Version

This is the Author’s Original version of the following article. This version is defined in the NISO recommended practice RP-8-2008
http://www.niso.org/publications/rp/

Suggested Reference


Copyright

Items in ResearchSpace are protected by copyright, with all rights reserved, unless otherwise indicated. Previously published items are made available in accordance with the copyright policy of the publisher.

For more information, see General copyright, Publisher copyright, SHERPA/RoMEO.
Teaching and learning data structure concepts via Visual Kinesthetic Pseudocode with the aid of a constructively aligned app

Ogen Odisho, Mark Aziz, Nasser Giacaman
Department of Electrical and Computer Engineering
The University of Auckland, New Zealand

Abstract—Data Structures is an integral topic for any Computer Science or Software Engineering degree, identified as a Core Tier-1 topic of the ACM/IEEE Computer Science Curricula. The underlying concepts are inherently abstract, making them especially difficult to understand for novice programmers. This paper proposes a cognitively challenging technique to help students understand the thought process that the learning outcomes of fundamental data structure units aim to achieve. The development of this thought process is using a technique we term Visual Kinesthetic Pseudocode, with the overarching goal of helping students code without coding, yet providing the necessary scaffold to guide them in implementing the data structures with real code. This was implemented in the form of INTERACTIVE DS, an app for students and teachers to guide the learning of fundamental data structure concepts. The evaluations demonstrate that students strongly credited INTERACTIVE DS with the aid of a constructively aligned app in aiding their understanding of concepts and confidence in applying data structure concepts in practice. The study is also a step forward in revealing potential threshold concepts pertaining to data structure modules.

I. INTRODUCTION

The fundamental data structures unit is one of the most cognitively difficult topics in programming due to its high level of abstraction [1]. The learning is further exasperated since it builds on object-oriented programming principles and pointers -- both of which are concepts identified as troublesome threshold concepts in computer science [2], [3]. While visualization has been attempted in helping students, they are deemed ineffective unless they cognitively engage and challenge students [4], [5]. This paper proposes Visual Kinesthetic Pseudocode (VKP) as an engaging technique to scaffold students transitioning “from concept to code”.

Programming in general is a difficult subject to learn: it involves many abstract concepts and students rarely receive sufficient amounts of personal instruction [1]. The ACM and IEEE Computer Science Curricula [6] has identified data structures as one of the most essential topics, with the recommendation that 12 of the introductory hours dedicated to it. The issue in learning data structures lies in the high-level concepts rather than at the low-level programming technicalities [7]. Such topics are not limited to Computer Science and Software Engineering; Electrical and Electronics, Mechatronics, and Computer Systems Engineering are also examples incorporating CS2 level topics.

Since some of the difficulties associated with learning data structures is also acknowledged by teachers [1], [3], this raises some important questions. How can the underlying concepts of the fundamental data structure unit be taught, and therefore learnt by students, effectively? How can teachers ensure that their explanations are correct without inadvertent errors that could potentially further confuse students? Answering these questions requires a careful approach that is educationally sound, both in terms of motivating students as well as meeting the correct learning outcomes. Instructors teaching programming courses seek and welcome scaffolding tools that support their delivery of such abstract concepts [8]. This scaffolding also plays an essential role for students, particularly when attempting to implement the concepts with practical programming exercises.

A driving motivator for this work is that textual pseudocode lacks engagement and is easily misinterpreted [9]. VKP aims to reinforce the correct thought processes, which is a precursor for students understanding how to implement something in code. Using the ideas of VKP, INTERACTIVE DS has been implemented without losing sight on delivering the intended learning outcomes pertaining to data structures. INTERACTIVE DS ensures only programmatically-correct steps may be executed, with interactions strongly aligned to pseudocode such that any given action correlates to a “real line of code”. This serves two purposes:

• Eliminate any inadvertent errors likely to arise when explaining data structure concepts, especially when the explanation is attempted in a visual manner.
• Provide a scaffolded learning environment to reduce ambiguity when transitioning from the concept to implementation, especially when conceptual explanations tend to be too abstract in helping students independently apply the concept in practice.

The rest of this chapter is organized as follows. Educational foundations guiding this work are overviewed in section II. Visual Kinesthetic Pseudocode is presented in section III, before introducing INTERACTIVE DS in section IV. Student experiences and potential threshold concepts are discussed in section V before presenting a brief overview of related work in section VI, then concluding in section VII.

II. BACKGROUND

A. ACM & IEEE CS2013: Fundamental Data Structures unit

The ACM and IEEE Computer Science Curricula (CS2013) [6] is a guideline for Computer Science and Software Engineering undergraduate degree curriculum design. The core topics are decomposed into Core Tier-1 and Core-Tier 2, with most of the Core Tier-1 topics covered in introductory courses. CS2013 identified 165 hours dedicated to Core Tier-1 topics, 43 of which are dedicated to Software Development
B. INTERACTIVE DS educational theories

1) Levels of thinking about teaching: The primary design decision for INTERACTIVE DS was to promote an active learning approach. This is motivated by the three levels of thinking about teaching that an instructor may take in teaching [11]:

1) What the student is: the teacher believes responsibility to learn lies solely with students, and little can be done to motivate them.
2) What the teacher does: the teacher attempts to interest students by focusing on the way material is presented.
3) What the student does: the teacher recognizes that learning comes from student engagement.

To foster learning, focus needs to be on what the student does [12]. Here, students learn by actively applying older knowledge as they encounter new knowledge [13]. Such active learning is vital for long-term retention, in which students make meaning of the new information by relating it to existing knowledge [14]. INTERACTIVE DS promotes this, as it comes in the form of activities to engage students, but is still also useful as a teaching aid for the instructor when explaining concepts in class.

2) Constructive Alignment: Engagement is not the only ingredient required for learning; an engaged student may not necessarily be learning the intended outcomes. Constructive alignment [15]. Rather than focusing on topics the teacher should teach, constructive alignment first focuses on the Intended Learning Outcomes (ILO): what students should learn and to what extent. In the case of the data structures topic, the learning outcomes are provided by the CS2013 learning outcomes (as discussed in section II-A). The second step of constructive alignment involves the identification of Activities that engage students in order to meet those ILO. LO#3 states that students should be able to “write programs that use the data structures”. While an obvious activity might entail a lab exercise requiring students to code and use the data structure, this is likely to be a daunting task for weaker students. INTERACTIVE DS is built on activities that serve as stepping stones. Finally, Assessment Tasks are required to provide feedback informing students about their learning progress. In the context of INTERACTIVE DS, this is implemented with randomly generated exercises to manipulate the data structures. This process of accomplishing sub-goals (that lead to larger goals) helps promote student self-motivation, ultimately nurturing self-efficacy [16], [17].

III. VISUAL KINESTHETIC PSEUDOCODE

This section introduces Visual Kinesthetic Pseudocode (VKP) as a strategy to scaffold novice programming students when studying abstract and concept-rich topics that require a high level of mastery. The goal is to provide a stepping stone in understanding the concept (i.e. the thought process), as this is a prerequisite to the actual coding (i.e. practical application of the concepts). This is illustrated in Figure 1, showing the support steps a student requires before being able to apply to real programming exercises (such as LO#3). Although the intention of textual pseudocode is to make it easier to understand real code by expressing it informally, it still possesses some problems. Novice programming students find pseudocode easy to misinterpret, and it does little in helping them detect their misunderstanding due to the lack of feedback it provides [9].

Applying VKP is not limited to learning outcomes that incorporate code development; VKP can also provide a stepping stone when targeting a high level of conceptual understanding, such as appreciating the performance consequences of alternative data structure implementations (e.g. LO#4). The kinesthetic aspect of VKP would allow students to “experience” the performance associated with different implementations. The important components of the VKP strategy are:

- Visual: This recognizes that code (whether it be real code or pseudocode) can be daunting for novice programmers, and the peculiarities surrounding each programming language’s syntax only further inhibits understanding of the underlying concepts. While textual pseudocode has the benefit of being programming language-neutral, visualizations help students develop mental models [5].
- Kinesthetic: Visualizations alone are insufficient in helping students learn; for them to be effective, they need to cognitively engage and challenge students [4], [5]. This component therefore promotes active learning.
Figure 2. An example incorporating visual and kinesthetic considerations to produce a VKP. The careful alignment to the textual pseudocode ensures the interactive visualization meets the intended learning outcome. While the VKP is more engaging and less daunting for novice programmers, the direct alignment will still allow the automatic generation of textual pseudocode from the corresponding actions.

![Visual kinesthetic pseudocode: adding a new value to the front of a singly linked list](image)

**Textual pseudocode:** adding a new value to the front of a singly linked list

```
begin add_to_start(value)
  create a new node storing value;
  set the new node’s next reference to the current head node;
  set the new node as the head node;
end
```

**Visual kinesthetic pseudocode:** adding a new value to the front of a singly linked list

![Diagram of VKP](image)

Figure 3. An activity using VKP to add a new node to the front of a singly linked list. Each kinesthetic action is directly related to corresponding code (shown at the bottom). The “sliding” effect (e.g. Step 2 in Figure 2) is shown in (a) to (b). This aids the understanding of pointers, by avoiding the incorrect approach of attempting to directly access nodes.

![Image of VNPP](image)

- **Pseudocode:** For engagement to contribute to learning, it needs to be closely aligned to the ILO. The Fundamental Data Structures unit requires a high level of mastery that requires students to apply the thought process in manipulating data structures. This process, if documented as pseudocode, is a helpful precursor to writing “the real code” [18]. The difference with VKP is that pseudocode is constructed by students visually and kinesthetically.

**Aligning visualization and kinesthetics to pseudocode:** Figure 2 illustrates traditional textual pseudocode for a learning activity to manipulate a data structure, in this case adding a new value to the front of a singly linked list. This example illustrates the potential misinterpretation that can result with using textual pseudocode; does line 3 mean “head = newNode”, or does it mean “newNode = head”? Due to the informal language of pseudocode, it is easy to see how misinterpretation might creep in. Below the textual pseudocode is the same learning activity, this time infused with visual and kinesthetic components to produce the VKP. Here, misinterpretation has been eliminated since each kinesthetic action is intentionally designed to be mapped to real code.

![Images of VNPP](image)

![Images of VNPP](image)

![Images of VNPP](image)

![Images of VNPP](image)

(a) (b) (c) (d)

(a) The plan: update head
(b) “Expectation failure”

Figure 4. The learner’s goal is to add a new node to the front of the singly linked list. The learner’s first step in the plan, making nodeToAdd as the head, immediately results in an expectation failure.

IV. INTERACTIVE DS

The design recommendations discussed in the previous sections have been implemented in INTERACTIVE DS, in the form of an app freely available for both students and teachers. Figure 3 illustrates INTERACTIVE DS screenshots in correctly adding to the front of a singly linked list. Each step in the VKP interaction (outlined in Figure 2) directly maps to pseudocode.

**Learning with expectation failure:** To promote deep learning, students need to have their existing mental models challenged [19]. Long term understanding is aided with the intellectual stimulation of grappling with a problem. As an essential part of this learning process, INTERACTIVE DS capitalizes on the concept of “expectation failure” [20]. This is where the learner has a goal, follows a plan, but fails when the result does not meet their expectation. This is best illustrated with Figure 4, where the user’s goal is to add a node to the front of a singly linked list. It may seem simple enough, but this particular activity catches out many users when attempting it for the first time. When the wrong action is taken, the linked list loses reference to the first 3 elements. INTERACTIVE DS recognizes that the data structure is in an incorrect state, so it immediately highlights the lost nodes and vibrates to inform the user (Figure 4(b)).

**Reducing misconceptions:** A challenge in learning a conceptual topic is that, if the explanation is too abstract (in the aim of focusing on the conceptual explanation), it does little in helping the learner apply the concepts. Furthermore, visually explaining concepts may lead to misconceptions when the visualizations do not correctly reflect the intended and correct

---

1 www.ece.auckland.ac.nz/~ngia003/dsapp
meaning. For example, when removing a node from the middle of a linked list, it may be tempting to “directly point” to the middle node. However, just because “we can see it” in the visualization, it does not mean we can actually access it yet. The kinesthetic interactions of INTERACTIVEDS help learners understand they cannot manipulate or directly reference nodes until they access them by traversing the data structure; Figure 5 shows two steps of this activity, here the use of a probe pointer to navigate to the middle of the doubly linked list. If a teacher was explaining this activity without reinforcing these rules (e.g. by using a freehand sketch), students might be misled into believing operations on the middle of a linked list are constant-time operations.

Another added challenge in learning how to manipulate data structures is its heavy use of pointers, which is a threshold concept in itself that novice programmers struggle with [2], [3]. Consider again Figures 3(a) and 3(b). In setting the value of nodeToAdd->next, the user learns that they cannot directly point to Node 1; instead, the value of head is used (Figure 3(a)), which effectively causes it to point to the node (hence the “sliding” effect that results in Figure 3(b)). INTERACTIVEDS ensures that the correct intention of using pointers is respected, avoiding misunderstanding of pointers.

V. EVALUATION

INTERACTIVEDS was offered to a CS2 course consisting of 250 students. The course is compulsory and includes Electrical and Electronic, Mechatronics, and Computer Systems Engineering students. The data structures unit spanned two teaching weeks, focusing on the core learning outcomes presented in section II-A. The logs in this section were gathered from the Android and WebPlayer platforms, from 159 and 38 unique devices respectively. At the end of the data structures unit, students were asked to complete an anonymous questionnaire delivered in-app, where 49 responses were received.

Limitations: It was an informed design decision to encourage student uptake by ensuring privacy was respected. Since data gathered was anonymous, this meant that the questionnaire and activity logs could not be correlated to particular individuals, let alone to course assessments. Although individuals could not be identified, it was possible to distinguish usages on different devices (based on unique device ID). It is therefore possible that a student may have installed the app on multiple devices (e.g. a tablet and a smartphone), or used the WebPlayer from multiple browsers; in such cases, this study treats different devices as different users. Only logs coinciding with the delivery of the data structure module are included, maximizing the likelihood that they were from students genuinely enrolled in the course. Finally, receiving logs was at the user’s discretion: users are allowed to opt out of sending logs by disabling this from within the app settings. While the app may have been used more than is portrayed by the logs, but it is not possible to determine how many users disabled sending anonymous data.

A. Student perceptions

Table I summarizes anonymous responses (using the 5-point Likert scale) from the in-app questionnaire. The questionnaire was visible only to users that completed at least half the activities (to ensure only genuine users completed it). The app most effectively developed student confidence for Array, Vector and Linked List activities (Q1-Q3). These activities were the most cognitively challenging activities, requiring users to interact using the VKP technique proposed earlier. This is best illustrated with the following student comment, emphasizing learning through expectation failure:

“I found this easy to learn because I was able to interact with the data structure, make mistakes, then see how the order of doing things played a part.”

The other activities were less cognitively challenging, especially the Queue and Stack activities since the aim of their semi-automated interaction was to emphasize the code-reuse of other data structures (Q5). Many students commented on the lack of interactivity for these activities, therefore confirming mobile app activities need to be cognitively challenging:

“The fact they were only animations as opposed to exercises meant less thinking was required.”

Q6-Q10 show that students generally agreed that data structure concepts were effectively conveyed with INTERACTIVEDS. The CS2013 Learning Outcomes however require a much deeper level of understanding; in this regards, Q12 and Q13 show that INTERACTIVEDS helped provide some scaffolding for students to bridge concepts with coding activities. This is emphasized with Q14, where a fair portion of students actually wanted to see more code. The app was useful in guiding students towards implementing with code:

“Being visually able to work through linking forward and behind, then deleting the nodes, really helped concrete the process required to then implement it into my lab and assignment. Also being able to see pseudocode after completing the task is a real help.”

Some students hinted they wanted to see more engaging and challenging exercises (Q15), but otherwise had a highly positive attitude to the educational benefits of using mobile apps in their learning (Q16-Q19). The combination of visual and kinesthetic pseudocode was a powerful combination:

“The use of visual diagrams with code provided after successful completion showed what was going on with each step. Drag and drop feature made it clear what steps I was doing to get my intended result.”

B. Logged usage

During the two week data structure module, over 5600 activity logs were recorded from the Android and WebPlayer platforms. Table II shows the time distribution of all activity attempts, to help convey the granularity of engagement.
operations in the middle of the SLL are most challenging, while operations on the end of the DLL least challenging. It may come as a surprise that the successful completion rates of the SLL operations are higher than that of the DLL operations, but this is likely attributable to predominantly stronger students attempting DLL activities. Removing from the middle of a DLL is clearly the most challenging operation, while operations on the end of the DLL least challenging.

For a user having successfully completed a given linked list activity, we next investigate the total number of overall app activities completed. This aims to understand if the user’s general engagement is an attribute that contributes to them completing the more challenging activities. Building on the assumption made earlier that DLL activities are predominantly targeted by stronger users, we distinguish between two groups of users: those that have successfully completed half of DLL activities (namely 3+ of the 6), versus those that have not. Figure 6 shows a significant distinction of general engagement in completing activities between the two groups. The value above each whisker is the total number of students in that group that successfully completed the respective linked list operation, while the y-axis shows the group’s distribution of overall app engagement. This figure helps explain, for example, why the DLL completion rates are higher than the SLL rates for operating in the middle of the linked list; it is due to the DLL being predominantly targeted by the more engaged (and presumably stronger) users.
scaffolding novice programmers to learn an abstract topic such as data structures. The activity logs also identified the most challenging operations in manipulating linked lists.

REFERENCES