre-theory experimentation. Keer had a rough idea for the kind of interface that might be needed for experimentation, in a similar way that Faraday may have had a rough idea of some of the artefacts he needed for understanding phenomena relating to electromagnetism. However, when it came to the procedures that Keer used for experimenting, these were very much influenced by the situation and could not have been preconceived beforehand. For example, Keer experimented with different ideas in the beginning to see what insights could be gained. In particular, he found that new ideas could be generated "by moving the ant myself and seeing what she could perceive of the environment" [Kee05]. Keer found that he "built up the complete working AI through this interaction and experimentation with the model" [Kee05].

As highlighted by Beynon & Russ, human engagement plays a central role in pre-theory experiment [RB07]. Keer confirms this sentiment: "Experimenting with the model from the 'point of view' of the ant was vital for generating ideas about how the ant could compare snapshots with current surroundings." Interaction such as this appears to be crucial to developing understanding and learning as Keer goes on to say, "I think it would be very difficult to develop such an AI without 'playing' with the model to generate such insights" [Kee05].

Gooding's account of scientific discovery emphasises the role played by developing artefacts and procedures for experimenting during the experiment itself. Although Keer developed some of the ant environment before experimentation began, there were some parts of the environment which developed during experimentation. Simple additions such as buttons for acting out a specific element of agency, like moving the ant to the nest, were added on-the-fly. The potential to change the construction environment whilst constructing is evident in the EM Presentation Environment discussed in §4.2.2.2 and §5.2, when changes were made to the underlying notation in response to newly identified needs in the current model (i.e. "in the stream of thought").

Beynon & Russ emphasise that while the traditional specification-led approach to programming is not well-suited to pre-theory experimentation, the EM approach has the right ingredients due to the close relationship with observation and experiment [RB07]. Furthermore, EM artefacts such as the ant navigation model have much in common with 'construals' in Gooding's sense [RB07]. Keer endorses these affinities in his conclusion that "the Empirical Modelling 'experimental' approach has significant merit for AI development." [Kee05].

6.2.2.3 Differences between ant navigation and planimeter modelling

In both the planimeter project and the ant navigation project, it should be clear from the above discussion that the students were learning about each of their respective domains (as well as improving their model-building skills). The nature of this learning is linked in that they were exploring areas unknown to themselves, but in terms of the areas themselves the nature of the learning is different. In the planimeter project it was the case that Care's referent in the world was a physical planimeter, something that (had he had one) he could have compared with his model in order to check it reliably followed its behaviour. A planimeter, having been designed and engineered, has a clear behaviour. When the problem with the corkscrewing wheel arose in the EM model, it was a discovered because the behaviour did not accord with the physical planimeter. In the ant navigation project Keer had a quite different task in that he was attempting to model a phenomena whose behaviour is not well understood. His exploration of building a model was experimenting to confirm possible theories, or even describe new theories, about the plausibility of desert ants navigating in particular ways.

6.2.2.4 Implications for learning

Not only is Keer using EM to construct a model of ant navigation, but he is also using EM to construct an environment for exploring and learning about the nature of ant navigation. A traditional educational technology approach to this would force the construction of the environment first, and later a learner could use the environment to explore ant navigation. This can be called post-theory experimentation. EM allows learners to be much less constrained by a specification (if indeed there is any specification at all), and enables learners to follow their own lines of inquiry, by considering all artefacts and procedures—observables and dependencies—in the environment to be open to change by

the learner during any experiment. Thus, EM can be considered as supporting pre-theory experimentation.

6.3 The basis for further empirical work

The coursework for the 'Introduction to EM' module shows the general applicability of EM for learning in a wide range of subjects and domains. It offers evidence on a macro-level that EM has potential in domains that have yet to be explored in any detail. The third year projects in planimeters and ant navigation examined particular examples of learning through EM in a specific domain. These studies show on a micro-level that EM has potential for facilitating learning that is different to, and complements, the learning from language-based sources (whether they be written as in books, or spoken as in teacher/classroom sources).