The K-Clustering Project

FAQ

| | |
|---|---|
| **Working enviorement** | |
| Can I download cplex to my personal computer? | The formal answer is no. Cplex is very expensive and the only way to use is to connect to the linux servers. However, at your on risk, you can try the trial version of IBM lhttp://www.ibm.com/developerworks/university/software/get_software.htm But we do not know to support the installation and usage. |
| Eclipse does not recognize cplex.h | Option 1. Try opening a new C project of the type makefile project (When creating a new project, choose C project, click next, and then choose the last option makefile project). A project contacting a makefile file would be created. Delete the content of the makefile provided, and put the content of the makefile we provided for you.<br>Option 2. Stand on the project name, and right click to choose properties. In the window opening choose the option c/c++ build. On the tab Builder Settings unchoose both bullets. |
| Working from home | http://www.cs.tau.ac.il/system/faq/services/servers/NX_CLIENT Fill yourself like you are in Schreiber lab while you are at home |
| | |
| **Fixing the project specifications document** | |
| In the results file you should print the network for the U clustering | The instruction document was fixed |
| What files to put in the bonus folder | 3 files: *k_clustering_solution.xgmml*, *best_clusters.xgmml* and *results* fot the network you estimate as real and for k=10 |
| How to I determine what is a "big" cluster | "big" refers to the number of vertices in the cluster. 5 biggest clusters are the clusters with the highest numbers or vertices in them. |
| | |
| **General Instructions** | |
| Error handling | The guidance is to handle errors with a wish to never stop. Regarding errors that you cannot overcome – for example one of the input files does not exists – exit the program with an appropriate message and memory release. Regarding errors that you can overcome – double edges for example – you inform the user, handle as best as you see fit, and carry on with the program. Of course, you never ever crash or exit the program in an unorderly fashion. |
| How to inform the user? | Print to stdio/stderr **"Warning: <your message>"** |
| Exe2 Code | You may change code or delete unneeded functions. Exe2 is still part of the project, and as so, parts of it might be checked. However you can assume there will be no deletion of edges/vertices – and can cut those functions out. And you may change the fashion in which you handle errors according to your new needs, if you have decided you need something different (for example, you may take in 2 nodes with the same name, and you may disregard the second one. |

| | |
|---|---|
| **Parameters** | |
| The parameters are passed as | argv |
| What they contain? | relative path + slash, and the k-ranges |
| | |
| **Project evaluation** | |
| Can we change the declaration of the k_cluster function | You most definitely can. Part of the project's checks are the design, and how you build the communication between the different parts.<br>Fill free to change whatever you want. |
| The clusters' numbers are not unique, that is clustering the nodes into 1 1 1 2 3 is the same as clustering into 2 2 2 1 3. What will happen with automatic tests? | The automatic tests take that into consideration and both clustering would be counted as equal. |
| | |
| Part 1 (Cplex and statistics) | |
| Which arrays should contain some variables with coefficient of zero? | The array obj, correlated to the optimization function, should contain all variables, even the nodes, which have no effect. For the nodes the coefficient would be zero. The rest of the arrays (matval for example) should not contain a variable with a coefficient of zero. To those arrays we would insert only the variables participating in the current constraint only. |
| What is the diameter of a cluster with isolated components (When at least one vertex in the cluster is unreachable from another vertex)? | The diameter is infinite. The output should be:<br>Cluster <cluster_no>: score - <cluster_score> diameter - inf |
| | |
| Part 2 (cytoscape and xgmml) | |
| Can I use cytoscape on windows? | Yes, you can download it for free from www.cytoscape.org/ |
| Cytoscape crashed when I try to upload any xgmml file (on linux) | Some versions of cytoscape have a problem with the 'label' attribute of the graph. Try, just for the sake of uploading to cytoscape, to change its value to "".<br>(label="") |
| When I open an xgmml file I cannot see some of the attributes | Try opening the file using notepad++ and not a browser |
| What do I need to do in order to free all of the xml elements | `Free any memory you allocated yourself`<br>`xmlFreeDoc(file);`<br>`xmlCleanupParser();`<br>`xmlMemoryDump();` |
| While we are on the subject, is there a tool we can use to validate | valgrind is considered nice and easy. Try http://valgrind.org/ .<br>It is not taught nor supported for the class, but you can try it anyway if you like. It should be helpful down the road. |

| | |
|---|---|
| memory release? | |