

Plagiarism

Sean Luke

Department of Computer Science
George Mason University

Students are often confused by what constitutes plagiarism in academia and the seriousness with which the offense is treated. The objective of this essay is to explain the issue and provide examples to explain what is plagiarism.

Plagiarism is the second worst thing you can do in academia (the worst thing is to fake your experiments). It is so bad of an offense that people who plagiarize are **often fired and their careers are over**. Authors who plagiarize in papers are often publicly denounced by their coauthors in public to save the coauthors' good names. PhD students who plagiarize are typically disowned by their advisors and thrown out of the program. Because it is so serious, not only do you want to avoid plagiarizing, **you want to avoid even the appearance of plagiarizing**. The last thing you want is for someone to even *suggest* that you might be *sort of* plagiarizing. It is that bad.

There are **two things** that can be plagiarized: someone's original words, or someone's descriptive idea. Lightly rewording the text is not sufficient: unless the ideas are well-known and general, you will be accused of taking someone else's descriptive idea and using them as your own.

When you plagiarize you have done harm to **two groups** of people. First, you have denied an original source the credit he is due for an idea or for words he said. Second, you have misled your readers into thinking that you came up with the idea or text.

Last, plagiarism brings up **two issues**. First, there is the plagiarism itself, which is a *moral or ethical offense*. Second, there may also be *copyright violation*, which is a *legal offense* for which you and your organization may be sued.

You cannot plagiarize anywhere in your paper. This includes not only the main work but also the abstract, introduction and conclusion, tutorial or explanatory material, and especially the previous work section. Even if *all* the original work of your paper is in your own words, if you have plagiarized from other sources *anywhere* in the text, it is absolutely, unequivocally unacceptable.

Below are some examples. The text, math, and pictures in the examples are from page 3 of a paper I co-wrote with Liviu Panait, entitled "A Pheromone-Based Utility Model for Collaborative Foraging" [1]. Before we go on, here's the original text:

As we will discuss later, although this model bears a strong similarity with reinforcement learning models, the assumption of symmetry makes possible update rules which are much more efficient than reinforcement-learning style "backup" rules like TD(λ) or prioritized sweeping. One such efficient update rule is essentially a modified form of value iteration though its application is somewhat unusual.

Imagine, as is shown in Figure 1 that the ant has just transitioned from location s to location s' , and at s' the ant now has available actions $a \in A$. The model provides the probabilities $T(s', a, s'')$ that performing a in s' will transition to s'' . Let γ be a learning rate constant between 0 and 1. The update rule would then be:

$$U_p(s') = R(s') + \gamma \max_{a \in A} \sum_{s''} T(s', a, s'') U_p(s'')$$

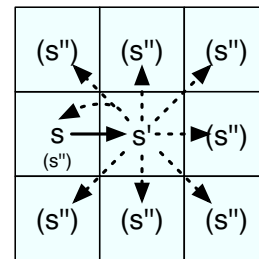


Figure 1: State transitions for a simple two-dimensional discrete grid environment. The ant was just in state s , and has just transitioned to state s' . From there the ant may transition to any state $s'' \in S''$. Note that s is also in S'' .

Example 1

In this and other examples, I list the plagiarizer's own text *in this font*.

The pheromone updating model is an effective technique for ant foraging. Although this model bears a strong similarity with reinforcement learning models, the assumption of symmetry makes possible update rules which are much more efficient than reinforcement-learning style "backup" rules like TD(λ) or prioritized sweeping. Another approach is to use reinforcement learning directly. For example...

Here the plagiarizer has simply cut and pasted text from the original source into his own work. This is plagiarism in its purest form.

Example 2

What if we include a citation?

The pheromone updating model is an effective technique for ant foraging. Although this model bears a strong similarity with reinforcement learning models, the assumption of symmetry makes possible update rules which are much more efficient than reinforcement-learning style "backup" rules like TD(λ) or prioritized sweeping. [1] Another approach is to use reinforcement learning directly. For example...

This is **also plagiarism**. The issue here isn't that you've cited your sources. The issue is that it appears to be *you saying this*, when in fact someone else did. Just because you cite someone doesn't mean you can then insert their text into your own.

Example 3

What if we mention the authors?

*The pheromone updating model is an effective technique for ant foraging. **As Luke and Panait have pointed out**, though this model bears a strong similarity with reinforcement learning models, the assumption of symmetry makes possible update rules which are much more efficient than reinforcement-learning style "backup" rules like TD(λ) or prioritized sweeping. [1] Another approach is to use reinforcement learning directly. For example...*

This is **still plagiarism**. You're no longer stealing the idea from the original authors: but you are still stealing their text. You have acknowledged their credit for their idea, but you have made it sound like you're rephrasing them in your own words when you have in fact just used their words.

Example 4

What if we use less of their text?

The pheromone updating model is an effective technique for ant foraging. As Luke and Panait have pointed out, the assumption of symmetry makes possible update rules which are much more efficient than reinforcement-learning style “backup” rules. [1] Another approach is to use reinforcement learning directly. For example...

Don't do this. If you use a very small number of words (less than a phrase), you're relatively safe from accusation. But in the example above, you are still open to accusation. Again, the issue is that you're using someone else's invented text and pretending it's your own speech.

Example 5

How about lightly rewording their text?

The pheromone updating model is an effective technique for ant foraging. As Luke and Panait have pointed out, when you assume a symmetric model, you can come up with update rules which are more efficient than the “backup” rules common in reinforcement learning methods like prioritized sweeping or TD(λ) [1] Another approach is to use reinforcement learning directly. For example...

If your text is still basically the same structure as the source's text, you're plagiarizing his ideas. A good guide here is: never reword text just “to avoid plagiarizing.” If you're doing that, you're probably still plagiarizing. Instead, write in your own words from the start without relying on the original text.

Example 6

What if I totally rewrite the idea in my own words without consulting the original text?

The pheromone updating model is an effective technique for ant foraging. As Luke and Panait have pointed out [1], this model looks and feels like reinforcement learning, but its update rules are much more efficient. This is made possible by the symmetric nature of the interactions in the model. Another approach is to use reinforcement learning directly. For example...

This is much better. But it's still not as good as the next example:

Example 7

What if I quote the authors?

The pheromone updating model is an effective technique for ant foraging. Luke and Panait [1] have pointed out:

Although this model bears a strong similarity with reinforcement learning models, the assumption of symmetry makes possible update rules which are much more efficient than reinforcement-learning style “backup” rules like TD(λ) or prioritized sweeping.

Another approach is to use reinforcement learning directly. For example...

This is not plagiarism and it’s fine here. But be careful not to quote too large a quote without very good reason, or your reviewers will wonder why you’re relying so heavily on someone else’s text. Additionally, a very large quote without very good cause will lead to charges of **copyright violation**, a legal situation you don’t want to be in.

Example 8

What if I quote the authors inline?

The pheromone updating model is an effective technique for ant foraging. Luke and Panait [1] have pointed out, “Although this model bears a strong similarity with reinforcement learning models, the assumption of symmetry makes possible update rules which are much more efficient than reinforcement-learning style ‘backup’ rules like TD(λ) or prioritized sweeping.” Another approach is to use reinforcement learning directly. For example...

This is okay, though larger quotes should probably be offset as in Example 7.

Example 9

Copying math equations:

The update rule we’re using is $U_p(s') = R(s') + \gamma \max_{a \in A} \sum_{s''} T(s', a, s'') U_p(s'')$

This is plagiarism. Math should be treated like very dense text. That is, a moderate equation should be treated with the same care as several sentences. You need to make it clear that the equation comes from a source.

Example 10

What if I mention and cite the authors:

As Luke and Panait have shown [1], the update rule is $U_p(s') = R(s') + \gamma \max_{a \in A} \sum_{s''} T(s', a, s'') U_p(s'')$

This is dangerously close to plagiarism. It sounds like you're deriving the *idea* from the source, but it's not made clear that you have derived the *text* from the source. Do not do this; it opens you up to accusation.

Example 11

What if I make it clear that this is the authors' equation?

The update rule (from Luke and Panait [1]) is: $U_p(s') = R(s') + \gamma \max_{a \in A} \sum_{s''} T(s', a, s'') U_p(s'')$

This is acceptable. If you need to cite a few equations, just say that: "This and the next four equations are from Luke and Panait [1]".

As it so happens, the notation and form of this particular equation example is very close to the "folk" notation for a standard broadly-used equation called the Bellman Equation. $U(s)$, $R(s)$, γ , and $T(s, a, s')$ are all common notation for this equation. If you are using standard notation common in a large portion of the literature, and a standard equation, you don't have to indicate a source. But you should be *very conservative* when making that judgement. When in doubt, **always cite the source**.

Example 12

How do I use more than a few equations?

*The next several equations are cited, **with permission**, from Luke and Panait [1]. The update equation is: $U_p(s') = R(s') + \gamma \max_{a \in A} \sum_{s''} T(s', a, s'') U_p(s'')$*

You can't cite more than a few equations without very good reason. If you wish to cite *many* equations, then you may be accused of padding your own paper with the ideas of your source. There is a way to do it though: **ask your source for permission**. Note that you may need to get permission *both from the author and from the publisher* if the publisher is the copyright holder, depending on the author's copyright agreement with them. If they give you permission, you may use multiple equations safely if you indicate where the source was as shown above. (But not large pieces of text).

Example 13

Can I use a figure?

Figure 5 shows the state transitions...

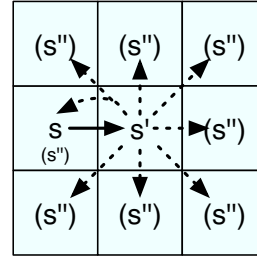


Figure 2: The State Transitions.

This is plagiarism. You have directly copied your source's ideas and art without citation.

Example 14

Okay, what if I cite the authors?

Figure 5 shows the state transitions reproduced Luke and Panait [1]...

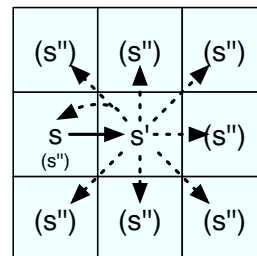


Figure 3: *The State Transitions.* (reproduced from Luke and Panait [1])

This is no longer plagiarism, but it is almost as bad: it is copyright violation. Figures are not like text for two reasons. First, a figure is set off of text, and so it appears to be a free-standing element without citation. But second and more importantly, figures are author's creations, and involve much more effort than mere text. You cannot just cut and paste a figure with citation. **You have to get the source's permission** (and possibly his publisher's permission).

Example 15

Getting permission.

Figure 5 shows the state transitions...

[or stronger]

Figure 5 shows the state transitions reproduced with permission from Luke and Panait [1]...

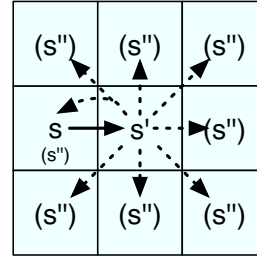


Figure 4: *The State Transitions* (reproduced with permission from Luke and Panait [1]).

You should only use a source's figure with permission, no matter how small it is, and also the publisher's permission if necessary. You must indicate this permission in the caption of the figure. You may wish to also include the permission in the body text as well.

Example 16

Using the caption text from another paper.

Figure 5 shows the state transitions...

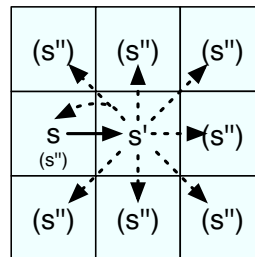


Figure 5: *The State Transitions* (reproduced with permission from Luke and Panait [1]). The ant was just in state s , and has just transitioned to state s' . From there the ant may transition to any state $s'' \in S''$.

This is plagiarism. **Treat figure captions exactly like body text.** The citation ([1]) does not give you permission to reproduce text in such a way that it sounds like your own words.

The next few examples use algorithms from a paper I wrote with Deepankar Sharma and Gabriel Balan called “Finding Interesting Things: Population-based Adaptive Parameter Sweeping” [2].

Example 17

What about a large (third-page or more) algorithm?

From Luke [2], with permission:

```
procedure GRADIENTMAGNITUDESEARCH(initializationSize, exploreProbability, numBrackets,  
                                     fitTourn, distTourn)  
    pop =  $\emptyset$  ▷ Population Initialization  
    for i = 1..initializationSize do  
        ind = CREATE a random individual  
        ASSESS(ind)  
        pop = pop  $\cup$  {ind}  
  
    loop for some time ▷ Parent Selection  
        with probability exploreProbability  
            parent1 = CREATE a random individual  
            ASSESS(parent1)  
            pop = pop  $\cup$  {parent1}  
        otherwise  
            parent1 = a random individual from pop  
            parent2 = SELECT(parent1, pop, fitTourn, distTourn)  
  
        for i = 1..numBrackets do ▷ Iterated Bracketing  
            child = CROSSOVER(parent1, parent2)  
            ASSESS(child)  
            pop = pop  $\cup$  {child}  
            parent1 = UNLIKEPARENT(parent1, parent2, child)  
            parent2 = child  
  
    return pop
```

A large algorithm is a significant effort of art and so should be treated like a figure. You need to get permission from the author and from the publisher.

Furthermore, because an algorithm is so large, it will look like you are padding your paper and reviewers don't like this. You should think twice about including a large algorithm from someone else.

Example 18

What about a small algorithm?

From Luke [2]:

```
procedure SELECT(parent1, pop, fitTourn, distTourn)
  best = SELECT2(parent1, pop, distTourn)
  for i = 2...fitTourn do
    new = SELECT2(parent1, pop, distTourn)
    if |Fitness(new) – Fitness(parent1)| >
      |Fitness(best) – Fitness(parent1)| then
      best = new
  return best
```

A sufficiently small algorithm may be treated similar to a large equation. You may include it if you make it clear that the algorithm text is not yours. The example above is okay, but it'd be better if you wrote this in your own words. It'd still be wise, but (if the algorithm is very small) not necessary to get permission from the author.

References

- [1] Liviu Panait and Sean Luke. 2004. A pheromone-based utility model for collaborative foraging. In *3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2004)*. IEEE Computer Society. 36–43.
- [2] Sean Luke and Deepankar Sharma and Gabriel Catalin Balan. 2007. Finding interesting things: population-based adaptive parameter sweeping. In *GECCO 2007*. ACM. 86–93.