GOFIRST Source Code Conventions and Style Guide

Max Veit
GOFIRST Senior Software Engineer

Last Updated: December 29, 2013

0. Contents

1. Introduction

1.1 Purpose and Audience

The purpose of this document is to establish a uniform style of writing source code for all programs written for GOFIRST projects, mainly the IGVC competition. This is intended to facilitate collaboration between programmers as well as maintainability of software, since a uniform style removes one of the main barriers to understanding the source code written by others (or yourself!).

This document is aimed at anybody interested in writing programs that will become a part of or aid GOFIRST projects. It assumes only a basic, general knowledge of programming, including the existence of different programming languages, the purpose of a source code file, and the existence of a language-specific syntax for the text contained therein. However, the style guides in Section B do assume a basic familiarity with the specific syntax of the language being described. To better understand these guides it might be helpful to find a simple introductory tutorial in the language and its syntactic constructs (such as function definitions, "if," and "for" constructs).

1.2 Typographic Conventions

Source code, variable names, command lines, and anything else that is only correct exactly as it appears (including spaces) will be shown in a `monospaced font`.

Definitions, special names, or other words to remember are set in a **bold** font (in text, as are section headings separate from the body text).

Commands you type at the operating system command line interface (ordinary terminal in Unix or GNU/Linux, Git Bash or Cygwin prompt in Windows) will be prefixed with a `$` symbol. You should not actually type the `$` or the space that follows it. It is only shown in this document to remind you that the text after it is a command (you should also see something like it in the actual terminal; the symbol is called the **prompt**). When you actually type the command, enter everything from the first word after the `$` onwards (including spaces, although not the one before the prompt).

If there is something in one of these command lines or program snippets you need to fill in yourself it will be indicated with *italics*. For example,

```
$ git config --global user.name "Your Name"
```

means you should replace *Your Name* with your actual name (but leave the quotes as they are; those aren't emphasized).

Command lines that are too long to fit on this page are continued on the immediately following line and without a leading `$`, like so:

```
    $ foo bar baz --quux --blah File Name Here
--hippopotomonstrosesquipedaliophobia
```

To type such a command line into the terminal, ignore the line break (replace it with a single space) or use your terminal's line-continuation character if you know what that is.

Longer examples of program code are typeset in a frame approximately 80 charaters wide, to mimic the actual appearance of code in a text editor. Omissions are indicated by an ellipsis (`...`) on its own line. For example:

```
VisionPacket VisionInterface:: getPacket() {
    ...
    // Exclude the sky
    rectangle(lines, Point(0, 0), Point(obsts.cols, params.hrzHeight),
              Scalar(0), CV_FILLED);

    // Find the contours and send the packet on its way
    vector<Mat> contours;
    VisionAlg:: findLineContours(lines, contours, lineOpenKrn);
    ...
    //// This comment was cleverly devised to be exactly eighty characters long.
```

The frame is the same as that used for section headings; however, the different font styles in each type of frame should be sufficient to distinguish the two types of constructs.

2. Language Classes

This section lists the classes of languages that will be discussed in this document.

C-syntax (C, C++, Java and C#)
: This is a family of languages with differing programming paradigms. The one thing they have in common is the basic syntax derived from C, so they will generally be discussed as a group in this document. Most GOFIRST code so far uses the C++ and C# languages in this family.

Python
: This language blurs the boundaries between programming and scripting; its rapid development cycle and interpreted nature make it ideal for high-level control of large, composite programs, an area known as the "glue layer." Its syntax is much simpler and more consistent than that of the C-syntax languages. We are considering using this language in the 2014 competition code for the glue layer and more algorithmically focused sections of the program.

LabVIEW
: The language designed for use with National Instruments (NI) hardware. In the past, GOFIRST has used NI controllers and boards, necessitating the use of this langauge. NI's controllers and language are no longer used in the IGVC competition robots, so this document does not discuss the language in detail. What distinguishes LabVIEW from most other languages is that it is a graphical language; programs are created not by writing statements to a source code file but rather by visually connecting virtual blocks with wires. This graphical nature makes it difficult to set specific style guidelines for LabVIEW code. Some conventions are already imposed upon the programmer, but the visual layout of the blocks and wires themselves is completely up to the programmer. "Spaghetti code" that looks like a literal pile of spaghetti can quickly ensue if you are not careful to organize the blocks and wires in a logical, readable fashion. A general rule: If you cannot understand the code that you wrote yesterday, you need to put more effort into organizing your code layout. Of course, this rule applies to text-based languages as well!

3. Source Code Style and Conventions

This section focuses on the detailed mechanics of the style of GOFIRST source code. The term **style** as it is used here means the positioning, layout, whitespace, and naming aspects of source code not directly specified by the program's logical structure. Source code style is meant to facilitate the reading, understanding, and modification of a source code file. A source file can compile to a completely correct, consistent program and still be difficult or impossible to read. This is why programmers generally hold themselves to some sort of style guide, whether explicity written or not.

So as not to reinvent the wheel, this document references some already established and well-known standards; they are listed in Appendix B at the end of this document.

This section to be expanded soon.

3.1 General

The general GOFIRST code style bases on two style manuals: The Java style guide for C-syntax languages and the Python style guide for Python; both are listed in Appendix B.

For Python, the entire style guide applies. Please read it in its entirety if you plan on working with Python at all.

For C-syntax languages, please read the Java style guide, sections 4-9. Also, observe section 3 insofar as it applies to your language, and take section 10 with a grain of salt. It contains good advice, but not all of it is applicable to or compatible with languages besides Java. Finally, I have one important disagreement with Section 4.2 of the guide: Lines should be broken *after* binary operators, not before. This is the most important conflict between GOFIRST style and the Java style guide, so it's worth mentioning it here even though it's covered in the points below.

For both language classes, the specific recommendations in this section *override* any conflicting recommendations in the applicable style guide. To repeat, if there is any conflict between the points below and one of the two online style guides, this guide wins.

3.2 C/C++ Features

The languages C and C++ contain syntactic elements that are not present in Java, such as pointers and (for C++) templates. Also, the syntax for class definitions (especially when inheritance is involved) is somewhat different. This subsection sets conventions for using those features.

The C++ language also has the peculiarity that two styles of class declaration are possible: One with the declarations in a header file and implementations in a C++ source file of the same name, and one with the entire class declaration and definition **inline** in a single C++ source file. The former is preferred, as it is impossible to `#include` C++ source files in other files.

Another C++ peculiarity is that multiple extension types are possible for source files. Use `.cpp` for C++ source code (implementation) files, `.h` for C or C++ header files. If the project mixes C and C++ code (not recommended unless it's absolutely necessary), you may use `.hpp` for header files to indicate that they cannot be parsed by a plain C compiler.

The possibility of multiple inclusion of C++ header files in even moderately-sized projects can often cause problems with multiple definitions. To avoid this, use **header guards** to prevent the file from being included more than once. They look like this:

```
// Beginning of file
// Include statements, ''using'' namespace-related statements
...
#ifndef HEADERFILE_H_
#define HEADERFILE_H_
// Declarations for functions, variables, and classes
...
#endif
// End of file
```

The name of the defined constant should be related to the name of the header file in the following way: Put the name of the file in all caps, replace the period with an underscore, and add a trailing underscore. So the file `BufferThreadedP.h` should have a line like `#ifndef BUFFERTHREADEDP_H_` in it.

3.3 Comments

Comments serve several different purposes in source code. A good enumeration of the various types can be found in Steve McConnell's "Code Complete" (the standard

textbook for the U of MN introductory course in program design and development, CSCI 3081W). The Python style guide, section “Comments” (ignore the parts specific to Python syntax) provides a more concise overview; please read it. In short: Comments should *not* repeat the code. Comments should *never* contradict the code. Comments should not attempt to explain confusing code; please use tools such as variable names and defined constants to make your intent more clear. Comments are often helpful when used as headers for distinct logical sections of code. **Summary comments** and **documentation comments** (block comments, function comments) are more useful than **inline comments**.

Every function and *every* class needs a **documentation comment** explaining what it does and how to use it, including a short description of each of the parameters and of the return type (if applicable). Such comments should be in a format that Doxygen (ref needed) can parse. Please use the convention of comments starting with `/**`. An example of a well-formed Doxygen comment in C++ is:

```
/**
 * Find contours in the input binary image; first applying a morphological
 * open operator to remove small artifacts resulting from the detection
 * procedure.
 *
 * \param img Input binary image of detected lines, possibly including thin
 *        borders and other artifacts. This array will be modified by the
 *        morphological transformation.
 * \param contours Place to store the resulting contours
 * \param openKrn Morphological kernel for removing borders and artifacts
 *
 * \return The number of contours found.
 */
int VisionAlg:: findLineContours(
        InputOutputArray lines,
        OutputArrayOfArrays contours, InputArray openKrn) {
    // Use morphology to get rid of the small borders
    morphologyEx(lines, lines, MORPH_OPEN, openKrn);
    ...
```

The type and purpose of the returned data does not need to be explained for constructors, or if the function does not return anything (or returns `void`).

3.4 Line Length

The *maximum* number of characters allowed in one line of source code is 79. Don’t go over it, even by one character. Doing so will annoy every other person who tries to read the code with an 80-character terminal, as the resulting line wraps

visually disrupt and distort the text. Every decent language has a feature for long-line continuation; please use it. Also, please break your lines *after* binary operators such as && or =.

The only exception here is the source code for documentation to be typeset with TEX/ LATEX(like this document). Since such files contain long blocks of prose, it can often be desirable to let the editor handle the line wrapping.

3.5 Whitespace

The issue of whitespace is often so contentious that it is a good idea to set detailed and consistent rules for its use throughout source code.

Please read the Python style guide, section "Whitespace in Expressions and Statements". For C-syntax code, hold yourself to these rules where they are applicable to your language. For Python code, you should observe the entire style guide anyway.

3.5.1 Spaces vs. Tabs

For indentation, use four spaces and no tabs. To repeat, your source files should *not* contain the Tab character. Often, stupid "holy wars" flare up about what style of indentation is preferable or even superior; setting a convention throughout the project and sticking with it is the best way to avoid needless disputes, even if it inconveniences some developers' workflows. The spaces vs. tabs issue mostly comes down to personal preference and there is not much to be gained in a technical sense by choosing one convention over the other. I chose spaces because of a minor technical reason relating to the continuation of long lines of code (although you could probably make a convincing technical argument either way): Clearly you want to align the continuation in some logical way with the preceding line, such as a continued argument list of a function that you want to align with the first argument in the previous line. Plain tabs are often a recipe for disaster (different editors sometimes throw off the indentation wildly), and mixing spaces and tabs for indentation in one file is generally frowned upon.

So please, for the sake of consistency, set your editor to expand tabs to spaces up to the next column that is a multiple of four (most decent editors should be able to do this). Those who claim to have a real problem with this in terms of readability will have to deal with it, i.e. either get used to it or take the time to program/script a solution.

3.5.2 Line Continuation

Our limit on line length necessitates a consistent convention for the whitespace mechanics of line continuation. The Python style guide, section "Code lay-out" subsection "Indentation" suggests a convention for continuing long function calls or multline constructs. The hanging-indent convention, a variant on the "Pretty-Printing" convention from Lisp-like languages, is preferred. For the hanging indent, please use one indent (4 extra spaces) in addition to the indent level that a normal line in that position would have. For example, if you are writing a long function declaration (the first logical line of the definition, which specifies the function name, arguments, and sometimes the return type) that wraps onto the next line (or lines), indent the continuation *two* levels (8 spaces) farther in than the first line of the declaration, like so:

```
// Function declaration
VisionInterface:: VisionInterface(
        Matx33f cameraMat, Mat distCoeffs, Matx33f perspTrans,
        std::string camAddress) :
        cameraMat(cameraMat), distCoeffs(distCoeffs), perspTrans(perspTrans) {
    // Start of function body
    if (!source.open(camAddress)) {
    ...
```

The first line following a function declaration is normally indented in one level relative to the line above it; the extra indent is necessary to distinguish the continuation of the declaration line from the first line of the function body that would normally occupy that place.

3.5.3 Blank Lines

No more than necessary, but please do use them to separate logical "paragraphs" of code (e.g. 5-10 lines inside a function that perform a distinct task). Do not use blank lines directly after a function declaration line.

3.5.4 Miscellaneous Notes

Avoid whitespace at the end of lines. This makes it harder to copy and paste entire lines in some editors. If your editor has a feature to make trailing spaces visible, turn it on.

4. Technical Notes

The notes below don't necessarily relate to the layout and style of code; however, they are still important for compatibility and standardization purposes.

4.1 File Encoding

Please use **UTF-8** for general symbol support and inter-system compatibility. No, I don't anticipate any need to write comments in Russian, Finnish, or Japanese; nevertheless, it is the modern, extensible way to encode text in a file. Also, the encoding happens to coincide with ASCII over the entire set of 128 pure ASCII characters, so you have no excuse.

Using Unicode strings in programming storage and communications is a little trickier, as some languages' support for the feature is still lagging behind. Please use Unicode for all strings intented for communication with other programs, e.g. for strings sent via the network or via message queues. For communication with external hard-programmed hardware devices (e.g. an LCD status display), please use whatever encoding the device accepts by default.

4.2 Line Endings

Line-endings usually aren't a problem if your editor is not Notepad (if it is, shame on you; go see Appendix A). However, in order to preempt any possible problems or inconsistencies, please use the automation facilities available to you. The Git version control system includes transparent handling for line ending formatting. If you are working on Windows, make sure you set Git's `core.autocrlf` setting to `true` (You can do this by typing

```
$ git config --global core.autocrlf true
```

into the Git Bash prompt).

4.3 Git setup

Please set up your Git settings so that your commits are tagged with your name and email. This is useful if we want to know who made a certain change e.g. to find out what it was intended to do. You can set this up by typing:

```
$ git config --global user.name "Your Name"

$ git config --global user.email
your.address@emaildomain.tld
```

into the Unix or Git Bash command line.

4.4 General Rules

We have programmers using different systems (primarily *nix, but occasionally Mac and Windows), so if you have a choice between convenience and compatibility you should almost always choose compatibility.

5. Miscellaneous Recommendations

This is a miscellaneous collection of recommendations concerning programming in general, and not any specific language.

- Version control: Just use it. Learn a modern version control system (in GOFIRST we use Git) and use it for every software project on which you ever work. These systems give you a complete history of the revisions your software has seen (as long as you commit often enough!), allowing the recovery of lost work, faster fixing of bugs, and in general saving you hours and hours of development time in every project, big or small.
- "Magic Numbers:" It's generally considered bad practice to embed numbers other than 0, 1, 2, or -1 directly into a program's control flow, rather than as constants at the global, file, or class level. Other numbers (integers!) that are tied closely with the structure of the surrounding code block (for instance, as an increment for a step of a specific algorithm) are exempt from this rule *if* this use is documented with a comment briefly explaining the significance of the value. Just writing a statement like `increment = sensorLevel * 67`

(example in Java) will leave anyone else who reads the code scratching their head trying to figure out just what on Earth the value 67 even *means* in this context. Even worse, if the value 67 is used at multiple points throughout the code, and tests determine that a value of 72 gives more accurate results, replacing every single occurrence of the value would be a pain. And no, find-and-replace is often not an option if the example were using a more common value, such as 4, which you really wouldn't want to replace throughout your project. It would be much better to first declare 67 as a constant at the top of the class: `public static final int sensorScaling = 67`, then using `increment = sensorLevel * sensorScaling`. This way, the semantics of the sensor scaling value are made a bit more clear (and can be made even more clear by a comment next to the constant). Also, changing the scaling to 72 is nearly trivial, and tuning this value experimentally is much easier than it would be in the previous example.

- Don't rely on letter case to distinguish between two unique identifiers. For example, if you have already defined a variable called `httpConn`, don't make another variable called `HTTPconn`. Many languages regard these two identifiers as distinct, but you shouldn't. For one, it's hard to remember exactly what case you used for what variable. It's also not portable to case-insensitive languages or situations where the case is stripped from the name (as in the case of constant names for header guards).

- Just because a certain code construct or statement is redundant in a specific case does not mean you should leave it out, even if the two forms are equivalent *after* compilation. In some cases, leaving out such redundant statements can greatly hinder source code readability. For a simple example, take variable initialization. Java automatically initializes fields (variables belonging to a class) to a default value of 0, `false`, or `null`, depending on the type. Do *not* rely on other people (or yourself, some time later) to recognize this fact immediately! Another programmer reading the code may be confused by the fact that a variable is used apparently without being initialized, an error in many other languages. Worse still, if this hypothetical programmer doesn't see the code that relies on this implicit assignment, they may change the variable's value elsewhere, creating an incredibly elusive bug. In addition, programmers with relatively little Java experience may not always remember the exact initial assigned value for a field. With integers and objects this is usually easy to figure out, but the boolean value `false` is not at all intuitively obvious. **Note:** This does not apply to redundant statements that actually worsen the performance of the code, as opposed to those that are automatically compiled away to a more efficient form.

A. Editor and IDE recommendations

This section lists some text editors and Integrated Development Environments (IDEs) that members of GOFIRST have found to be useful and productive for programming diverse languages on different platforms.

A.1 Text editors by platform

Windows
: As mentioned above, you want to move away from Notepad. It's not even good for editing configuration files. One alternative is Notepad++ (http://notepad-plus-plus.org/).

Mac
: The built-in text editor is usable, but unsuited to programming. Luckily, quite a few cross-platform editors integrate well with OSX. You could also try TextWrangler (http://www.barebones.com/products/TextWrangler/), although I have personally never used it.

GNU/Linux
: The built-in editors GEdit (GNOME) and Kate (KDE) are both good, extensible, full-featured, and usable for programming in many languages. However, it might be worth it to take a look at some of the cross-platform editors as well.

Cross-platform
: There are two classic editors for Unix-like systems that have been extensively ported:

  - Vim, the "improved version" of Vi (http://www.vim.org). Max's editor of choice; many people claim it's by far the most efficient editor they've ever used. However, Vim is notorious for its learning curve. If you have a free weekend and feel like learning to use the editor efficiently, check out the program "vimtutor" which should be installed along with the editor (try typing `:help tutor` after starting Vim). Features include everything you'd expect from a programmer's text editor (syntax highlighting, search/replace with regexes, code folding, and diff to name a few). Built-in support for most known programming languages, often with extensions available for specific languages. Originally designed for use in

a terminal, although good graphical versions are available on the download page.

- (x)Emacs (http://www.xemacs.org), perhaps the more full-featured (according to some, bloated) of the two classic editors. Some people (such as Max; now you know where he stands in the "editor wars") find it unusable, although if you're a boss with chained keyboard shortcuts in Word or whatever you might like this editor. Lots of programmer's features and extensions available. It even has a built-in psychologist (try `M-x doctor` or something like that). The graphical version is xEmacs.

In addition, jEdit (http://www.jedit.org/) is a full-featured, cross-platform text editor with support for many programming languages. It appears to to work well for editing Java code.

A.2 IDEs by language

Java

- Eclipse (http://www.eclipse.org/) is a popular cross-platform IDE for Java and C++ development. It has all the features you'd expect from a standard IDE, plus a (somewhat clumsy) plugin installation system. It can integrate with version control systems such as Git (which we will most likely be using) through the plugin framework. It's a large piece of software – it's several hundred megabytes and slow to start up on any system. However, many Java (and C++) programmers think the rich feature set, which includes integration with a sort of task-tracking software, is worth it.
- NetBeans (http://netbeans.org/) is definitely more lightweight than Eclipse but specializes on Java. It also has a plugin system with a large database of user-contributed plugins (Git support is naturally among them). Some programmers might find the relatively minimalistic approach of NetBeans more appealing than the one-stop-shop philosophy of Eclipse.

C/C++

- As mentioned above, Eclipse also works well with C++.
- If you're on a Mac, you could check out XCode (it costs $ 5 on the Mac App Store). However, it's frustrating to use for development not involving a shiny Mac application.

- On Windows, Visual Studio is the standard all-in-one IDE available for CSE students for free at http://cselabs.umn.edu/software/msdnaa. Although designed primarily for Windows development, it's usable for general C/C++ program development.

C# (If we do end up using it):

- Those with Windows should use the Microsoft Visual Studio IDE.
- Those who don't have Windows (or don't want to install it) can try the Mono development environment, see http://www.monodevelop.org/.

Python

Again, we might not end up using Python at all, but if we do, here are my IDE recommendations: First, it's perfectly possible to develop Python with nothing but a text editor and terminal (that's what I do, and I prefer it to using an IDE). However, if you're used to using IDEs and want to use one for Python, there is one available: IDLE should be installed with the default Python installation (Linux users might need to install an additional package). It's about as minimalistic as IDEs get; nevetheless, it does simplify and integrate the Python development process. Users new to Python or lacking a decent command-line interface might find this helpful.

General

If you like lightweight IDEs, try Geany (http://www.geany.org/). It's not tied to a specific language, which is both an advantage and a disadvantage depending on your personal preferences. In any case, it's got pretty much everything you need to edit and build most types of source code effectively.

Others

Search the web and use your best judgement :)

B. Bibliography

This is a list (with commentary) of the documents and Web resources referred to throughout the body of this document. Also listed here are a few lighthearted presentations of general tips on programming.

- Sun Code Conventons for Java: http://www.oracle.com/technetwork/java/codeconventions-150003.pdf

This is a good, but not perfect, introduction to basic style for Java source code. It provides a guide for the general “cosmetics” of source code files in C-like languages so programmers don’t have to fight with something as trivial and infuriating as brace placement. We will be using it with a few modifications (detailed below) for all of our Java, C and C++ code because of the syntactic similarities between the languages. Please read this thoroughly if you have no experience with C syntax (come on, it’s only 18 pages. A TL;DR response will disqualify you from our programming team). If you have programmed in the C-syntax languages before, at least skim the guide to see what specific conventions well be using.

- Style Guide for Python Code: http://www.python.org/dev/peps/pep-0008/
  A very good guide for Python code with general advice that applies to other programming languages as well. Please at least read the introduction, first section, and section on comments. If we do any coding in Python, the style should follow this guide. Comments in other languages should follow the advice in this guide; for all other areas the language’s style guide naturally has precedence.
- Pasta Code: http://www.gnu.org/fun/jokes/pasta.code.html
  If you’re doing object-oriented programming, please try for ravioli code. Canderli code and Ristto code should be avoided because of poor maintainability. Polenta code is nice, too, if it never needs to be modified (which is generally never the case, so watch out!)
- Good Code: http://xkcd.com/844/
  Please try to stay in the "Code Well" loop. Yes, requirements do change, and we’ve had to throw out a few projects because of this, but trust me: this is not always the case!