Teaching Discrete Structures: A systematic review of the literature

Thomas Whelan
Dept of Computer Science
National University of Ireland
Maynooth, Co. Kildare, Ireland
thomas.j.whelan@nuim.ie

Susan Bergin
Dept of Computer Science
National University of Ireland
Maynooth, Co. Kildare, Ireland
susan.bergin@nuim.ie

James F. Power
Dept of Computer Science
National University of Ireland
Maynooth, Co. Kildare, Ireland
jpower@cs.nuim.ie

ABSTRACT
This survey paper reviews a large sample of publications on the teaching of discrete structures and discrete mathematics in computer science curricula. The approach is systematic, in that a structured search of electronic resources has been conducted, and the results are presented and quantitatively analysed. A number of broad themes in discrete structures education are identified relating to course content, teaching strategies and the means of evaluating the success of a course.

Categories and Subject Descriptors
K.3.2 [Computers and Education]: Computer and Information Science Education—Computer science education, Curriculum; G.2.0 [Mathematics of Computing]: Discrete Mathematics—General

Keywords
Computing curriculum, discrete structures, discrete mathematics

1. INTRODUCTION AND MOTIVATION
The 2008 ACM Computer Science curriculum defines discrete structures as foundational material for computer science, that is required by many other areas in the syllabus [1]. This curriculum defines 43 core hours of topics, covering sets, logic, counting, graphs and probability. As a product of a long process of standardisation, adopted by the ACM and the IEEE Computer Society, it constitutes the standard definition of the topic.

However, in attempting to implement such a curriculum, many decisions remain for the educator. For example, frequently a choice must be made of which sub-topics to cover, bearing in mind the particular needs of a given programme, and the possibility that some topics may be covered elsewhere in the curriculum. Beyond this, the correct strategies for teaching and assessment must also be worked out. Finally, the ACM curriculum documents, while firmly based on years of experience, do not themselves provide a guide to the literature available in the field.

In this paper we address the last of these points, providing a literature survey, in order to provide a basis for addressing the others. In particular, we present the results of a systematic literature review of the teaching of discrete structures, based on the guidelines outlined by Kitchenham and Charters for performing literature reviews in software engineering [20]. Our evaluation contains cues from other sources of analysis, in particular the meta-analysis presented by Valentine [44].

The starting point of our review is to identify the main research questions:
Question 1: What approaches have been taken to designing the curriculum for discrete structures courses?
Question 2: What teaching methods have been employed in these courses?
Question 3: How has the success in teaching discrete structures courses been evaluated?

Rather than presenting the entire systematic literature review process, this paper provides an overview of the process and a summary of the results; fuller details can be found in the associated technical report [48].

We believe this review will be of interest on at least three levels: first, as a review of the literature spanning many decades of teaching discrete structures, second at the meta-level, providing a meta-analysis of publications in the area, and third at the meta-meta level, as an example of applying the systematic review process applied to CS education.

Phases of paper selection and analysis
The paper selection strategy proceeded in three phases:
Phase 1: Use an automated search engine to identify papers that deal with topics relevant to our research. For simplicity we have restricted our search to electronically-available resources, but, given the topic, we do not believe this to be a major limitation of our review. This phase is described in Section 2.
Phase 2: Manually filter the results based on reading the abstracts, in order to identify the subset of these papers that are relevant to our research questions. This phase is described in Section 3.
Phase 3: Extract and synthesise the research by reading the remaining papers. Section 4 presents the results of this phase, discussing some common themes in the literature concerning course structuring, delivery and evaluation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGCSE '11, March 9–12, 2011, Dallas, Texas, USA.
Copyright 2011 ACM 978-1-4503-0500-6/11/03 ...$10.00.
2. AUTOMATED SEARCH RESULTS

The **search space** for sources consisted of, in order, the ACM Digital Library (ACM DL), ACM Guide to Computing Literature (ACM Guide) and the Google search engine. All searches were carried out between July and September 2010.

It quickly became apparent that both “Discrete structures” and “Discrete mathematics” were good search terms for the kind of papers we wanted. Beginning with these search terms (referred to as “base terms” from now on) we also tried searches including the phrases “teaching”, “education” or “learning” to enhance the search precision.

2.1 Search Results

When displaying results in the ACM DL there are two options. The first, “Sort by” was left at its default setting (“relevance”), while the second, “Form” was set to “condensed form”, which contains 50 paper titles per page.

Since the search for each of the base terms returned about 20,000 papers, we needed to establish a cut-off point for manual filtering. In the remainder of this paper, we use a cut-off density of 4%: that is, we did not consider more pages of results after two or less papers in a page of fifty were judged relevant. In total, 65 papers were manually identified by title alone from the ACM DL.

The ACM Guide contains over 1.2 million citations from over 3,000 publishers, enabling a larger search space than the ACM DL alone. However, many of the results were not relevant to the topic or the research questions, or did not have a readily available on-line copy. Only two new papers were found in the ACM Guide (references [24, 50]).

The final search strategy used the search engine Google, keeping in mind that many results returned from such a search engine will not be relevant and/or academically published. In order to counteract this, the base terms were only ever queried along with either “teaching”, “education” or “learning”. All queries were done as individual words and also as complete phrases, and the results were examined manually up to the 5th page. Interestingly, having removed unrefereed or non-relevant sources such as lectures and coursework, no new papers were identified by this search.

Following these searches a total of 67 unique papers were identified by title alone as meriting further analysis.

2.2 Analysis of search terms

Table 1 shows the results of searching the ACM DL using the base terms and their extensions. As can be seen, the vast majority of relevant papers were retrieved using the two base terms, with the extra terms yielding just six new papers in total, only two of which were unique.

Figure 1 contains a bar chart showing the distribution of relevant papers retrieved, broken down by search term over five year periods. In this bar chart, when a paper was found using more than one search term its score was divided between these terms. From this figure it can be seen that whereas “discrete structures” was the dominant term until 1985, since then “discrete mathematics” has become more common, with an increasing number of these being additionally qualified by one of the secondary search terms.

<table>
<thead>
<tr>
<th>Search Term</th>
<th>Number of Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inspected</td>
</tr>
<tr>
<td>Discrete Structures</td>
<td>100</td>
</tr>
<tr>
<td>+ Teaching</td>
<td>150</td>
</tr>
<tr>
<td>Discrete Mathematics</td>
<td>200</td>
</tr>
<tr>
<td>+ Education</td>
<td>250</td>
</tr>
<tr>
<td>+ Learning</td>
<td>200</td>
</tr>
<tr>
<td>+ Teaching</td>
<td>150</td>
</tr>
</tbody>
</table>

Table 1: Numbers of ACM DL papers judged relevant based on examining the results using various search terms.

![Figure 1: Distribution of relevant papers from the ACM DL by search term. Each bar represents the number of relevant papers in a given five-year period; subdivided by the search term used.](image)

3. FURTHER MANUAL FILTERING

Following their selection by title, the 67 papers were then manually filtered by reading their abstracts and determining whether they contained information relevant to the research questions. The main reason for rejection of a paper in this phase was the nature of the paper (e.g. book review, poster and discussion sessions, documentation) rather than the content. This phase resulted in the exclusion of a further 20 papers, leaving 47 papers for full consideration.

**Precision and Recall**

In order to measure the exactness and completeness of our search for sources we can calculate the precision and recall associated with each search term. Precision is defined as the number of relevant documents retrieved by a search divided by the total number of documents retrieved by that search, while recall is defined as the number of relevant documents retrieved by a search divided by the total number of existing relevant documents (which should have been retrieved).

Calculating precision requires determining the total number of documents retrieved by a search, but there are a number of possible choices here. For example, the search of the ACM DL for the term “discrete mathematics” yielded 19,013 papers, but only 200 were actually examined by title, using the 4% density cut-off mentioned above. Thus, dividing by
19,013 gives a slightly lower result than is fully accurate, but dividing by 200 inhibits comparison with other terms whose density may have reduced at a different rate. Bearing this caveat in mind, Table 2 uses the full total figure to facilitate comparison between search terms.

Table 2 gives the precision and recall values as percentages, assuming that the relevant papers are the 47 remaining after filtering based on abstract. As can be seen from this tables, the terms “discrete mathematics” teaching” and “discrete mathematics education” provide the most precise search terms. Consistent with the results shown above in Figure 1, the term “discrete mathematics” provides the greatest recall value, yielding 81% of the final 47 relevant papers.

Source of publications

Finally, Table 3 shows a breakdown of the 67 title filtered papers according to the search term used and the source of the publication. Just under three-quarters of the relevant papers came from four sources, with the top eight sources accounting for nearly 90% of the papers. These results must be qualified by noting that most came from the ACM DL, thus causing a natural clustering, but are nonetheless broadly in line with expectations.

4. ANALYSIS AND SYNTHESIS

Having acquired the 47 source papers, the final step is to read, analyse and synthesise the results. While we cannot provide universal recommendations for any syllabus in the subject, we hope to at least provide a framework within which they can be worked out.

In this section we discuss the research in teaching discrete structures under five broad themes:

• Whether discrete structures should be taught as a single course, or have its content distributed over multiple courses.
• When students should study discrete structures
• What should the material be linked with: we have identified two main overlapping areas: with programming and with data structures.
• How should the course be taught; in particular, what teaching methods have been identified.

Table 4 presents a summary of the papers by each of these themes, and Figure 2 shows the distribution of the themes over time.

A further theme, the evaluation of course success, is discussed at the end of this section.

Whether to teach DS

This somewhat existential question is not so much related to whether the topics in discrete structures should be taught, as to whether they should be taught in a single course, or moved into the courses where they are applied. Examples include moving logic into software verification or formal methods or moving all study of trees and graphs into data structures. Similar questions arise when considering the mathematical techniques relevant to computer science, such as calculus or statistics, and appear across a large time span, from Berztiss, Tremblay and Manohar in the 1970’s to Fleury and Neff in the 1990’s and 2000’s [6, 43, 12, 32].

It is highlighted quite frequently that one of the main ideas of teaching students discrete structures is to prepare them for more theoretical content they meet in courses later on [25, 27], which brings up the question that if computer science majors are finding the content too difficult or aimless so early in their studies, why not push it back to “where it belongs”? With Tremblay and Manohar for example, counting techniques, permutations and probability are omitted from their discrete structures course (although part of the curriculum) as they are covered in other courses [43].

A more novel approach is taken specifically by Neff [32], where parts of the discrete structures course are outsourced and the subject is then taught with an approach of rather than “Here’s the tools, you’ll need them sometime”, it says “Here’s a computer science problem we need to solve, what tools do we need to solve it?”.

When students should learn DS

Marion draws us back to our previous point about dismantling the subject, but instead segments the course rather than completely dissolving it [27]. This can allow an eas-
ier introductory course for discrete structures and a subsequently more in-depth course later on when students are more experienced. But the appreciation of the subject is always a problem early on, with a difficulty noted in teaching students material that they believe they may never use [13]. If discrete structures is to serve as a basis for a wide variety of more in-depth modules later on, it needs to justify its existence by demonstrating its necessity and application.

**DS with Programming**

A number of papers describe curricula that involve coding up a discrete structures related problem. One approach uses real world examples in order to aid students in the applicability of what they are being taught and motivate them to excel at the course work [39]. Martin proposes allowing students to implement an algorithm of their choice which they learnt in their discrete structures course, enforcing their understanding of the material [28]. Cigas and Hsin, introduce the use of specification for problems; one of the primary applications of the logic learnt in discrete structures [10].

There is also the opportunity to exploit the link between discrete structures and declarative programming languages. For example, Berry describes a more advanced application of coding involving the use of Scheme [5]. As a language which contains elements of functional programming there is already a clear link between this approach and the functions aspect of discrete structures. Similarly, Hein outlines the use of FP, ML and Prolog to teach students discrete structures concepts [15]. Clearly, an issue here is how this interacts with the teaching of programming and language paradigms in the syllabus as a whole.

McMaster et al. draw the conclusion that teaching discrete structures with programming makes the subject easier for students - or at least for those students who already know how to program [31]. Implementing a theory learnt in discrete structures as a program requires a full understanding of it, rather than the regurgitation of a rote learnt mathematical proof. However, it is important not to turn discrete structures into another programming course, since omitting the theoretical material could lose much of the intended flavour of the subject.

**DS with Data Structures**

Another noticeable feature of many of the papers reviewed is an apparent crossover between topics covered in discrete structures and data structures (CS2). The ACM curriculum for discrete structures contains a number of intersections with data structures such as graphs and trees and such a crossover (as well as some others) is further explored by Decker and Ventura [11].

Even though the context in which these topics are to be studied is different between the two subjects, this again raises questions regarding the coherence of a discrete structures course in a syllabus. A number of the papers note that the concept of coding up actual algorithms in discrete structures is very similar to the way in which data structures are studied [39, 28, 5, 31]. Teaching an inductive structure such as a list or tree, and then asking students to implement them, particularly in a functional language [34], is similar to implementing these as data structures, albeit often with different motivation and analysis mechanisms.

A different angle on this issue arises from studies of factors influencing student success at discrete structures. There is significant coupling between data structures and discrete structures when evaluating student performance [6, 40, 34, 32].

---

**Figure 2: Distribution of topics over time. Each bar represents the number of relevant papers over a five-year period, subdivided by the topic covered.**

---

11], giving strong support to closely examining this link in a syllabus.

**How to teach DS**

Finally, in response to the uncertainties of educating students in discrete structures a number of teaching methods are described in the literature.

There is much focus on reducing the unfamiliarity of the subject. For example Zeng and Jiang builds on the foundation of staying in a student’s comfort zone when introducing a new idea [51]. This appears to ease students into the formality of certain areas of discrete structures. Berry, Cigas and Hsin use visual tools to try to reduce the complexity for students [10, 5]. Exploitation of a student’s familiarity with programming is also an interesting way to tackle teaching problems since students may be a lot more comfortable in a software engineering setting than in a pure mathematical one in which discrete structures is sometimes presented. To deal with unfamiliar notation, a simple reinforcement of exactly what a symbol means can uncloud a problem for a student, allowing them to focus on understanding the theory rather than the theory’s representation [21].

The collaborative approach discussed by Buchele takes a more traditional pen and paper approach but the increased student engagement can certainly be seen as beneficial [9]. Collaborative homework also enables mingling of those who understand certain concepts and those who don’t, hopefully making the educational experience “viral”. Similarly, Neff suggests that a problem-directed approach can go some way to tackling the appreciation and motivation problems with discrete structures [32].

It is interesting that while the previous themes explored the relationship of discrete structures with other subjects, many of the papers on approaches to teaching seek to exploit this relationship positively.
Table 5: Ways of evaluating course success.

<table>
<thead>
<tr>
<th>Category</th>
<th>Sources</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade Based Statistics</td>
<td>[11, 19, 30, 34, 35, 40, 50]</td>
<td>7</td>
</tr>
<tr>
<td>Survey Based Statistics</td>
<td>[9, 26, 28, 37, 38, 42, 45]</td>
<td>7</td>
</tr>
<tr>
<td>Anecdotal Feedback</td>
<td>[6, 7, 8, 14, 16, 17, 21, 24, 31, 41, 43]</td>
<td>11</td>
</tr>
</tbody>
</table>

4.1 **Evaluation of Course Success**

Given that a very large number of papers discuss some kind of discrete structures course it is important to analyse the ways in which these courses are evaluated. We found that a total of 25 papers discussed some form of evaluation, though this varied between papers. Overall we can identify three broad categories of information regarding the evaluation of a teaching approach: based on student grades, based on a quantitative analysis of student feedback, and based on anecdotal information from students. The papers falling into each of these three categories are shown in Table 5.

Notably, it appears that statistics in some cases can be quite sparse when it comes to discrete structures. With Martin for example, a set of 20 students may not tell us as much as a set of 100 students would when it comes to certain questions [28]. Going by statistics alone, in taking radical approaches such as scattering discrete structures course content to later on modules we effectively move into unknown territory. For this reason, it is necessary to collect a significant amount of statistics for a wide variety of approaches taken to teaching the subject. Having said that, there are some more sizable data sets which analyse the final grades for students [11, 34, 40].

As well as these grade-based statistics there are also some survey-based course evaluations which provide more anecdotal type feedback on courses [28, 9]. This anecdotal type evaluation is a lot stronger in some papers than in others, sometimes with just a sentence or two given for student thoughts on a course [6, 21, 31, 43]. There appears to be quite some variety in course evaluation techniques and as a result it can be difficult to fully evaluate, much less compare, course outcome analyses.

5. **CONCLUDING REMARKS**

This paper has presented a review of research on discrete structures teaching. Since the initial submission of this paper our search parameters were extended to include the ERIC collection, however this did not result in any significant changes to this review; details can be found in the associated technical report [48].

We hope this work will be of use as a reference point for those preparing to teach discrete structures, and as an example of a systematic literature review in CS education.

6. **ACKNOWLEDGEMENTS**

This work was partially funded by the National Academy for Integration of Research, Teaching and Learning (NAIRTL).

7. **REFERENCES**


[19] L. Jamba-Joyner and W. F. Klostermeyer. Predictors...


