RAMA UNIVERSITY

www.ramauniversity.ac.in

FACULTY OF EGINEERING AND TECHNOLOGY

Soft Computing

LECTURE -30

Umesh Kumar Gera
Assistant Professor
Computer Science & Engineering

OUTLINE

- **Application of soft computing**

- **Genetic algorithm based Internet Search Techniques**

- **References**

- **MCQ**

Problem Definition

Client-based searching agents can be agents for local search (which means that they search for documents linked to the input documents) and agents for non-local search (which can find documents that are not linked to the input documents).

Broadly defined, an agent is a program that can operate autonomously to accomplish unique tasks without direct human supervision.

The agent presented in this paper named *Tropical* is an agent for global search, written in the Java programming language. For set of input documents, the agent finds and displays similar documents and the information regarding how similar they are. It uses services of following packages:

Spider [3] for document fetching

Agent [5] for Best First Search algorithm

Generator [6] for designing and communication with the database of topic sorted URLs.

GENETIC ALGORITHM BASED INTERNET SEARCH TECHNIQUES

Genetic Algorithm for Internet Search

Algorithm performs following steps:

I.   User submits input set of documents like string of URLs.

II.   From the Web are fetched the input documents and parsed in order to extract keywords.

III.   From the Web are fetched documents that are pointed to by the hyperlinks in input documents and they form the current configuration set H.

IV.   From the database is randomly picked desired number of URLs that cover the same topic as do the input documents (this topic is specified by the user at the beginning of the program) and corresponding documents are fetched from the Web and inserted in the set H.

V.   Determining the documents from the set H that are most similar to the input ones by calculating Jaccard Score for each of them.

VI.   Inserting the most similar document in the output set and adding the documents linked to it in the set H.

VII.   Repeating steps 4-6 until desired number of output documents is obtained.

Some Relevant Details

Linking the database - just once before you first start the program

Start button in Windows environment, Settings, Control Panel, ODBC, User DSN Window, Add, Microsoft Access Driver,

Finish, In the field Data Source Name put *URLbase,* Select, select the database *url.mdb* in the tropical directory, Advanced,

in the fields for username and password type *mladen* and *java*
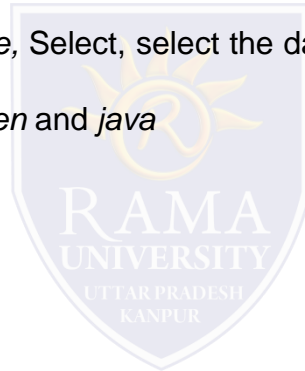
Starting the program

❑In JDK 1.1.4 Package:

❑java tropical -dProxySet=true

❑-dProxyHost=*proksi*

❑-dProxyPort=*port*

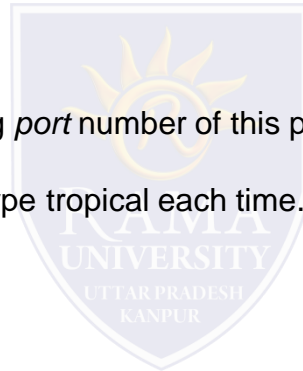❑*proksi* is IP address of proxy server you are using *port* number of this proxy server.

❑you put this line in tropical.bat file and then just type tropical each time.

❑Example of the command line is:

❑java tropical -dProxySet=true

❑-dProxyHost=*147.91.8.62*

❑-dProxyPort=*8080*

GENETIC ALGORITHM BASED INTERNET SEARCH TECHNIQUES

Input parameters

❑Input Set - field for input set of URLs. It is necessary that you enter either a set of URLs or a name of the file containing URLs or both.

❑Input file - field for the name of the file with URLs (each in the new line)

❑Number of outputs - field for the required number of output documents. Default number is 10.

❑Database mutation - list for desired number of mutated documents per iteration. Default number is 0.

❑Spatial mutation - not realized yet.

❑Temporal mutation - not realized yet.

❑Topic - list for choosing one or more topics covered by mutation.

❑Comments - comments on/off field.
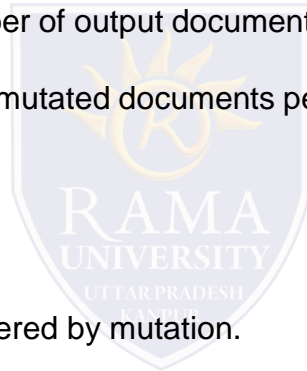
Output window

Program Flow - some program messages that keep you informed about program execution.

Result Set - result set of URLs.

Medium Jaccard's Score - medium Jaccard's Score for the result set.

Result of the program work can also be found in the file izlaz.html in the Tropical directory and can be viewed during program execution.

1. Bacteriophage Mu is also a phage, choose the correct statement for it.

a) It is an example of temperate phage

b) It is packaged into heads containing 50 kbp of DNA

c) When the phage replicates it goes transposition only a few times

d) Mu can't be used for cloning in vivo

2. Choose the incorrect statement for 'Mini Mu'?

a) It is produced by cloning of bacteriophage Mu in vivo

b) It contains reduced phage genome

c) It doesn't contain replication of origin of plasmid

d) A selectable marker is added

3. Which of the following is not included as a step for cloning mini Mu in vivo?

a) A culture of target strain, whose DNA we want to clone is infected with mini Mu

b) The phage is inserted at multiple sites in the genome

c) Some of the packaged strain is flanked by two mini Mu genes

d) The packaged phage is introduced into a second strain and Mu would replicate as a phage

4. If recombination takes place between elements in the second strain, it would replicate by the origin of replication of mini Mu.

a) True

b) False

5. Which of the following is a restriction for cloning in bacteriophage Mu?

a) It is limited to handling only small number of phage particles

b) It relies on whether the DNA is flanked by two mini Mu genes or not

c) Recombination between the elements in the second strain is not a restriction

d) Fragments of any genomic DNA can be cloned irrespective of the fact whether or not the phage can replicate in it or not

REFERENCES

❑https://shodhganga.inflibnet.ac.in/bitstream/10603/41504/12/12_chapter%202.pdf

❑https://rti.etf.bg.ac.rs/rti/ebi/dipl/jelena/paper.htm