Collaboration, Collusion and Plagiarism in Computer Science

Robert Fraser
David R. Cheriton School of Computer Science
University of Waterloo

Abstract
This report contains an overview of university policies toward academic dishonesty, and the efficacy of these policies upon student education. We discuss the idiosyncratic nature of how collaboration, collusion and plagiarism are defined by students, instructors and administration. After considering some of the common reasons for dishonest behaviour among students, we look at some methods that have been suggested for mitigating them. Finally, we consider some ideas for improving computer science courses in this context. We suggest emphasizing the intended learning outcomes of each assignment, providing tutorial sessions to facilitate acceptable collaboration, and finally to deliver quizzes related to assignment content after each assignment is submitted.

Robert Fraser may be contacted at r3fraser@uwaterloo.ca
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>3</td>
</tr>
<tr>
<td>Academic Dishonesty in Computer Science</td>
<td>3</td>
</tr>
<tr>
<td>Collaboration</td>
<td>4</td>
</tr>
<tr>
<td>Collusion</td>
<td>5</td>
</tr>
<tr>
<td>Plagiarism</td>
<td>6</td>
</tr>
<tr>
<td>Mitigating Dishonesty</td>
<td>7</td>
</tr>
<tr>
<td>Detection &amp; Discipline</td>
<td>8</td>
</tr>
<tr>
<td>Foundational Knowledge</td>
<td>9</td>
</tr>
<tr>
<td>Entitlement</td>
<td>10</td>
</tr>
<tr>
<td>Motivation</td>
<td>11</td>
</tr>
<tr>
<td>Discussion</td>
<td>11</td>
</tr>
<tr>
<td>The Status Quo</td>
<td>12</td>
</tr>
<tr>
<td>Unlimited Collaboration</td>
<td>12</td>
</tr>
<tr>
<td>Improving Tutorials</td>
<td>13</td>
</tr>
<tr>
<td>Intended Learning Outcomes</td>
<td>14</td>
</tr>
<tr>
<td>Quizzes</td>
<td>14</td>
</tr>
<tr>
<td>Conclusions</td>
<td>15</td>
</tr>
<tr>
<td>References</td>
<td>16</td>
</tr>
</tbody>
</table>
Introduction

This paper presents an examination of group work policies and academic dishonesty in computer science. The field is subject to a significant amount of cheating, and we begin by looking at why computer science may be prone to some of these practices. We then take a detailed look at collaboration, collusion, and plagiarism in turn, and determine where the lines are drawn between them. We follow up with a section on mitigating dishonesty, where we review some of the reasons that students cheat in general, and some of the practices that have been used to try to reduce dishonesty. Finally, we present a discussion on how these lessons may be used to improve the design of a course in computer science, both in terms of reducing academic dishonesty and improving student learning.

Academic Dishonesty in Computer Science

It is telling that there exist a plethora of studies dedicated to academic dishonesty specifically within computer science and information technology. There are many reasons for this focus, and it has been shown that computer science accounts for more than its share of incidents. For example, Roberts (2002) reviewed incidents of dishonesty at Stanford University over a decade, and found that 37% of incidents were due to computer science, while their students represent less than 7% of the student body at the university. MIT, one of the top computer science schools in the world, made headlines when an introductory computer science course was found to have rampant plagiarism, resulting in disciplinary action against 73 students.

Whether the high rates of collusion and plagiarism incidents among computer

---

1 Business programs also claim to be prone to significant dishonesty, see for example (Levy & Rakovski, 2006).

2 At Stanford University, such incidents are referred to as honor code violations.
science students are due to higher rates of dishonesty in computer science or due to better detection tools is open to debate. Barrett and Cox (2005) hypothesize that mathematics and computer science are subject to higher rates of collusion because of the nature of the work. While arts and humanities assignments typically involve personal reflection on the parts of the students, assignments in this discipline typically have an ideal solution which the students are seeking. For the same reason, it is difficult to detect dishonesty amongst correct solutions; it is similar incorrect or strange solutions that usually raise alarm bells. Striving to find the ideal correct solution can be an exhausting and frustrating experience. The nature of computer science assignments was encapsulated beautifully by (Roberts, 2002): “The computer is a relentlessly unforgiving arbiter of correctness.”

Collaboration

Collaboration is distinguished from collusion in this discussion by defining both activities as a group of students working together on an assignment, but collusion is group activity that is unpermitted. As an example, most courses permit collaboration in the sense that students may meet to discuss ideas for an assignment, but students are prohibited from writing the assignment solutions together, which would be considered collusion.

Collaboration can be beneficial for student learning. As educators, our primary objective is for the students to truly understand the material that we are presenting, which may be achieved through active learning. Fink (2003, p.108) promotes active learning by a variety of means, but a key aspect is through discussions of the course material with their peers. Assignments provide perhaps the ideal setting for such discussion, and the challenge is to encourage discussion to the point that it is acceptable.
Collusion

Overall, the boundary between collaboration and collusion is ill defined and highly idiosyncratic. Indeed, surveys have shown that although students and educators are able to clearly identify plagiarism when it occurs, the same cannot be said for collusion (Barrett & Cox, 2005).

Collusion may be regarded as middle ground in a spectrum of practices ranging from collaboration to outright plagiarism, and it is perhaps best defined as unpermitted collaboration. At the University of Waterloo, the Office of Academic Integrity (OAI) places the onus upon instructors to specify how much collaboration is allowed (Mitchell, 2011):

For assignments, clearly indicate if group collaboration is acceptable (and the level of collaboration permitted) or if students must do all work independently.

On the other side of the equation, the OAI offers students advice for determining what level of collaboration is permitted when conducting group work (Mitchell, 2011):

Working in study groups or teams can be very beneficial, but you need to be very careful of the limitations and expectations. Check the course syllabus carefully to determine if group work is acceptable; if so, make sure you understand the level of collaboration allowed. ... If you do work with others without explicit authorization from your professor, you run the risk of collusion and may face academic penalty.

It is evident from this statement that as far as the OAI is concerned, the boundary between collaboration and collusion is flexible, and varies depending on the desires of the instructor for a given course. Often the instructions for completing an assignment
are something of the form “You may discuss your ideas as a group, but the assignments must be written individually.” This may lead to confusion among students, for the boundary between what constitutes discussion and writing varies.

As an example, Barrett and Cox (2005) provided a set of scenarios to students and instructors, and asked whether the students had acted acceptably, or whether they are guilty of plagiarism, collusion, or both. The following scenario (Scenario 3 in their study) elicited the highly personal divide between what defines acceptable practice or collusion:

Student A doesn’t know how to start the assignment and so he asks student B who helps him by showing him his own work. Student A writes up the assignment in his own words but there are some similarities with student B’s work.

For this scenario, 51% of instructors thought that this would be fine, while 39% defined the activity as collusion. Interestingly, among students, only 38% thought that it would be acceptable, 33% felt that it is collusion, and 24% considered it to be outright plagiarism (Barrett & Cox, 2005, Figure 3). These results underscore the importance that an instructor establishes a clear definition of what constitutes collusion in the course being taught.

In general, collusion is regarded in a softer light than plagiarism. In fact, many educators are willing to tolerate collusion, since students are actually thinking and learning in the process (Barrett & Cox, 2005).

Plagiarism

Plagiarism is “the action or practice of taking someone else’s work, idea, etc., and passing it off as one’s own; literary theft” (OED, 2011). For this discussion, we
consider plagiarism to be the act of copying someone else’s work, whether or not the original author is aware of the act. This is particularly a problem in computer science, where the ideal solution to a problem may be obtained with a simple cut-and-paste operation.

Plagiarism is regarded as a serious academic offense, and institutions go to great pains to impress upon the students the severity of the penalties for students who are caught (Roberts, 2002). In general however, students have a poor perception of what constitutes plagiarism and will often err too heavily on the side of caution (Ashworth, Bannister, & Thorne, 1997). This has been reinforced by observing that the amount of plagiarism actually taking place has been declining, at least amongst computer science students (Culwin, 2008), although the general perception is the opposite (Dick et al., 2003). Certainly, evidence presented by Dick et al. (2003) suggests that students tend to cheat when given the opportunity to do so.

Mitigating Dishonesty

In this section, we review some of the common reasons for dishonest behaviour, and we examine some of the practices that are commonly used to discourage or deter such practices. This topic has been the subject of numerous studies, many of which are based upon questionnaires which ask students to self-report on their cheating or ask students and instructors to speculate on the cheating behaviour of students.

Franklyn-Stokes and Newstead (1995) conducted a study where they asked students and educators to rank different types of cheating behaviour by seriousness, and also by their perceived frequency. They identified an inverse relationship between these rankings, so that behaviours that are considered more serious violations occur less frequently. One observation based upon these results is that cheating on exams is considered more serious and occurs less frequently than cheating on coursework,
but another is that plagiarism is generally considered more serious than collusion.

Detection & Discipline

One of the significant challenges for mitigating academic dishonesty is detection. Dick et al. (2003) refer to statistics showing the rate of cheating in courses ranging from 40% to 96%, while the rates of detection may be as low as 1.3% (Björklund & Wenestam, 1999). The improvement of detection techniques is being aggressively undertaken, as it is believed that high rates of detection and prosecution will deter cheaters. If students believe that cheating is commonplace, then they are more likely to cheat themselves, as they feel that this creates a level playing field.

There are a number of commercial software applications which have been created to aid in the detection of plagiarism, and a recent study found that over a third of instructors are using such tools (Dick et al., 2003). At Waterloo, the Office of Academic Integrity (Mitchell, 2011) recommends the use of Turnitin, and has made the tool available to all instructors on campus.

The penalties if students are caught cheating range in severity from receiving a zero on the offending assignment or test, to expulsion from the university. While severe penalties are a deterrent, a survey of students by Sheard, Markham, and Dick (2003) found that it ranks fifth on a list of deterrents, below factors such as knowing the value of their work and pride in their work. Simply appealing to the moral code of the students does not reduce cheating rates, but the credible threat of detection and punishment has been demonstrated to reduce the amount of cheating in a course (Hollinger & Lanza-Kaduce, 2009). However, excessive punishment, exemplified by a zero-tolerance approach to prosecuting academic dishonesty, may deter even honest students from enrolling in a course (Levy & Rakovski, 2006).

Professors are busy and some regard dealing with cases of academic dishonesty
as a waste of their time. In a survey of nearly 500 university professors in 1996 (Björklund & Wenestam, 1999), it was found that 20% had ignored cases of blatant cheating. More recently, a 2005 survey of 59 professors revealed that 51% had ignored cases of suspected cheating (Barrett & Cox, 2005). The policy at the University of Waterloo is very clear on how much personal judgement is allowed when an instructor is considering whether to prosecute suspected cheaters (Mitchell, 2011):

UW faculty are **required to report** any instances of academic misconduct - even relatively minor cases, or ones which you prefer to handle informally.

**Foundational Knowledge**

Students who are more comfortable with the course material are less likely to cheat (Ashworth et al., 1997). In particular, in computer science, students entering the first year are assumed to have a certain level of background knowledge. Students who fall short of the prerequisites are more likely to cheat than their better prepared peers (Dennis, 2004). Also, the level of maturity and personal motivation of students is inversely related to their tendency to cheat (Sheard, Carbone, & Dick, 2003).

While this may be a factor, some students will elect to cheat even in classes where students all students begin the course with roughly equivalent skill sets. This leads to poorer learning amongst the cheaters, and subsequently higher rates of attrition (Palazzo, Lee, Warnakulasooriya, & Pritchard, 2010). Roberts (2002) echoes this sentiment, in that cheaters may be students who have fallen behind in the material during the progression of the course and are completely lost when attempting assignments later in the term.
Entitlement

Students may be inclined to cheat because they feel a sense of entitlement to their grades. Such students may regard paying tuition as a transaction to purchase their grades and degree. In the cheating incident at MIT, many students felt that they deserved a good grade because they had invested a significant amount of time in the assignment (Butterfield, 1991).

However, perhaps unsurprisingly, students who cheat generally do poorly in courses. Palazzo et al. (2010) observed the performance of students with varying rates of cheating on the course examinations as the term progressed. As might be expected, students who were the worst offenders on the assignments performed the poorest on their exams. In particular, students who copied more than 50% of their assignment work earned nearly two letter grades lower on the final exam than students who copied less than 10% of the work on assignments. Further, while only 20% of students cheated on over 30% of the material, they represented 47% of the students who fail the course. These results are intuitive and hardly surprising, and evidence suggests that students are aware that cheating hurts them but they don’t care. In the Palazzo et al. (2010) study, students were shown a graph at the beginning of the term clearly demonstrating a decline in performance with increased cheating behaviour, but no decline in cheating was observed when compared with previous sessions of the course. Further, some students who were caught cheating rationalized their actions in precisely this context: “cheating isn’t bad because it only hurts you at test time” (Palazzo et al., 2010, p.8).

Some students confess that they are simply lazy (Sheard, Markham, & Dick, 2003). Dennis (2004) found that a related factor, that students ran out of time because they started the assignment too late, was the top reason that students plagiarized.
However, they specify that this is more attributable to poor time management than outright laziness.

Motivation

Perhaps the most interesting suggestion for reducing cheating is to ensure that the students are aware of the Intended Learning Outcomes (ILOs) for a given assignment. Dick et al. (2003) found that when students understand the reason that the work has been assigned to them, they are less likely to cheat.

Underscoring this value may reduce the sense of entitlement among students. Those students who feel entitled to a grade regard the assignment as a transaction with the instructor, in the sense that the instructor wants a good set of solutions from the student and will provide a good grade in return. By emphasizing the ILOs as an objective, the students may instead see the assignment as a means to understanding the material and hence to obtaining a better grade in the course.

Some students who cheated on assignments rationalize their actions by claiming that the assignment is a waste of their time or that they aren’t motivated to complete the task (Palazzo et al., 2010). To mitigate this, Palazzo et al. (2010) suggest that a course design which involves more teacher interaction results in students believing that the instructor is more concerned with their learning (rather than simply assigning a grade based on their performance). They assert that such modifications to the physics courses at MIT is a primary factor in the roughly 75% reduction in cheating that they have observed since implementing the changes.

Discussion

As the university has developed effective policies for dealing with plagiarism, this discussion will focus on choosing the boundaries for collaboration. The status quo
is that the professor is free to set the boundary between collaboration and collusion where they see fit, but there is no real guidance on what is an appropriate choice. We begin by discussing the drawbacks of the current model.

The Status Quo

In theory, the current model that many computer science courses use is ideal. As we mentioned previously, many courses encourage students to discuss the assignment questions, but the writing must done individually. The problem with this model is that many honest students err on the side of caution, and will avoid group work altogether to be safe. The result is that some of the best students are missing the active learning opportunities associated with the assignment because they feel that it is the most honest course of action.

Presently course instructors emphasize the acceptable limits of collaboration, but do little to facilitate or encourage students to work near the boundary of acceptable behaviour. The time constraints on a course is a significant reason for this, as there is hardly enough time in lectures for all of the course material as it is.

Unlimited Collaboration

One approach to reducing collusion would be to allow any amount of collaboration among the students, so that there is no need to check for plagiarism. In the extreme, this would permit outright copying of another students solutions if the student so desires. However, as discussed previously, Palazzo et al. (2010) demonstrated that providing students with the knowledge that cheating only hurts them and appealing to their maturity has no effect on reducing the level of cheating that occurs. It seems that simply permitting any level of collaboration would increase the amount of copying that would take place in the course, and damage the average student’s
learning experience.

*Improving Tutorials*

A first suggestion for improving course assignments is to encourage collaboration on the assignment questions during tutorial sessions where instructors and/or TAs are available to facilitate group discussion. Further, these sessions could be used for discussion of assignments that have been returned so that students may compare the solutions that they developed to the ideal solutions that were used for grading. If managed properly, these tutorial sessions would create active learning opportunities, so that students who are comfortable with the material may assist other students or engage in high level discussions with their peers. There are several challenges that would have to be addressed with such a proposition.

The first problem is simply a matter of time. It may be difficult to introduce tutorial sessions into courses which lack them at present. However, to begin with, this format could be attempted in a first or second year course which already has tutorial sessions. An added benefit of this proposal is that it may improve tutorial attendance. Tutorial sessions in these courses are often sparsely attended because many undergraduates don’t see the value of tutorials. A correlation between the tutorial sessions and assignments may help improve attendance and reinforce the intended learning outcomes for the assignment.

Another challenge is to minimize plagiarism, and these tutorial discussions may actually facilitate such actions. However, the prohibitions that are enforced presently would be maintained, and the same plagiarism detection techniques would be used. The tutorial sessions, if used correctly, may actually reduce the tendencies toward plagiarism. One of the major reasons identified earlier as to why students cheat was that they had poor time management or procrastinated, and they did not have enough
time to finish the assignment on their own. By timing the assignment discussions in tutorials so that they occur at least several days before the assignment is due, the students will be pushed into actively thinking about the assignment material with adequate time to complete the problem set properly.

Furthermore, the tutorials can help students keep up to date on the course material, which was identified earlier as one of the primary reasons for dishonest behaviour.

**Intended Learning Outcomes**

As educators, we often feel that the lessons to be learned by course work are self-evident. A course is usually designed so that the sections fit together well, and each section has a set of lessons and values that contribute to the skill set of the student. However, it is often the case that students lose sight of these values, particularly as they become pressed for time while juggling the workload resulting from taking a number of courses.

With each lecture and assignment, it is worthwhile to emphasize the intended learning outcomes. With respect to our discussion, the students should be provided with a short section at the beginning of the assignment which explicitly states what they should learn by completing the questions. This makes it clear that the students should master these concepts if they wish to do well in the course, and that this assignment is a vehicle for that purpose rather than just some busy work.

**Quizzes**

To ensure that students are putting thought into their assignments, the instructor could provide quizzes to the class after each assignment is submitted which tests their knowledge of the assignment material. To be successful, this quiz should
be fairly easy for someone who dutifully completed the assignment. The intention is that during collaboration sessions, weaker students would be compelled to understand the material more since they would be tested on it shortly.

Implementing quizzes would be particularly easy if the class were using a student response system such as clickers, so that the quiz could be marked instantly and students could compare their level of comprehension to that of their peers. Furthermore, this would facilitate some review during lecture at that time if it appeared that there was a common misunderstanding or knowledge gap in the class.

Conclusions

In this work, we have examined how and why students cheat in computer science. After distinguishing between collaboration, collusion, and plagiarism, we found that the differences are subjective in nature, and that a course instructor is responsible for ensuring that the students in the course are aware of where the boundaries lie. We identified that perhaps the most promising avenue for mitigating academic dishonesty at present is to reinforce the intended learning outcomes of the assignment material. Finally, we presented concrete suggestions for improving a computer science course so that the amount of cheating is reduced, and simultaneously, the learning environment for the students is improved.
References


