

# Cluster Computing for Humans White Paper: ECJ and the Royal Tree Problem

Digital Scholarship Lab  
Brock University  
St. Catharines, Canada  
dsl@brocku.ca

**Abstract**—This paper is meant to serve as a proof of concept and example use-case for deploying ECJ programs to Brock University's Library mini-cluster, and why it is useful to use a mini-cluster versus running the problem on a desktop PC. The problem tested was the Royal Tree Problem, which is a useful benchmark program for genetic programming.

## I. INTRODUCTION

The department of Computer Science at Brock University uses Java Evolutionary Computation Toolkit(ECJ) for it's fourth year and graduate level genetic programming classes (COSC 4P82/5P71), and consequently is also pinnacle to a lot of graduate level research that goes on at Brock within the Computer Science department.

One problem that consistently can come up with using toolkits like this is the length of time it takes to run operations, for example, if one needs to run an operation with 10 different parameters (to compare results of the different parameters for example) at least 5 times, at an operating time of 10mins, on a single computer it could take up to 500 minutes. This is a constant and real problem that occurs within the department.

On a mini-cluster, we can take that time and slash it drastically by spreading out each parameter set to a node(Workstation in the cluster). This allows a variety of parameter sets to run synchronously.

Solving this problem, allows us to use the library's available resources more effectively, e.g. instead of having computers sit idly and not be used, we can constantly be computing and thus getting more of out the existing machines. Solving this problem was a first step towards the ability to solve large real-world problems.

## II. METHODS AND PROCESS

Utilizing HPC Pack 2016, we can achieve a mini cluster capable of executing multiple genetic algorithm instances with different parameters.

The library cluster consists of NUMof on premise library computers which a head node dispatches jobs too. Each workstation node is non-dedicated to the cluster, and can be scheduled to become available during specific times outside of library hours.

### A. What is ECJ

ECJ is an Evolutionary research computation system that is written in Java. The Framework supports a vast quantity of computational techniques and processes such as evolution strategies, genetic programming, genetic algorithms, co-evolution, and particle swarm optimization. The Framework uses a series of Pipelines to represent an iterative evolution for individuals with a selection process. It also provides operators for mutation and crossover techniques that represents breeding of another individual in a population.

### B. Setting Up Jobs

ECJ execution source files were placed into the network storage drive the head node has access to. In our sample we have nine separate instances of ECJ algorithm running with different parameters set. Thus nine separate parameter files were created and nine .bat files to execute each parameter set.

Each of the .bat execution files are run through the HPC cluster on separate nodes dispatched by the head node.

### C. Gathering Results

Once several Parameter files were created the program was run producing output files, Each output file contains some ECJ output text followed by a print out of each Generations best fitness score. Once the final generation was run an overall best generation and run fitness score was printed.

The output text files contained useful data, but the file format with other filler text made processing the data difficult. Therefore some parsing and cleaning of the data was needed. Python was used to loop through each file in our folder producing a CSV file containing only the generation alongside the best fitness score. This allowed for the use of clean tabular data for the data processing and statistic gathering.

## III. PREPARE YOUR PAPER BEFORE STYLING

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections III-A–III-E below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads— $\LaTeX$  will do that for you.

### A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

### B. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: “Wb/m<sup>2</sup>” or “webers per square meter”, not “webers/m<sup>2</sup>”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
- Use a zero before decimal points: “0.25”, not “.25”. Use “cm<sup>3</sup>”, not “cc”.)

### C. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus ( / ), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \quad (1)$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use “(1)”, not “Eq. (1)” or “equation (1)”, except at the beginning of a sentence: “Equation (1) is . . .”

### D. $\LaTeX$ -Specific Advice

Please use “soft” (e.g., `\eqref{Eq}`) cross references instead of “hard” references (e.g., (1)). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

Please don’t use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Please note that the `{subequations}` environment in  $\LaTeX$  will increment the main equation counter even when there are no equation numbers displayed. If you forget that, you might write an article in which the equation numbers skip from (17) to (20), causing the copy editors to wonder if you’ve discovered a new method of counting.

$\BibTeX$  does not work by magic. It doesn’t get the bibliographic data from thin air but from .bib files. If you use  $\BibTeX$  to produce a bibliography you must send the .bib files.

$\LaTeX$  can’t read your mind. If you assign the same label to a subsubsection and a table, you might find that Table I has been cross referenced as Table IV-B3.

$\LaTeX$  does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it’s supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` (there won’t be any anyway) and it might stop a wanted equation number in the surrounding equation.

### E. Some Common Mistakes

- The word “data” is plural, not singular.
- The subscript for the permeability of vacuum  $\mu_0$ , and other common scientific constants, is zero with subscript formatting, not a lowercase letter “o”.
- In American English, commas, semicolons, periods, question and exclamation marks are located within quotation marks only when a complete thought or name is cited, such as a title or full quotation. When quotation marks are used, instead of a bold or italic typeface, to highlight a word or phrase, punctuation should appear outside of the quotation marks. A parenthetical phrase or statement at the end of a sentence is punctuated outside of the closing parenthesis (like this). (A parenthetical sentence is punctuated within the parentheses.)
- A graph within a graph is an “inset”, not an “insert”. The word alternatively is preferred to the word “alternately” (unless you really mean something that alternates).
- Do not use the word “essentially” to mean “approximately” or “effectively”.
- In your paper title, if the words “that uses” can accurately replace the word “using”, capitalize the “u”; if not, keep using lower-cased.
- Be aware of the different meanings of the homophones “affect” and “effect”, “complement” and “compliment”, “discreet” and “discrete”, “principal” and “principle”.
- Do not confuse “imply” and “infer”.
- The prefix “non” is not a word; it should be joined to the word it modifies, usually without a hyphen.
- There is no period after the “et” in the Latin abbreviation “et al.”.
- The abbreviation “i.e.” means “that is”, and the abbreviation “e.g.” means “for example”.

An excellent style manual for science writers is [?].

### F. Authors and Affiliations

**The class file is designed for, but not limited to, six authors.** A minimum of one author is required for all conference articles. Author names should be listed starting from left

to right and then moving down to the next line. This is the author sequence that will be used in future citations and by indexing services. Names should not be listed in columns nor group by affiliation. Please keep your affiliations as succinct as possible (for example, do not differentiate among departments of the same organization).

#### G. Identify the Headings

Headings, or heads, are organizational devices that guide the reader through your paper. There are two types: component heads and text heads.

Component heads identify the different components of your paper and are not topically subordinate to each other. Examples include Acknowledgments and References and, for these, the correct style to use is “Heading 5”. Use “figure caption” for your Figure captions, and “table head” for your table title. Run-in heads, such as “Abstract”, will require you to apply a style (in this case, italic) in addition to the style provided by the drop down menu to differentiate the head from the text.

Text heads organize the topics on a relational, hierarchical basis. For example, the paper title is the primary text head because all subsequent material relates and elaborates on this one topic. If there are two or more sub-topics, the next level head (uppercase Roman numerals) should be used and, conversely, if there are not at least two sub-topics, then no subheads should be introduced.

#### H. Figures and Tables

*a) Positioning Figures and Tables:* Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation “Fig. 1”, even at the beginning of a sentence.

TABLE I  
TABLE TYPE STYLES

Table Head	Table Column Head		
	<i>Table column subhead</i>	<i>Subhead</i>	<i>Subhead</i>
copy	More table copy <sup>a</sup>		

<sup>a</sup>Sample of a Table footnote.

Fig. 1. Example of a figure caption.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity “Magnetization”, or “Magnetization, M”, not just “M”. If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write “Magnetization (A/m)” or “Magnetization {A[m(1)]}”, not just “A/m”. Do not label axes with a ratio of quantities and units. For example, write “Temperature (K)”, not “Temperature/K”.

#### ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

#### REFERENCES

Please number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [?]. Refer simply to the reference number, as in [?]<sup>1</sup>—do not use “Ref. [?]” or “reference [?]” except at the beginning of a sentence: “Reference [?] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [?]. Papers that have been accepted for publication should be cited as “in press” [?]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [?].

#### REFERENCES

- [1] M. Shell. (2007) IEEEtran homepage. [Online]. Available: <http://www.michaelshell.org/tex/ieeetran/>

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove the template text from your paper may result in your paper not being published.