tree_converge a program implementing the Successive-Approximations approach to parameter optimization in maximum-likelihood phylogeny estimation
(Program Manual)

By: Zaid Abdo

For more detailed description of this approach please refer to:

Sulivan, J., Z. Abdo, P. Joyce, and D. Swofford. 2005. **Evaluating the Performance of a Successive-Approximatioon Approach to Parameter Optimization in Maximum-Likelihood Phylogeny Estimation.** *Mol. Biol. Evol.* 22(6) 1386-1392.

This is a command-line PERL program that implements the successive-approximations approach presented in [1] where you will find a detailed description and evaluation of it.

Pre-requirements:

1) Need to have PAUP* [1] installed. PAUP can be found at the following website: http://paup.csit.fsu.edu/
2) Need to have PERL installed. PERL can be found at the following web site: http://www.perl.com/

Again this is a command line program and not a menu driven one. This program exists in two different formats. The first format requires the exact equality (both topology and branch lengths) of the trees resulting from two successive searches as a stopping rule (this program is referred to as "tree_converge.pl"). The other requires the two successive trees to have the same topology, regardless of the branch lengths (it uses symmetric distance to compare and is referred to as "tree_converge_sumDist.pl"). The second program requires downloading our tools library "library.tgz" along with the program because it uses some of the functions existing in that library.

To use "tree_converge.pl" you need to follow the following steps:

1) Create a nexus file call it data.nex (see the example file provided).
2) Choose an appropriate model of evolution to analyze this data set (you can use DT_ModSel to perform that task, for example.)
3) Once the appropriate model is identified create a file called "Model_Command_Initial_Tree." The main purpose of this file is to provide the starting tree for the successive-approximation. An example using the GTR+I+G model is provided in the example "Model_Command_Initial_Tree" file found on this website.
4) Create a second file call it "Model_Command_Iter_Trees." This file includes the code that tree_converge will use, along with the starting tree (that we refer to as tree.tre) to iterate the search process until convergence is attained. An exemplary file is also available on this website with the GTR+I+G as the model of choice.

5) Make sure that the "data.nex", "Model_Command_Initial_Tree", "Model_Command_Iter_Trees" and the "tree_converge.pl" files are all in the same directory.
6) To run "tree_converge.pl" type:
   perl tree_converge.pl
7) The output is made up of three files: a "score_file" showing the parameter estimates associated with the final run; a "tree.tre" file that includes the optimum tree associated with the optimized parameters; and a log file showing the details of the paup run.

To use "tree_converge_symDist.pl" you need to follow the following steps:

1) Down load and unzip the library.gtz file. On windows use WinZip, on Lynix/Unix (including Apple) use tar –xzfv library.gtz.
2) Follow steps 1-5 above.
3) Before running "tree_converge_symDist.pl" make sure that the program can find the library: On windows you will want to change the directory in the statement "use lib "$ENV{HOME}/DTProgLibrary";" the correct directory that you saved the library under (for example, C:\DTProgLibrary). In a Lynix/Unix environment just make sure that you save the library in a directory called "/DTProgLibrary" on your account (and not within any other directory).
4) The output of this program is as described in step 7 above.

If you have any question or input, or if you find a bug in the program please send me an e-mail at:

zabdo@uidaho.edu