Sungmin Rhee
srhee4@uw.edu
CSE 403
Project 1: Pitch document

Tree merge for version control

More often than not, coding projects require a team of engineers to work on it simultaneously. Even with a good version control system and a good practice, a merge conflict is almost inevitable. Many of current version control systems are using line-based methods to analyze updates. This means that the system will detect any difference in a line of code, and consider it to be a change. This is not always correct in the user's perspective. If we wrapped a block of code with an if statement, we are really making changes only at the boundary of that block of code by adding that if statement. However, the current system will consider this as a change in the entire block of code simply because they were indented. This would cause a merge conflict if another member changed a line within that block of code. Merge conflicts will require manual resolution, which will add more workload to the user.

One way to resolve this problem is to analyze updates using a tree structure. If we convert the code into a syntax tree, it is easier to see that the previously discussed conflict is not a real conflict. Diagrams below show how the version control system would perceive the same changes in the code when using line-based method and tree-based method. The diagram on the left shows the case of using line-based method. Because adding the if statement in update 1 indented the line 'x = 2', the system thinks that it conflicts with update 2 which changes the line 'x = 2' to 'x = 3'. However, on the right, it is more clear that update 1 is not actually changing the line 'x = 2' thus there is no conflict with update 2.

One challenge would be to manage complexity. Below examples are simple ones, but if we were to use this analysis on an entire codebase, the tree would grow very large and complex. This might make it more difficult to figure out which part in update 1 corresponds to which part in the original code or update 2, making the comparison much harder.