Global Journal of Advanced Computer Science and Technology

Volume No. 12

Issue No. 3

September - December 2024

ENRICHED PUBLICATIONS PVT.LTD

JE - 18,Gupta Colony, Khirki Extn,
Malviya Nagar, New Delhi - 110017.
E- Mail: info@enrichedpublication.com
Phone :- +91-8877340707

Global Journal of Advanced Computer Science and Technology

Aims and Scope

Global Journal of Advanced Computer Science & Technology deals with information technology, its evolution and future prospects and its relationship with the Business Management. It addresses technological, managerial, political, economic and organizational aspects of the application of IT in relationship with Business Management. The journal will serve as a comprehensiveresource for policy makers, government officials, academicians, and practitioners. GJACST promotes and coordinates the developments in the IT based applications of business management and presents thestrategic roles of IT and management towards sustainable development.

Global Journal of Advanced Computer Science and Technology

Global Journal of Advanced Computer Science and Technology

(Volume No. 12, Issue No. 3, September - December 2024)

Contents

Lexicographical Order/Sorting using Boost Library Implementation

Jayant Kumar*
* Director of Platform Integration and Architecture, Bidtellect Inc.,
Delray Beach, USA - 33432

A B S T R A C T

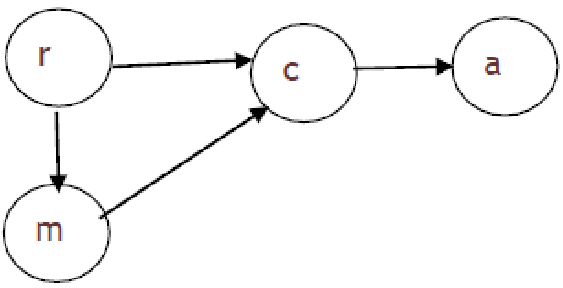
Background:
the lexicographic or lexicographical order (also known as lexical order, dictionary order, alphabetical order or lexicographic(al) product) is a generalization of alphabetically ordered based on the alphabetical order of their component letters
.
This generalization means that the order is not based on alphabetical order but based on relationship between two letters or entities.

Example:
Given that r < c, c < a, r < m, m < c
So, the lexicographical order would be r,m,c,a

I. INTRODUCTION:

Boost provides free peer-reviewed portable C++ source libraries. We can create an Adjacent graph by using boost adjacency list.

So, if we have relationship like r < c and further relationship to create graph as below.



We can apply topological sorting to get a ordered list.

Topological sorting for Directed Acyclic Graph (DAG) is a linear ordering of vertices such that for every directed edge uv, vertex u comes before v in the ordering.

Since Boost is a C++ library, we will use C++ for implementation.

IMPLEMENTATION:

Let's implement the same by using a file which will provide relationship between letters.

The file is comprised of a sequence of words that are arranged in alphabetical order (for some arbitrary alternate alphabet)

File will have content like as below.

rcrtv

rcrmb

rcrdsfd

rcxrbw

rcxrbws

rcxrbwn

rcxrbkdz

rsqxbwarbw

rsqxbwarbws

rsqxbwarbwn

rsqxbwarbkdz

rsqxbwawa

rsqxbwawan

rsqxbwafqn

rsqxbwafqnxh

rjwdkbkwn

rjwdkbh

rjwdfaapwr

rjwaktkqj

rjkrcxw

rjktrcxw

rjktrcxh

rjks

rjksw

rjksnb

By reading each word one by one, we find the relationship between two letters and created a directed graph.

Once we are done reading all words, we will apply topological sort to get final output. Implementation is divide into various steps.

STEP1:

Let's create a header file named "LexcoSorting.h" as below.

We will be declaring labeled_graph Graph which will be used to create graph while reading various words from give file and fnding relationship between two letters.

vertex_iter would be used to traversed through various vertex of the graph AddVertex would be used to add new vertex if vertex doesn't exist in the graph AddEdge would be used to add Edge between two give vertex

```
#ifndef LexcoSorting_H
#define LexcoSorting_H

#include "boost/graph/adjacency_list.hpp"
#include "boost/graph/labeled_graph.hpp"
#include "boost/graph/topological_sort.hpp"
#include <deque>
#include <iterator>
#include <iostream>
#include <fstream>
#include <string>

using namespace std;
using namespace boost;


class LexcoSorting
{
    struct VertexProperty
    {
        char c_literal;
    };

    typedef boost::labeled_graph<boost::adjacency_list< boost::vecS,
boost::vecS, boost::directedS,VertexProperty>,char> Graph;
    typedef boost::adjacency_list<>::vertex_descriptor Vertex;
    //typedef boost::labeled_graph<boost::setS,
boost::vecS,boost::directedS,VertexProperty> Graph;
```

```
typedef boost::graph_traits<Graph>::vertex_iterator vertex_iter;
typedef std::vector<Graph::vertex_descriptor>  Vcontainer;

Graph g;
//    list<Vertex> lVertex = new list<Vertex>(30);

void AddVertex(char c_temp);
void AddEdge(char sFirst,char sSecond);
void PrintOutPut();

public:
    string Compare(string sPrevious,string sCurrent);
    void Execute(string sInputFile);
};

#endif // LexcoSorting_H
```

STEP2:

Create Another file LexcoSorting.cpp with below code which will define all methods declared in LexcoSorting.h

AddVertex – Add a new vertex in the graph

PrintOutPut – Apply topological sort and prints the final sorted output

AddEdge – Add Edge between two letters passed as input

Compare – Compare two strings and find the relationship (order) between two letters by comparing two words as both words are in alphabetical order.

Execute: It's the main method which read word by word in the file and create Vertex if it doesn't exist or compare previous and current letters and create Edge between two if doesn't exist already.

Code:

```
#include "LexcoSorting.h"

void LexcoSorting::AddVertex(char c_temp)
{
    VertexProperty v1 ;
    v1.c_literal = c_temp;
  //Vertex vtemp = boost::add_vertex(c_temp,g);
  boost::add_vertex(c_temp,v1,g);
  //g[boost::add_vertex(c_temp,g)].c_literal = c_temp;
}
```

```
void LexcoSorting::AddEdge(char sFirst,char sSecond)
{
    boost::add_edge_by_label(sFirst, sSecond, g);
}

void LexcoSorting::PrintOutPut()
{
    Vcontainer c;
    Vcontainer::iterator ii;

    topological_sort(g.graph(), std::back_inserter(c));

     std::cout << "A topological ordering: ";
     for ( ii=c.begin(); ii!=c.end(); ++ii)
     {
         //cout << *ii << " ";
         //cout << *ii << " ";
         cout << g.graph()[*ii].c_literal << " ";
     }

 }

string LexcoSorting:: Compare(string sPrevious,string sCurrent)
{
    string op;
    int cnt=0,i=0;
    while(sPrevious[i] !='\0' || sCurrent[i] !='\0')
    {
            if(sPrevious[i] == sCurrent[i]) {
                    cnt = 1;
            }
            else {
                break;
            }
            i++;
         }
         if(cnt > 0) {
            op[0] = sPrevious[i];
            op[1] = sCurrent[i];
            op[3] = '\0';
        }
        return op;

    }
```

```
void LexcoSorting::Execute(string sInputFile)
{
    /* //For Testing
    AddVertex('a');
    AddVertex('b');
    AddVertex('c');
    AddVertex('d');
    AddEdge('a','b');
    AddEdge('d','c');
    AddEdge('b','c');
    //AddEdge('a','b',*g);
    PrintOutPut();*/

     std::ifstream infile(sInputFile);
     std::string strcurrent,strprevious,strCompare;

      while (std::getline(infile, strcurrent))
      {
          if(!strcurrent.empty())
          {
if(strprevious.empty())
{
      for(char& c : strcurrent) {
          AddVertex(c);
      }
}
else
{
      strCompare = Compare(strprevious,strcurrent);
      if(!strCompare.empty())
      {
          AddEdge(strCompare[0],strCompare[1]);
      }
      AddVertex(strCompare[0]);
      AddVertex(strCompare[1]);

        }
        strprevious = strcurrent;
        //cout << strcurrent;
        //file_contents.push_back('\n');
      }
```

STEP4:

Create a makefile with below code to compile the program which declare all dependencies.

Code:

```
# compiler:
            CC =   g++


            # compiler flags:
            CFLAGS = -std=c++11 -g -Wall

            #Linking Flag
            LFLAGS =      -Wall

            INCLUDES =    -I C:/boost_1_59_0 -I C:/MinGW -I C:/boost_1_59_0/boost/graph
            LIBS =
            # the build target executable:
            TARGET = Lexicograph

            $(TARGET):    main.o LexcoSorting.o
                   $(CC) $(CFLAGS) -o $(TARGET) main.o LexcoSorting.o

             main.o:main.cpp LexcoSorting.h

                   $(CC) $(INCLUDES) $(CFLAGS) -o main.o -c main.cpp


            LexcoSorting.o:       LexcoSorting.cpp LexcoSorting.h
                   $(CC) $(INCLUDES) $(CFLAGS) -c LexcoSorting.cpp


            clean:
                   $(RM) $(TARGET) *.o *~
```

USE CASE AND IMPACT:

GENETIC SCIENCE:

It can be used to create genetic database as if we know relationship like FATHER -> SON, MOTHER-DAUGHTER we can get entire genetical order.

HEALTH CARE:

We can also create use the relationship between cause and symptom of a disease. That information can be used to order the symptoms of disease in a perfect order and we can keep track of hour health and discover any disease in early stages.

```
}
PrintOutPut();


}
```

STEP3:

Edit/Create file main.cpp which would be the main file compiled and executed. It will take path of the input file which have alphabetical order of words.

Code:

```
#include <iostream>
#include "LexcoSorting.h"
#include <exception>


using namespace std;

int main()
{
    try
    {
        LexcoSorting l;
        cout << "Enter File Path and Name" << endl ;
        char cfilename[200];
        cin.getline(cfilename,sizeof(cfilename));
        cout << "File Name : " << cfilename << endl;
        l.Execute(cfilename);
              getchar();
                 //"C:\\Users\\Jayant Kumar\\Documents\\alphabet.txt");
    }
    catch (const std::exception& e)
    {
        cout << e.what() << endl;

              }


              return 0;
          }
```

CRIMINOLOGY;

Same techinque can be applied to various crime cases and we can solve few complex criminal cases by ordering every aspect and story point of a crime.

REFERENCES

[1] https://en.wikipedia.org/wiki/Lexicographical _order
[2] http://www.boost.org/
[3] http://www.boost.org/doc/libs/master/libs/gr aph/doc/adjacency_list.html
[4] https://en.wikipedia.org/wiki/Glossary_of_gr aph_theory_terms#adjacent
[5] http://www.geeksforgeeks.org/topological- sorting/

"Transforming Data Flow Diagram to use Case Diagram"

Shinde Ajaykumar D.[*], Dr. M.S. Prasad[**]
[*] Dept. of Computer Studies, CSIBER, Kolhapur, Research Student
[**] Professor, Bharati Vidyapeeth's Institute Of Management And Entrepreneurship Development, Erandavane, Pune

ABSTRACT

In this paper we present an approach that transforms the DFD to the Use Case diagram of UML. DFD is a structured approach which provides the functional view of the system whereas Use Case diagram is an object oriented approach provides the functional view of the system under consideration. Structured methods are very commonly used by the developers and if there is need to expand the functionality of the systems then object oriented approach is used which is very useful. So the transformation of one approach to the other will be beneficial for the developers. For this we present an approach that will transform Data Flow Diagrams (Major tool of Structured Approach) with Use case diagram.

Key Words: DFD (Data Flow Diagram), Object Oriented Approach, Structured Approach, UML (Unified Modeling Language), UCD (Use Case Diagram)

1. INTRODUCTION

Many organizations are using the software that was designed and developed at least a decade ago. The approach used to design and develop these systems was procedure oriented and the same approach was reflected in documents [18]. The procedure oriented approach has become outdated and the focus is shifting to object oriented. As replacing the existing software puts extra burden on the users in terms of cost. Many customers are continuing with existing software. It is possible to implement changes in the code so as to shift software from procedure oriented to object oriented, but this will create a new problem document and code can not match. The only feasible solution for up-gradation and maintenance is to preserve system design and incorporate it with latest software development strategies [7]. It is possible to generate design using reverse engineering with the help of available code. But if frequent changes are made to code it becomes inconsistent with the design. The user also feels the original system is irreplaceable and trustworthy [9].

In procedure oriented approach DFD is treated as main artifact for system representation. A DFD is must for each and every system designed using procedure oriented approach. The main advantage in using DFD is it shows dynamic approach of the system [9]. Procedure oriented is being replaced by object oriented approach and is becoming the only approach for design and development. Many organizations are shifting from procedure oriented to object oriented approach [11]. . For design of object-oriented systems and creation of model, Unified Modeling Language[17] has now become the industry standard [2][3]. UML is a collection of diagrams used to represent different aspects of the system under consideration. UML allows us to represent static structure of the system as well as dynamic behavior of the system.

In [12] Liu and Wilde, have proposed type base and global base object finder methodologies for identifying object from non-object-oriented languages. In [8] Jacobson and Lindstrom, discuss reverse engineering strategies for object-oriented model to incorporate changes. Newcombe and Kotik [13] present a tool for abstract object-oriented model generation. Subramanian and Bwirne [15] generate objects from FORTRAN code. They discuss constraints like private, virtual, and pure virtual. Cimitile and others [4] and De Lucia and others [5] present approaches that revolve around data stores. Authors propose approaches that consider functions and subroutines, interacting with tables, data-store and use them as objects methods. De Lucia and others in [6] propose an approach to recover class diagram from system code that is highly data intensive. From the above discussion it is clear that all the techniques are dependent on code. In our approach, we are more interested in procedure oriented design than code. In design, we have observed that in literature, both structured design to non-UML design and structured design to UML design transformations exist.

2. DATA FLOW DIAGRAM AND USE CASE DIAGRAM

a) Notations used in Data Flow Diagram

The notations for DFD were proposed and popularized by Yourdon, DeMarco, and others are described below:

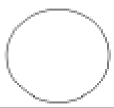
symbol	Name	purpose
	processes	To show work to be carried out in the system
	Source and consumer of data	To show external users of system
	Data flow	To show data flowing through the system
	Data store	To show where the data is stored

Table 1 : Symbols used to draw DFD

b) Use Case Diagram

The Use case diagram is drawn to identify the primary elements and processes that form the system. Primary elements are termed as "actors" and the processes are called "use cases", Use case diagram shows which actors interact with each use case.

A use case diagram captures the functional aspects of a system. More specifically, it captures the business processes carried out in the system. As we discuss the functionality and processes of the system, we discover significant characteristics of the system that we model in the use case diagram. Due to the simplicity of use case diagrams, and more importantly, because they are shorn of all technical jargon, use case diagrams are a great tool for user meetings. Use case diagrams have another important use. Use case diagrams define the requirements of the system being modeled and hence are used to write test scripts for the modeled system.

A use case diagram is quite simple in nature and depicts two types of elements: one representing the business roles and the other representing the business processes. Let us take a closer look at use at what elements constitutes a use case diagram.

Element	Name	Meaning
<<role>>	Actor	External user of the system
Work name	Use case	Work carried out in the system/ functionality expected by user
———	connecter	Connection between user and use cases

Table 2 : Symbols used in Use Case diagram

3. REPRESENTATION OF DFD IN FRAMEWROK

All transformation processes are dependent on representation of DFD. In order to simplify the transformation process we have designed and used a framework [14]. A data flow diagram uses very limited number of symbols and may be represented as a set of symbol sets. In the framework the DFD is represented as a graph using atomic relational grammar [1]. In the framework [14] DFD is defined as

DFD = {{SS}, {PS}, {DS}, {TS}, {RS}, {PR}} where,

DS represents set of data flows

PS represents set of processes that may be either atomic or aggregate

TS represent set of data stores

SS represent set of source consumers

RS represents the set of relationships

PR represents the set of productions

The framework treats DFD as directed graph in which the sets {SS}, {PS}, {TS} are the vertices and the set {DS} is the set of edges. {RS} is a set of relationships between the atomic elements of DFD and {PR} is set of productions derived from elements of {RS}. A member of the set {SS} will be a start symbol [14].

a) Transformation of DFD to Use Case Diagram

For transformation of DFD to use case diagram strategy is defined in this paper. Every transformation strategy should be based on concrete rules [10][16]. The transformation strategy presented in this paper is also based on rules. As all diagrams in procedure oriented and object oriented design are represented as graphs. We have designed a strategy that is based on patterns, as patterns strategy deals with graphs and representation of graph is based on concrete or abstract syntax of source or target model language [16]. The transformation strategy adopted in this paper constructs a intermediate model using tagged language [14]. Transformation rules are framed and applied to entire model rather than a specific location in the model. Application of the transformation rule generates a new model; same set of rules is applied iteratively to all matching locations in the source model. The transformation rules are applied in phases where each phase has a specific purpose and it invokes a definite rule of transformation. The transformation rules are organized according to the source language and target language i.e. DFD and use case diagram. The rule application strategy used here is unidirectional [16] i.e. transformation rule are used to transform source model to the target model reverse is not possible.

The transformation strategy adopted in this paper is Hybrid; it is a mixture of direct manipulation, relational approach and graph transformation approach [16]. In direct manipulation approach API and internal model representation is provided. In relational approach the type of the element is specified using relational constraints. This approach also has relational specification and mapping rule. In graph transformation approach LHS and RHS sides are used. The LHS pattern is matched in the model being transformed and replaced by the RHS pattern. As these properties are incorporated in the framework, it becomes hybrid strategy for transformation [16].

The transformation strategy designed in this paper starts with the scanning of existing model/graph. The scheme adopted for transformation is model-to-model mapping of symbols.This approach offers an internal model representation plus some API to manipulate it. Since the diagram is represented internally as distinct sets of symbols in the framework, each symbol set is scanned separately. Framework goes on scanning each and every element from DFD and identifies its type. The scanning process identifies attributes of the elements and is very essential process for mapping the symbols from DFD to use case diagram.

As the first step of transformation, the framework starts scanning source consumer set(SS set). For each source and consumer in the data flow diagram a new actor is added to the actor set in use case diagram. The source consumer in DFD and actors in use case diagrams represent the external users of the system. The attributes of source and consumer from DFD are recorded as attributes of actors in use case diagram. After completion of scanning of the source and consumer set the frameworks starts scanning the process set.

In the second step of transformation, process set(PS set) is taken. For each process in process set of DFD framework adds a new use case in the use case set of use case diagram. While scanning the process set the framework looks for the expansion attribute of the process in DFD. If the expansion attribute of process is true, it implies that the process is expanded in the next level of the DFD, i.e. the process expanded is made up of many other sub processes. For each level of the DFD the scan process is applied iteratively to the next level of the DFD. The framework opens the expanded DFD and starts scanning it. For each process in DFD a new use case is added to the use case diagram. In case of expanded processes the framework maintains a relationship between expanded and processes in new DFD as generalization. The generalization relationship is used to show relationship between more general and more concrete use case. The generalization relationship may be named as extends or uses depending upon, whether it adds something new to existing use case or it is integral part of existing use case.

In the third step of transformation, framework starts scanning relationships set(RS set). Relationship set has start and end attributes. If the relationship is either starting or ending on the data store the relationship is not recorded in relationship set of the use case diagram. This is because use case diagrams do not have data stores concept. For all the other relationships it goes on adding a new relation in the relation set in the use case diagram. The fourth step of scan starts on the data store set(TS set) of DFD. As data stores are not part of the use case diagrams this entire set is skipped. In the last step framework scans data flows; data flow coming into the data stores and the data flows that are coming out of the data stores(DS set) are omitted from the use case set. All the other data flows are stored as links in the use case set.

b) Algorithm for Transformation of DFD to Use Case

We have proposed an algorithm for transformation of DFD to use case diagram. It starts by reading the symbol set from DFD. It looks at the shapes and other characteristics of symbol to identify its type i.e. rectangle is source consumer of information, circle is process, arrow is a data flow with start and end characteristics, parallel lines represent data stores. This information is used for mapping the symbols from DFD to use case diagram.

Algorithm DFDtoUCD
Input : DFD subsets {SS}, {PS}, {DS}, {TS}, {RS}
Output : use case set
//start reading the DFD from first symbol set to
	//if circle in DFD is expanded in next level
	Else
	Get next level DFD
	Call algorithm recursively

last symbol set
If symbol set is empty
return
else
For each symbol in DFD set
// if the symbol from DFD set is rectangle then add stickman symbol to use case set
If symbol.type = source consumer then
Add actor symbol to use case set.
//if symbol from DFD is circle then add oval to use case set
Else if symbol.type = process then
// if circle in DFD is not expanded in next level
If process.expanded = false then
Add use case to use case set
	//If the symbol from DFD is arrow
	Else if symbol.type = dataflow then
	//do not record links with datastore
	If dataflow.source != datastore or data store.destination != data store
	Then
	Record link to use case set
	//do not record links with datatore
	Else if relation.start != data store or relation.end != data store then
	Record relationship to the use case set
	//do not consider the datastore symbol for conversion
	Else if symbol.type = data store then
	Skip the symbol
	Get next symbol
	Endif

Algorithm-1: Algorithm to transform DFD to Usecase diagram.	

The algorithm starts reading the DFD. As the DFD is represented using sets, it starts reading these sets one by one. It reads the terminal symbol sets first. After reading the symbols from terminal set, it identifies the type of the symbol. Once the type is known, it finds out the mapping symbol from use case set and adds it to the use case. The mapping set is given in Table- 3. For the data store symbol there is no equivalent symbol. All data store symbols and the links associated with data store symbol are skipped by the algorithm. The mapping process obtains the basic symbols required for drawing the use case diagram. The framework accepts these symbols and by looking at the link set goes on drawing the final use case diagram. The mapping symbol set used by the algorithm for transformation of DFD to use case diagram is given in Table 3. While mapping the symbols from DFD to use case the framework makes it sure that the symbol is not duplicated. This helps us in reducing the number of diagram elements and removing redundancy of element in the translated diagram.

DFD		Use Case Diagram	
Symbol	Meaning	Symbol	Meaning
	Source and consumer of data		User of the system
	Processes in the system		Use case in the system
	Data flow and its direction		Link between user & use case
	Storage of the data	No corresponding symbol available in use case model. The data store becomes class in class model.	

Table 3 : Mapping of symbol set from DFD to use Case Digram

The mapping of the symbol is done by looking at the purpose of each symbol e.g. the process symbol in DFD use used to show the functionality performed by the system similarly the use case symbol in use case diagram is used to show functionality in the use case diagram. As the purpose of both the symbols is same the process symbol from DFD is mapped to use case symbol in use case diagram. The source and consumer of data in DFD plays the role of the user. Actor in the use case diagram is user of the system. The source consumer symbol is mapped to actor symbol in use case diagram. A data flow is a link between two elements in the DFD. The data flow in DFD is mapped to a link between the elements of the use case diagram. As use case diagram has no concept of data store all the data stores and their associated links are dropped from the use case set.

4. EXAMPLES FOR TRANSFORMATION OF DFD TO USE CASE DIAGRAM

The framework [14] provides a toolbox to draw a data flow diagram. The elements from toolbox have to be selected by the users and paste on the drawing area. While connecting the symbol a line must start within the boundaries of the element and also must end in the boundaries of the element. The drawing procedure validates the connection between atomic elements of the diagram. i.e. if we try to connect sources and consumers to each other, error message is displayed and connection is rejected. All validations are based on the relationship between elements using ARG [1]. When the drawing is over the diagram is saved using a tagged language as shown in section 4(a).

a) Representation of DFD in Framework

The DFD is represented in the framework using tagged language. In this section we present tagged language generated by the framework for the Figure 1. The DFD is represented as multiset of symbols with attributes the symbol sets are written separately. A set of relationship is additional set written as a result of drawing it includes the binary relationships between the symbols in DFD.

Consider DFD drawn in Figure 1, it is represented in the framework as follows. Internal representation of DFD in figure 1 is given below it. It uses a tagged language for representation of DFD.

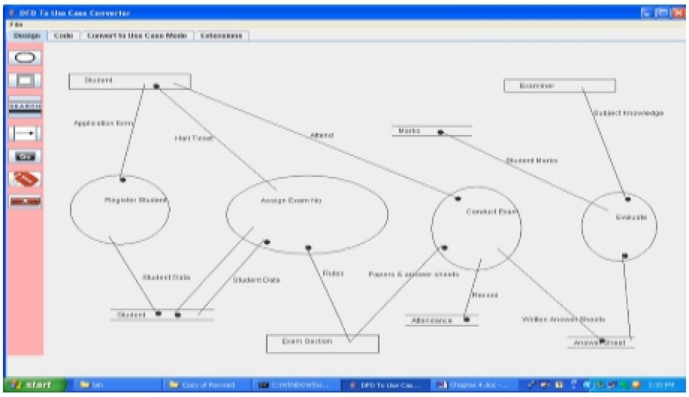


Figure 1 Level 1 DFD for Examination System

```
<DFD>
        <Source-Consumer>
        <%rectangle id=0 nm=" Student ", x=93 ,
y=62 , w=181 , h=30 %>
        <%rectangle id=1 nm=" Examiner ",
x=746 , y=73 , w=166 , h=24 %>
        <%rectangle id=2 nm=" Exam Section ",
x=389 , y=579 , w=168 , h=35 %>
        </Source-Consumer>
    <Relationship>
        <%relation start=Student, end=Register
Student %>
        <%relation start=Assign Exam No,
end=Student %>
        <%relation start=Student, end=Conduct
Exam. %>
        <%relation start=Examiner, end=Evaluate
%>
        <%relation start=Register Student,
end=Student %>
        <%relation start=Assign Exam No,
end=Student %>
        <%relation start=Student, end=Assign
Exam No %>
        <%relation start=Exam Section,
end=Assign Exam No %>
        <%relation start=Exam Section,
end=Conduct Exam. %>
        <%relation start=Conduct Exam.,
end=Attendance %>
        <%relation start=Conduct Exam.,
end=Answer Sheet %>
        <%relation start=Answer Sheet,
end=Evaluate %>
        <%relation start=Evaluate, end=Marks %>
        .</Relationship>

        <Process>
        <%oval id=0 nm="Register Student",
x=95 , y=262 , w=148 , h=138 %>
        <%oval id=1 nm="Assign Exam No",
x=329 , y=263 , w=242 , h=156 %>
        <%oval id=2 nm="Conduct Exam.",
x=636 , y=285 , w=135 , h=166 %>
        <%oval id=3 nm="Evaluate", x=861 ,
y=296 , w=113 , h=137 %>
        </Process>

        <Line>
 <%Joiner x1=206 , y1=82 , x2=169 ,
 y2=267 , nm="Application form" , x=101 , y=164
%>
 <%Joiner x1=404 , y1=292 , x2=219 , y2=82 ,
nm="Hall Ticket" , x=252 , y=194 %>
 <%Joiner x1=249 , y1=80 , x2=671 , y2=305 ,
nm="Attend" , x=455 , y=188 %>
 <%Joiner x1=862 , y1=89 , x2=926 , y2=306 ,
nm="Subject Knowledge" , x=879 , y=144 %>
 <%Joiner x1=153 , y1=384 , x2=222 , y2=532 ,
nm="Student Data" , x=203 , y=470 %>
        <%Joiner x1=369 , y1=366 , x2=251 ,
y2=535 , nm=" " , x=0 , y=0 %>
 <%Joiner x1=286 , y1=539 , x2=385 , y2=391 ,
nm="Student Data" , x=339 , y=475 %>
 <%Joiner x1=514 , y1=593 , x2=447 , y2=402 ,
nm="Rules" , x=473 , y=462 %>
 <%Joiner x1=514 , y1=593 , x2=651 , y2=402 ,
nm="Papers & answer sheets" , x=542 , y=464 %>
 <%Joiner x1=710 , y1=429 , x2=684 , y2=544 ,
nm="Record" , x=697 , y=505 %>
 <%Joiner x1=735 , y1=408 , x2=887 , y2=586 ,
nm="Written Answer Sheets" , x=773 , y=553 %>
        <%Joiner x1=935 , y1=589 , x2=922 ,
y2=418 , nm=" " , x=0 , y=0 %>
 <%Joiner x1=900 , y1=333 , x2=645 , y2=173 ,
nm="Student Marks" , x=748 , y=238 %>
        </Line>

        < DataBase>
        <%Database nm="Marks", x1=577 , y1=164
, x2=696 , y2=164 %>
        <%Database nm="Student", x1=155 ,
y1=529 , x2=310 , y2=529 %>
        <%Database nm="Attendance", x1=597 ,
y1=540 , x2=706 , y2=540 %>
 <%Database nm="Answer Sheet", x1=840 ,
y1=584 , x2=941 , y2=584 %>
        </DataBase>

        <ID relationship>
        <% sid=0 eid=0 %>
        <% sid=0 eid=1 %>
        <% sid=0 eid=2 %>
        <% sid=1 eid=3 %>
        <% sid=2 eid=1 %>
        <% sid=2 eid=2 %>
        </ID relationship>
</DFD>
```

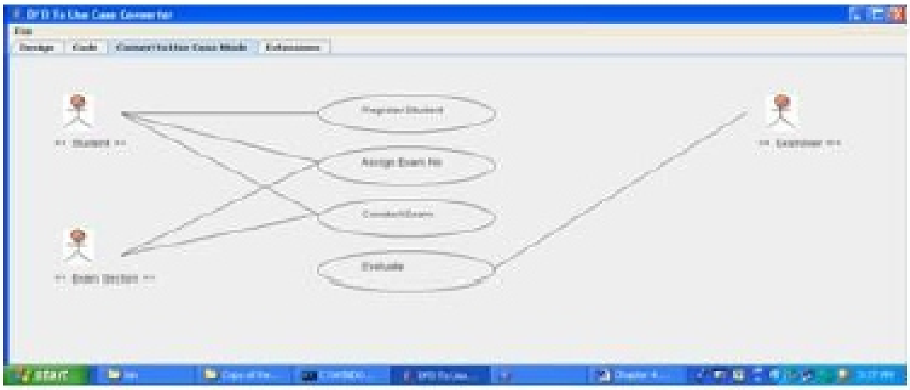


Figure 2 : Conversion of level 1 DFD to use Case Digram

The representation of DFD in the framework uses multi-sets of symbols. The tagged language records these sets separately using a definite format. The sets include set for source and consumers, set for processes, set for data stores, set for data flows and the data flow set is used to generate another set called relation set. The output of drawing of data flow is shown with the tagged language. Figure 2 shows conversion of level 1 DFD shown in figure 1. The conversion of DFD to use case diagram uses the algorithm described in section 3(b).

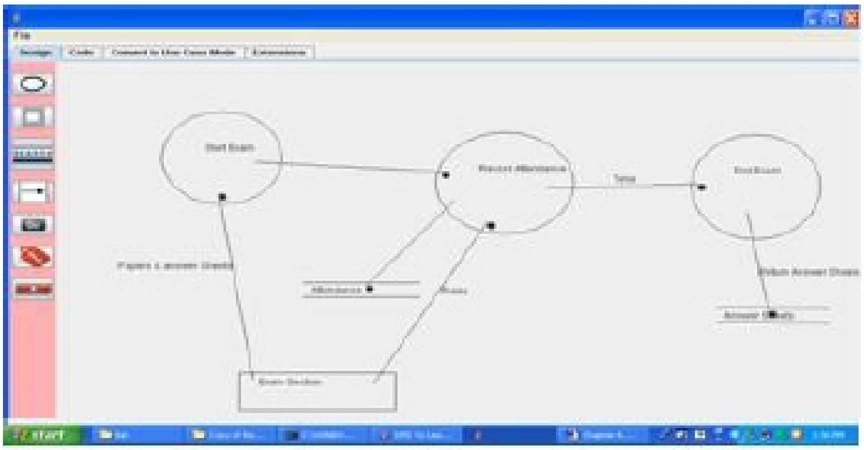


Figure 3 : Expansion of Process 3 to Level 2 DFD

Figure 3 shows the expansion of process 3 from level 1 DFD. The expansion of a process in framework is done by selecting a process and clicking on 'Go' button in tool box.

b) Representation of Level2 DFD in Framework

Representation of level-2 DFD shown in figure 3 in the framework

```
<DFD>
        <Source-Consumer>
        <%rectangle id=0 nm=" Exam Section ",
x=274 , y=431 , w=186 , h=51 %>
        </Source-Consumer>
        <Relationship>
        <%relation start=Exam Section, end=Start
Exam %>
        <%relation start=Exam Section,
end=Record Attendance %>
        <%relation start=Start Exam, end=Record
Attendance %>
        <%relation start=Start Exam, end=Record
Attendance %>
        <%relation start=Record Attendance,
end=Attendance %>
        <%relation start=Record Attendance,
end=End Exam %>
        <%relation start=Record Attendance,
end=End Exam %>
        <%relation start=End Exam, end=Answer
Sheets %>
        .</Relationship>
        <Process>
        <%oval id=0 nm="Start Exam", x=181 ,
y=68 , w=144 , h=129 %>
        <%oval id=1 nm="Record Attendance",
        x=509 , y=97 , w=165 , h=141 %>
        <%oval id=2 nm="End Exam", x=814 ,
y=100 , w=155 , h=145 %>
        </Process>
        <Line>
        <%Joiner x1=290 , y1=439 , x2=249 ,
        y2=182 , nm="Papers & answer Sheets" ,
        x=129 , y=287 %>
        <%Joiner x1=434 , y1=446 , x2=570 ,
        y2=222 , nm="Rules" , x=515 , y=322 %>
        <%Joiner x1=293 , y1=138 , x2=516 ,
y2=152 , nm=" " , x=0 , y=0 %>
        <%Joiner x1=531 , y1=192 , x2=425 ,
y2=310 , nm=" " , x=0 , y=0 %>
        <%Joiner x1=642 , y1=174 , x2=822 ,
        y2=169 , nm="Time" , x=721 , y=167 %>
        <%Joiner x1=881 , y1=208 , x2=907 ,
        y2=345 , nm="Return Answer Sheets" ,
        x=895 , y=295 %>
        </Line>
        < DataBase>

        <%Database nm="Attendance", x1=350 ,
y1=306 , x2=490 , y2=306 %>
        <%Database nm="Answer Sheets", x1=845 ,
        y1=341 , x2=979 , y2=342 %>
        </DataBase>
        <ID relationship>
        <% sid=0 eid=0 %>
        <% sid=0 eid=1 %>
        <% sid=-1 eid=1 %>
        <% sid=-1 eid=2 %>
        </ID relationship>
</DFD>
```

Figure 4 is transformation of level2 DFD shown in figure 3. The procedure for transformation is same as that of level1 DFD to use case diagram.

Figure 4 : Transformation of expansion of Process3 from fig.

Figure 5 shows the relationship between process in level-1DFD and its expansion in level 2 DFD. The process in level 1 and its sub-processes are in relations. The only relationship possible between these processes is generalization. Figure 5 shows this relationship between a process and its sub-process using use case notations.

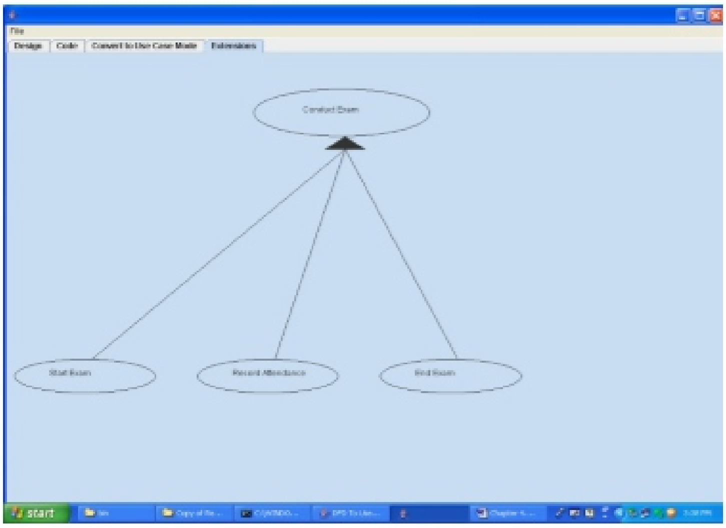


Figure 5 Relationship between Process3 and it's Expansion

CONCLUSION

As software industry is shifting from procedure oriented paradigm to object oriented. It is becoming essential to find out commonalities and differences between these two approaches. An attempt has to be made to establish connection between these approaches. The approach suggested in this paper is an attempt to establish connection between these approaches. The transformation strategy presented in this paper allows the user to draw DFD also it understands syntax and semantics of DFD. The main advantage of this framework is it stores diagram using tagged language which is easy to read and understand. On the other hand because storage is done in textual format is saves disk space. The transformation algorithm presented in this paper converts DFD to correct Use case diagram. The framework also understands relationship between leveled DFD's and this information is used to establish generalize relationship between use cases.

REFERENCES :

[1]. Antti-Pekka Tuovinen Object-Oriented Engineering of Visual Languages, PhD Thesis, Series of Publications A, Report A-2002-, Helsinki, February 2002, 185.

[2]. Booch, G., Jacobson, Rumbaugh, J. I. (2005). Unified Modeling Language User Guide. (2nd Edition), Addison Wesley, Upper Saddle River, NJ.

[3]. Booch, G., Rumbaugh, J., Jacobson, I. (1999). The Unified Modeling Language User Guide. Addison Wesley.

[4]. Cimitile, A., De Lucia, A., Di Lucca, G.A., Fasolino, A.R. (1997). Identifying objects in legacy systems. In: Proceeding Of 5th International Workshop on Program Comprehension, Dearborn, Michigan, IEEE CS Press, pp. 138-147.

[5]. De Lucia, G.A., Di Lucca, G.A., Fasolino, A.R., Guerra, P., Petruzzelli, S. (1997). Migrating Legacy Systems towards Object-Oriented Platforms. In: 13th International Conference on Software Maintenance ICSM'97.

[6]. De Lucia, G.A., Fasolino, A.R., Carlini, U. D. (2000). Recovering Class Diagrams from Data-Intensive Legacy Systems. In: 16th IEEE International Conference on Software Maintenance ICSM'00.

[7]. Dietrich, W., Nackman, I., Gracer, L. (1989), Saving a legacy with objects. In Proceedings of OOPSLA, pp. 77-88.

[8]. Jacobson, I., Lindstrom, F. (1991). Re-engineering of old systems to an object-oriented architecture. In: Proceedings of OOPSLA, pp. 340-350.

[9]. Kendall, K. E., Kendall J. E. (1999). System Analysis and Design. (3rd Edition), Prentice Hall.
[10]. Krzysztof Czarnecki and Simon Helsen, A classification of model transformation approaches OOPSLA' 03 Workshop on Generative Techniques in the Context of Model- Driven Architecture.
[11]. Larman, C. (2004). Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design. (3rd Edition), Addison Wesley Professional.
[12]. Liu S., Wilde, N. (1990). Identifying objects in a conventional procedural language: an example of data design recovery. In: Proceedings of Conference on Software Maintenance, San Diego, CA, IEEE CS Press, pp. 266-271.
[13]. Newcombe, P. and Kotik, G. (1995). Reengineering procedural into object-oriented systems. In: Proceeding of 2nd Working Conference on Reverse Engineering, Toronto, Canada, IEEE CS Press, pp. 237-249.
[14]. Shinde Ajaykumar and Dr. M.S. Prasad Automic Relational Grammer(ARG) as a means for Representing Data Flow Diagram International Journal of Information Systems, ISSN 2229-5429, Vol. III Issue II, December 2012.
[15]. Subramaniam, G.V. and Bwirne, E. J. (1996). Deriving an object model from legacy FORTRAN code. In: Proceeding of International Conference on Software Maintenance, Monterey, CA, IEEE CS Press, pp. 3-12.
[16]. Thu Nga Tram, Khaled M. Khan, Yi-Chen Lan, A Framework for Transforming Artifacts From Data Flow Diagram to UML, Technical Report No. CIT/04/2004. IASTED International Conference on Software Engineering Innsbruck, Austria.
[17]. UML (Unified Modeling Language): Superstructure. Version 2.1.2. Object Management Group (OMG). Document 07-11-02.pdf From http://www.omg.org/spec/UML/2.1.2/ [Accessed: March 1, 2008].
[18]. Yourdon, E. (1989). Modern Procedure oriented Analysis. Yourdon Press. Upper Saddle River, NJ.

Comparative Study of FFT/IFFT Processor for High Throughput-Rate Application

Krishan Kumar[1], Sonal Dahiya [2]
[1] ASET, Amity University Haryana, India
[2] ASET, Amity University Haryana, India

ABSTRACT

This paper proposes a FFT processor with multiple pipeline architecture for high throughput-rate applications. The power consumption and hardware cost can also be save in this processor by using the higher radix FFT algorithm and less memory and complex multipliers.

INTRODUCTION:

Fast Fourier transforms (FFT) and inverse fast Fourier transform (IFFT) are the key components ultra wideband system. However, it is a challenge to realize the UWB physical layer in VLSI implementation. So, it is important to desire high data rate transmission in time dispersive or frequency selective channels without having complex time domain channel equalizer but also can provide high spectral efficiency.

Although the memory-based architecture is considered most area efficient, it requires many computation cycles. Therefore we propose a multipath pipeline based architecture for high- speed application. The pipeline architecture has also the advantage of being highly regular, which can be easily scaled and parameterized in hardware design.

Design issue of the FFT processor for high-throughput rate application

Various architectures such as single-memory architecture, dual-memory pipeline architectures, array architecture and cached memory have been proposed over the last three decade.

Single memory architecture

Single memory architectures have one processor and memory. Its in this case memory size is big for large sized FFTs and hence results in slow processing, hence not used generally.

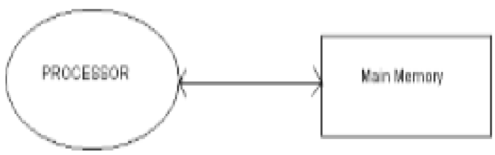


Fig.1 Single Memory Architecture

Dual Memory Architecture

Dual memory architectures are faster than the single memory architecture. The contrast circuitry is somewhat complex and also is costlier than single memory Architecture.

Fig2. Dual Memory Architecture

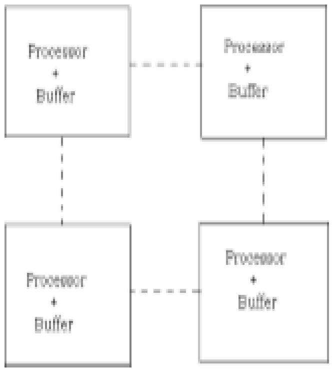


Fig.3 Array Architecture

This architecture is interconnected network results in the consumption of large chip area.

That is why these type of architecture are not commonly used.

Cache memory Architecture

Cache memory Architecture is faster than main memory and it also increases the effective bandwidth of memory overall speed of the processor is increased. But this type of architecture having disadvantage of increased controller complexity as well as makes the processor costly .

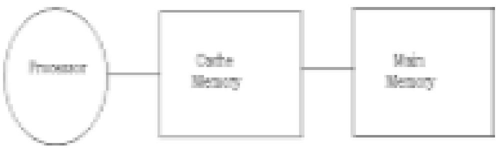


Fig.4 Cache Memory Architecture

Pipeline architecture offers a good tradeoff between hardware complexity and processing rate. It is very attractive in implementing the FFT processor for high speed multimedia communication system. Figure shown below is basic pipeline architecture for FFT processor.

It consist of number of butterfly computational units interspersed with delay commutator elements foe interstate data reordering with this architecture it is necessary to select the appropriate read for butterfly computation.

In our view the pipelined architecture should be best choice for high throughput rate application. Since it can provide high throughput rate with acceptable hardware cost. The pipelined FFT architecture typically falls one of two following categories. One is multipath delay commutator (MDC) and the other is Single path delay feedback (SDF) as shown in figure below.

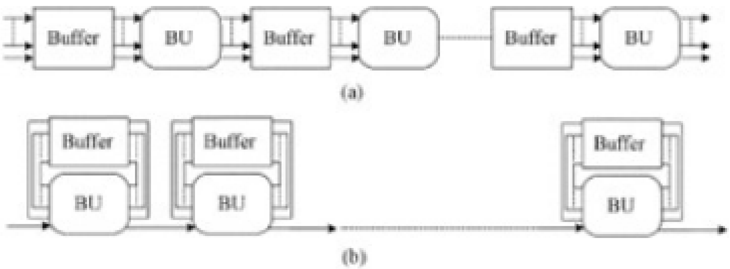


Fig.5 (a) MDC (b) SDF

In general, if appropriately reordered M parallel input data can be supported simultaneously in the MDC Scheme, this scheme provides M times throughput rate of SDF Scheme. However there are some limitations on the number of data path, FFT size and radix-2 FFT algorithm in MDC architecture.

Beside, the requirement of memory and complex multiplier in MDC scheme is more than that of SDF scheme. In general, the MDC scheme needs less memory and hardware cost.

For high throughput rate application, the MDC architecture is more suitable than SDF architecture in high frequency range application. If the input data are reordered in the input buffer they are loaded into the MDC processor. Unfortunately traditional R2 (Radix-2) MDC architecture cannot provide the available throughput rate unless it raise the work frequency. The R4 (Radix-4) MDC architecture, which needs a power of four, has the limitations on FFT size and split-radix (SRJ MDC) has higher hardware cost. In addition, the higher radix FFT algorithm is difficult to be implemented in the traditional MDC architecture.

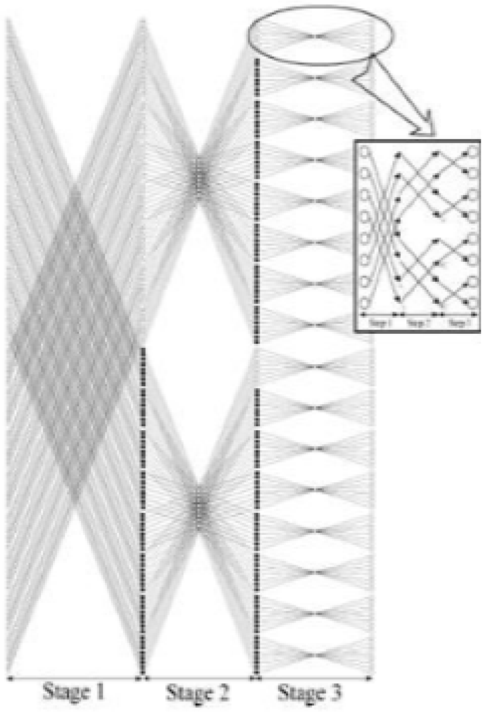


Fig.6 SFG of 128 point mix Radix FFT algorithm

In general, the higher throughput rate of FFT processor can be provided by increasing the number of data paths in MDC pipelines architecture. However the hardware cost is also significantly increased because more memory and complex multiplier are needed to allow multiple data to be operated simultaneously.

In this paper four data path pipelined FFT architecture which is called mixed radix multipath delay feedback (MRMDF) is studied which here combined the features of SDF and MDC architectures. This FFT/IIFT processor provides a throughput rate meet to ultra sideband specification. The MRMDF architecture has lower hardware cost compared with traditional MDC approach and adopts the high-radix FFT algorithm to save power dissipation.

Architecture of FFT for high throughput rate

The block diagram of proposed 128 point FFT/IIFT processor is derived from equation 1,2,3,4 and signal flow graph shown in fig.1 ..

$$\underbrace{\sum_{n_2=0}^{63}\left\{\underbrace{\sum_{n_1=0}^{1} x(64n_1+n_2)W_2^{n_1k_1}}_{\text{2 point DFT}}\ \underbrace{W_{128}^{n_2k_1}}_{\text{twiddle factor}}\right\}W_{64}^{n_2k_2}}_{\text{64 point DFT}}$$

$$\sum_{n_2=0}^{63} BU_2(k_1,n_2)W_{64}^{n_2k_2}. \qquad (1)$$

The operation of FFT/IIFT is controlled by control signal, as shown in fig7. . When the IIFT is performed in our processor , the sign of imaginary part of input sequence will be changed and then they will be performed by the process in treating FFT . The sign of imaginary part of output data from FFT will be changed again and then will be divided by 128point. Because 128 is a power of 2. The operation of division is implemented by shifting the decimal point location. The function of module 1 is to implement a radix 2 FFT algorithm corresponding to the first stage of signal flow graph. Module2 and module3 are to realize the radix8 FFT algorithm corresponding to second and third stage of signal flow graph. In order to minimize the memory requirement and to ensure the correction of the FFT output later.

Comparison of 128-point pipelined FFT architecture

	MRMDF	SRMDC*	R2MDC*	$R2^3SDF$
No. Of registers (complex words)	124(38.9%)	318(100%)	190(60%)	127(40%)
No. of complex multipliers	2+4*0.62(44.8%)	10(100%)	6(60%)	3(30%)
No. of complex adders	48(100%)	34(70.8%)	14(29%)	14(29%)
Algorithm	Radix-2 Radix-8	Split-radix	Radix-2	Radix-2 Radix-8
Input data format	Parallel(4 input port)	Parallel(4 input port)	Parallel(2 input port)	Serial
Output data format	Parallel(4 output port)	Parallel(6 output port)	Parallel(2 output port)	Serial
Throughput rate(R: clock rate)	4R	6R*0.73	2R	R

(1) No. of registers excluding the input buffers as listed.

(2) The desired throughput rate can be achieved if employing appropriately reordered parallel input data.

Table1: Comparison of 128 point pipeline FFT architecture

COMPARISON

In general the performance and hardware cost of pipeline architecture are increased by using the multiple data path approach . Thus the multiple path architecture usually provides the higher throughput rate with higher hardware cost if the parallel input data can be supported in this approach, the proposed MDRMDF architecture hardware cost in term of 128point FFT are as follows:

1. Register number 124
2. complex adder48
3. complex multiplier 2+4*0.62

Table shown above compares the hardware requirement FFT algorithm and throughput rate with several classical and the above approach in 128point FFT.

This equation can be considered as two- dimensional DFT . One is a 64 point DFT and other is a 2 point DFT .

$$X(2(\beta_1+2\beta_2+4\beta_3+8\beta_4)+k_1) = \sum_{\alpha_4=0}^{7} \mathrm{BU}_8(k_1,\beta_1,\beta_2,\beta_3,\alpha_4)W_8^{\alpha_4\beta_4} \quad (2)$$

In the above equation a complex multiplication with one of two cofficient can be computed using addition and a real multiplication whose hardware can be realized by 6 shifters and 4 adders .

$$X(n) = \frac{1}{N}\left\{\sum_{k=0}^{N-1} X^*(k)W^{nk}\right\}^* . \quad (3)$$

Where X(n) is the desried output sequence .

The proposed MRMDF architecture combining the feature of SDF and MDC architecture consist of module- 1, module-2 module-3, conjugate blocks, a division block and multiplexers. The features of proposed MDRDF architecture are the following:

1. Higher throughput rate can be provided by using 4 parallel data path
2. The minimum memory is required by using the delay feedback approach to reorder the input data and the intermediate results of each module.
3. The 128 point mixed radix FFT/IIFT algorithm is implemented to save power consumption.
4. The number of complex multiplier is minimized by using the scheduling scheme and the specified constant multipliers.

In the MRMDF architecture the input sequence and the output sequence are in specified order. The order of the output sequence is the bit reversal of the order of input sequence as in fig.7

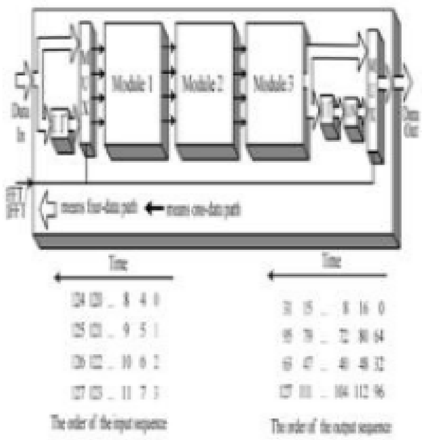


Fig.7 Block Diagram of 128 point FFT/IIFT processor

CONCLUSION

It is concluded that by MRMDF architecture high throughput rate can be achieved by using 4 data path . Furthermore the hardware costs of memory and complex multiplier can be saved by adopting delay feedback and data scheduling approach .In addition the number of complex multiplication is reduced effectively by using a higher radix algorithm.

REFERENCES

[1] Y.Jung, H.Yoon, and j.Kim,"New efficient FFT algorithm and pipeline implementation results for OFDM/DMT applications," IEEE Trans. Consum. Electron. vol.49,no.1,pp.14- 20,Feb.2003.

[2]. L.Jia, Y.Gao, J. Isoaho, and H.Tenhunen,"A new VLSI-oriented FFT algorithm and implement,"in Proc.11th Annu.IEEE Int. ASIC Conf.,Sep.1998, pp.337-341.

[3]. L.R. Rabiner and B.gold, Theory and application of Digital Signal Processing. Englewood cliffs, NJ:Prentice Hall, 1975.

[4]. S.Magar, S.Shen, G.Luikio, M.Fleming, and R.Aguilar,"an application specific DSP chip set for 100 MHz data rates,"in proc. Int. Conf. Acoustics, Speech, and Signal Processing, vol. 4, Apr.1988, pp.1989-1992.

[5]. Shousheng he and Mats Torkelson,"Designing pipeline FFT processor,"parallel Processing Symposium, pp.766-770, 1996.

[6]. J.Garcia, J.A.Michell, and A.M.Buron,"VLSI configurable delay commutator for a pipeline split radix FFT architecture,"IEEE Traans. Signal Processing, vol47, pp.3098-3107, Nov.1999.

[7]. L.R.Rabiner and B.Gold, Theory and Application of Digital Signal Processing. Englewood Cliffs,NJ:Prentice Hall,1975.

[8]. K.K. Parthi, VLSI digital Signal Processing Systems: Design and Implementation. New York: Wiley, 1999, ch.7, p.189.

[9]. E.H. Wold and A.M. Despain. Pipeline and parallel-pipeline FFt processors for VLSI implementation. IEEE Trans. Comput., C-33(5):414-426, May 1984.

Coding Theorems for the R-Norm Information Measure

[*]Dr Meenakshi Darolia , [**]Mr Vishesh Bansal

Assistant Professor, Isharjyot Degree College Pehova

Assistant Professor, MAIMT Jagadhri

In this paper, we will give coding theorems with respect to the R-Norm Information Measure. Suppose we have discrete memory less information source with an encoding alphabet with D Symbols and code words xi with word-lengths Ni, i = 1,2,….n, which fulfill the Kraft inequality

$$\left(\sum_{i=1}^{n} D^{-n_i}\right) \leq 1 \qquad (1.1)$$

Where D is the size of the code alphabet.

We next give a definition of average length LR of cord words.

DEFINITION: The average length LR with respect to R-norm information measure is for R€ R+ given by

$$L_R = \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right]$$

Clearly LR will increase for increasing word lengths. An important property of LR is that for R→1 it is equivalent with the average length of code words by Shannon, up to a constant.

Theorem: For all integer D > 1

$$L_R = \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right] > 0 \quad \text{for} \quad R \in R^{+}$$

Proof: To prove LR > 0, we consider the following cases:

Case I: when R > 1, then R – 1 > 0 and $\frac{R}{R-1} > 0$

From (1.1), we have $\left(\sum_{i=1}^{n} D^{-n_i}\right) \leq 1 \quad \Rightarrow D^{-n_i} < 1 \quad \Rightarrow D^{-n_i\left(\frac{R}{R-1}\right)} < 1$

Multiplying both sides of (1.3) by Pi , we get $\Rightarrow P_i D^{-n_i\left(\frac{R}{R-1}\right)} < P_i$

Summing over I = 1,2,3, ... ,N both sides, we get

$$\Rightarrow \sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R}{R-1}\right)} < \sum_{i=1}^{n} P_i = 1, \qquad \Rightarrow \sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R}{R-1}\right)} < 1$$

$$\Rightarrow -\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R}{R-1}\right)} > -1, \qquad \Rightarrow 1-\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R}{R-1}\right)} > -1+1=0$$

$$\Rightarrow 1-\sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R}{R-1}\right)} > 0 \qquad (1.4)$$

Multiplying (1.1) by $\frac{R}{R-1}$, , we get

$$\frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-n_i(R-1)/R}\right] > 0 \qquad (1.5)$$

But $\quad L_R = \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-N_i(R-1)/R}\right]$

Thus from (1.5), we get

$$L_R = \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-N_i(R-1)/R}\right] > 0 \quad \text{for } R > 1 \qquad (1.6)$$

Case II: when $0 < R < 1 \Rightarrow R-1 < 0$ and $\frac{R}{R-1} < 0$

From (1.1), we have

$$\left(\sum_{i=1}^{n} D^{-N_i}\right) \le 1, \quad \Rightarrow D^{-N_i} < 1, \quad \Rightarrow D^{-N_i\left(\frac{R}{R-1}\right)} > 1 \qquad (1.7)$$

Multiplying both sides of (1.7) by Pi , we get

$$\Rightarrow P_i D^{-N_i\left(\frac{R}{R-1}\right)} > P_i$$

Summing over i 1,2,3,. .... ,n both sides, we get

$$\Rightarrow \sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R}{R-1}\right)} > \sum_{i=1}^{n} P_i = 1$$

$$\Rightarrow \sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R}{R-1}\right)} > 1$$

$$\Rightarrow -\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R}{R-1}\right)} < -1$$

$$\Rightarrow 1-\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R}{R-1}\right)} < -1+1=0$$

$$\Rightarrow 1-\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R}{R-1}\right)} < 0 \qquad (1.8)$$

Multiplying (1.8) by $\frac{R}{R-1}$, we get

$$\frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R-1}{R}\right)}\right] > 0 \qquad (1.9)$$

But $L_R = \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R-1}{R}\right)}\right]$

Thus from (1.9) , we get

$$L_R = \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R-1}{R}\right)}\right] > 0 \quad \text{for } 0 < R < 1 \qquad (1.10)$$

Thus from (1.6) and (1.10), we get

$$L_R = \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R-1}{R}\right)}\right] > 0 \quad \text{for } R \in R^{+}$$

Theorem: If Ni , i=1,2,……..,n are in length of code words xi then

$$\lim_{R\to 1} L_R = \sum_{i=1}^{n} p_i N_i \log(D)$$

Proof: The average length LR with respect to R-norm information measure is for R€ R + given by

$$L_R = \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-N_i(R-1)/R}\right] \quad (1.11)$$

Taking limit both sides as R→1, we get

$$\lim_{R\to1} L_R = \lim_{R\to1} \frac{R}{R-1}\left[1-\sum_{i=1}^{n} P_i D^{-N_i(R-1)/R}\right] = \frac{0}{0}(form) \quad (1.12)$$

Thus by Bernoulli-L' Hospital theorem, we get

$$\lim_{R\to1} L_R = \lim_{R\to1}\left[1.\left[1-\sum_{i=1}^{n} p_i D^{-N_i(R-1)/R}\right] - R\left[0-\sum_{i=1}^{n} p_i \frac{dT}{dR}\right]\right] \quad (1.13)$$

Where $T = D^{-N_i(R-1)/R}$ (1.11)

Taking log both sides of (1.11), we get

$$\log T = \frac{-N_i(R-1)}{R}\log D \quad (1.15)$$

Diff w.r.t 'R' both sides of (1.15), we get

$$\frac{1}{T}\cdot\frac{dT}{dR} = -N_i \log(D)\left[\frac{R.1-1.(R-1)}{R^2}\right]$$

$$\frac{dT}{dR} = -TN_i\log(D)\left[\frac{1}{R^2}\right] = -\frac{1}{R^2}D^{-N_i(R-1)/R}[N_i\log(D)]. \quad (1.16)$$

Substitute (1.16) in (1.13), we get

$$\Rightarrow \lim_{R\to1} L_R = \lim_{R\to1}\left[1.\left[1-\sum_{i=1}^{n} p_i D^{-N_i(R-1)/R}\right] - \frac{R}{R^2}\left[\sum_{i=1}^{n} p_i D^{-N_i(R-1)/R} N_i \log(D)\right]\right]$$

$$\Rightarrow \lim_{R\to1} L_R = \lim_{R\to1}\left[\left[1-\sum_{i=1}^{n} p_i D^{-N_i(R-1)/R}\right] - \frac{1}{R}\left[\sum_{i=1}^{n} p_i D^{-N_i(R-1)/R} N_i \log(D)\right]\right]$$

$$\Rightarrow \lim_{R\to1} L_R = \left[\left[1-\sum_{i=1}^{n} p_i\right] - \left[\sum_{i=1}^{n} p_i \left[N_i\log(D)\right]\right]\right] = -\sum_{i=1}^{n} p_1 N_i \log(D) \quad (1.17)$$

Thus finally $\lim_{R\to1} L_R = -\sum p_1 N_i \log(D)$ (1.18)

Theorem: For all integer D > 1

$$H_R(P) \le L_R$$

Under the condition (1.1). Equality holds if and only if

$$N_i = -\log(P_i^R / \sum_{i=1}^{n} P_i^R)$$

Proof: To prove this theorem, we consider following cases:

Case I: when R > 1

We use Holder inequality [12]

$$\sum_{i=1}^{n} x_i y_i \geq \left(\sum_{i=1}^{n} x_i^{P}\right)^{\frac{1}{P}} \left(\sum_{i=1}^{n} y_i^{q}\right)^{\frac{1}{q}} \tag{1.19}$$

for all $x_i \geq 0, y_i \geq 0$, i = 1,2,....n when P <1(≠) and $p^{-1} + q^{-1} = 1$ .with equality if and only if there exists a positive number c such that

$$x_i^{p} = c y_i^{q} \quad \text{Setting} \quad x_i = P_i^{\frac{R}{R-1}} . D^{-N_i}$$

and $y_i = P_i^{\frac{R}{R-1}}$, P = $1 - \frac{1}{R}$, q = 1 R

$$\left(\sum_{i=1}^{n} P_i^{\frac{R}{R-1}} D^{-N_i} P_i^{\frac{R}{R-1}}\right) \geq \left(\sum_{i=1}^{n} (P_i^{\frac{R}{R-1}} D^{-N_i})^{1-\frac{1}{R}}\right)^{1/1-\frac{1}{R}} \left(\sum_{i=1}^{n} (P_i^{\frac{R}{R-1}})^{1R}\right)^{1/1-R} \tag{1.20}$$

$$\left(\sum_{i=1}^{n} D^{-N_i}\right) \geq \left(\sum_{i=1}^{n} P_i \quad (D^{N_i})^{\frac{-R}{R}}\right)^{1\frac{R}{R-1}} \left(\sum_{i=1}^{n} (P_i)^{R}\right)^{1/1-R} \tag{1.21}$$

Since $\left(\sum_{i=1}^{n} D^{-n_i}\right) \leq 1$ Thus (1.21) becomes

$$\left(\sum_{i=1}^{n} P_i \ (D^{-N_i})^{\frac{-R}{R}}\right)^{1\frac{R}{R-1}} \left(\sum_{i=1}^{n} (P_i)^{R}\right)^{1/1-R} \leq 1$$

$$\left(\sum_{i=1}^{n} P_i \ (D^{-N_i})^{\frac{-R}{R}}\right)^{1\frac{R}{R-1}} \leq \left(\sum_{i=1}^{n} (P_i)^{R}\right)^{1/R-1} \tag{1.22}$$

Raising power 1/R both sides of (1.22), we get

$$\left(\sum_{i=1}^{n} P_i \ (D^{-N_i})^{\frac{R-1}{R}}\right) \leq \left(\sum_{i=1}^{n} (P_i)^{R}\right)^{\frac{1}{R}}$$

$$\Rightarrow -\left(\sum_{i=1}^{n} P_i \ (D^{-N_i})^{\frac{R-1}{R}}\right) \geq -\left(\sum_{i=1}^{n} (P_i)^{R}\right)^{\frac{1}{R}}$$

$$\Rightarrow 1-\left(\sum_{i=1}^{n} P_i \ (D^{-N_i})^{\frac{R-1}{R}}\right) \geq 1-\left(\sum_{i=1}^{n} (P_i)^{R}\right)^{\frac{1}{R}} \tag{1.23}$$

We know $\frac{R}{R-1} > 0$ if R > 1

Multiplying $\frac{R}{R-1}$ by both side of (1.23) and we get

$$\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) \geq \frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i^{R}\right)^{\frac{1}{R}}\right) \tag{1.21}$$

But $\frac{R}{R-1}\left[1-\left(\sum_{i=1}^{n} P_i^{R}\right)^{\frac{1}{R}}\right] = H_R(P)$

and $\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) = L_R$

Thus from (1.21), $L_R \geq H_R(P)$ for R >1 (1.25)

Case II: when $0 < R < 1$

We use Holder inequality [12]

$$\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} x_i^{P}\right)^{\frac{1}{P}} \left(\sum_{i=1}^{n} y_i^{q}\right)^{\frac{1}{q}} \tag{1.26}$$

For all $x_i \geq 0, y_i \geq 0$, i=1,2,....N when P <1(≠) and $p^{-1}+q^{-1}=1$ with equality if and only if there exists a positive number c such that

$x_i^{P} = c y_i^{q}$ Setting $x_i = P_i^{\frac{R}{R-1}} . D^{-n_i}$

and $y_i = P_i^{\frac{R}{R-1}}$, $P = 1-\frac{1}{R}$, $q = 1-R$ Thus (1.27) becomes

$$\left(\sum_{i=1}^{n} P_i^{\frac{R}{R-1}} D^{-N_i} P_i^{\frac{R}{R-1}}\right) \leq \left(\sum_{i=1}^{n} (P_i^{\frac{R}{R-1}} D^{-N_i})^{1-\frac{1}{R}}\right)^{1/1-\frac{1}{R}} \left(\sum_{i=1}^{n} (P_i^{\frac{R}{R-1}})^{1-R}\right)^{1/1-R}$$

$$\left(\sum_{i=1}^{n} D^{-N_i}\right) \leq \left(\sum_{i=1}^{n} P_i (D^{N_i})^{\frac{-R-1}{R}}\right)^{\frac{R}{R-1}} \left(\sum_{i=1}^{n} (P_i)^{R}\right)^{1/1-R} \tag{1.27}$$

Since $\left(\sum_{i=1}^{n} D^{-n_i}\right) \leq 1$ Thus (1.27) becomes

$$\left(\sum_{i=1}^{n} P_i (D^{-N_i})^{\frac{-R-1}{R}}\right)^{\frac{R}{R-1}} \left(\sum_{i=1}^{n} (P_i)^{R}\right)^{1/1-R} \geq 1$$

$$\Rightarrow \left(\sum_{i=1}^{n} P_i (D^{-N_i})^{\frac{-R-1}{R}}\right)^{\frac{R}{R-1}} \geq \left(\sum_{i=1}^{n} (P_i)^{R}\right)^{1/R-1} \tag{1.28}$$

Raising power 1/R both sides of (1.28), we have

$$\left(\sum_{i=1}^{n} P_i \,(D^{-N_i})^{\frac{R-1}{R}}\right) \ge \left(\sum_{i=1}^{n}(P_i)^R\right)^{\frac{1}{R}}$$

$$\Rightarrow -\left(\sum_{i=1}^{n} P_i \,(D^{-N_i})^{\frac{R-1}{R}}\right) \le -\left(\sum_{i=1}^{n}(P_i)^R\right)^{\frac{1}{R}}$$

$$\Rightarrow 1-\left(\sum_{i=1}^{n} P_i \;(D^{N_i})^{\frac{R-1}{R}}\right) \le 1-\left(\sum_{i=1}^{n}(P_i)^R\right)^{\frac{1}{R}} \qquad (1.29)$$

We know $\frac{R}{R-1}$ < 0 if 0 < R < 1

Multiplying $\frac{R}{R-1}$ by both sides (1.29), we get

$$\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i D^{-N_i\left(\frac{R-1}{R}\right)}\right)\right) \ge \frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i^R\right)^{\frac{1}{R}}\right) \qquad (1.30)$$

But $$\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i^R\right)^{\frac{1}{R}}\right) = H_R(P)$$

and $$\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) = L_R \qquad (1.31)$$

Thus from (1.30),we get $L_R \ge H_R(P)$ 0 < R < 1 (1.32)

Thus from (1.25) and (1.32), we get

$L_R \ge H_R(P)$ for R€ R⁺ (1.33)

Theorem: For every code with length $\{n_i\}, i = 1,2,3,\ldots\ldots, N, and\ L_R$ made to satisfy

$$L_R < H_R(P).D^{\left(\frac{R-1}{R}\right)} + \frac{R}{R-1}\left[1 - D^{\left(\frac{R-1}{R}\right)}\right] \qquad \text{for R € R}^+$$

Proof: To prove this theorem we consider the following

cases: Case I: when R > 1

Let ni be the positive integer satisfying the inequality by (11)

$$-\log\left(\frac{p_i^R}{\sum p_i^R}\right) \le n_i < -\log\left(\frac{p_i^R}{\sum p_i^R}\right) + 1$$

Consider the interval

$\delta_i = \left[-\log\left(\frac{p_i^R}{\sum p_i^R}\right), -\log\left(\frac{p_i^R}{\sum p_i^R}\right) + 1\right]$ of length 1 .In every δ_i ,there lies exactly one positive number ni , such that

$$0 < -\log\left(\frac{p_i^R}{\sum p_i^R}\right) \le n_i < -\log\left(\frac{p_i^R}{\sum p_i^R}\right) + 1 \qquad (1.31)$$

It can be shown that the sequence $\{n_i\}, i = 1,2,3,......, N$ thus defined, satisfies (1.1).

Thus from (1.31), we get

$$n_i < -\log\left(\frac{p_i^R}{\sum p_i^R}\right) + 1$$

$$-\log D^{-n_i} < -\log\left[\left(\frac{p_i^R}{\sum p_i^R}\right) D^{-1}\right]$$

$$D^{-n_i} > \left[\frac{p_i^R}{\sum p_i^R}\right] D^{-1}$$

Raising above inequality by $\frac{R-1}{R}$ both sides, we get

$$D^{-n_i\left(\frac{R-1}{R}\right)} > \left(\frac{p_i^R}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)} \qquad (1.35)$$

Multiplying both sides of (1.35) by pi , we get

$$P_i D^{-n_i\left(\frac{R-1}{R}\right)} > P_i\left(\frac{p_i^R}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow P_i D^{-n_i\left(\frac{R-1}{R}\right)} > P_i .(P_i)^{R-1}\left(\frac{1}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow P_i D^{-n_i\left(\frac{R-1}{R}\right)} > (P_i)^{1-1+R}\left(\frac{1}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow P_i D^{-n_i\left(\frac{R-1}{R}\right)} > (P_i)^{R}\left(\frac{1}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)} \qquad (1.36)$$

Summing over i =1,2,3, ... ,N both sides of (1.36), we get

$$\Rightarrow \sum P_i . D^{-n_i\left(\frac{R-1}{R}\right)} > \left(\sum P_i^{R}\right)\left(\frac{1}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow \sum P_i . D^{-n_i\left(\frac{R-1}{R}\right)} > .\left(\sum P_i^{R}\right)^{1-1+\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow \sum P_i . D^{-n_i\left(\frac{R-1}{R}\right)} > .\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow -\sum P_i . D^{-n_i\left(\frac{R-1}{R}\right)} < -.\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow 1-\sum P_i . D^{-n_i\left(\frac{R-1}{R}\right)} < 1-.\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)} \qquad (1.37)$$

We know $\frac{R}{R-1} > 0$ if R > 1

Multiplying $\frac{R}{R-1}$ by both sides of (1.37), we get

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}\left[1-\left(\sum P^R\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}\right]$$

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}-\frac{R}{R-1}\left[\left(\sum P^R\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}\right]$$

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}-\frac{R}{R-1}\left[\left(\sum P^R\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}\right]$$

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}-\frac{R}{R-1}\left[1-1+\left(\sum P^R\right)^{\frac{1}{R}}\right] D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}\left[1-D^{\left(\frac{1-R}{R}\right)}\right]+\frac{R}{R-1}\left[1-\left(\sum P_i^R\right)^{\frac{1}{R}}\right] D^{\left(\frac{1-R}{R}\right)}$$

(1.38)

But $\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{N} P_i^R\right)^{\frac{1}{R}}\right) = H_R(P)$ And $\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{N} P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right) = L_R$

Thus (1.38) becomes $L_R < H_R(P).D^{\left(\frac{R-1}{R}\right)} + \frac{R}{R-1}\left[1-D^{\left(\frac{R-1}{R}\right)}\right]$ for R >1 (1.10)

Cases II: when 0 < R < 1 Let n_i be the positive integer satisfying the inequality

$$-\log\left(\frac{p_i^R}{\sum p_i^R}\right) \le n_i < -\log\left(\frac{p_i^R}{\sum p_i^R}\right) + 1$$ Consider the interval

$$\delta_i = \left[-\log\left(\frac{p_i^R}{\sum p_i^R}\right), -\log\left(\frac{p_i^R}{\sum p_i^R}\right) + 1\right]$$ of length 1 .In every δ_i, there

lies one positive number n_i , such that

$$0 < -\log\left(\frac{p_i^R}{\sum p_i^R}\right) \le n_i < -\log\left(\frac{p_i^R}{\sum p_i^R}\right) + 1 \qquad (1.11)$$

It can be shown that the sequence n_i , i 1,2,3,......., N thus defined, satisfies (1.1).

Thus from (1.11), we get

$$n_i < -\log\left(\frac{p_i^R}{\sum p_i^R}\right) + 1$$

$$-\log D^{-n_i} < -\log\left[\left(\frac{p_i^R}{\sum p_i^R}\right)D^{-1}\right]$$

$$D^{-n_i} > \left[\frac{p_i^R}{\sum p_i^R}\right]D^{-1}$$

Raising above inequality by $\frac{R-1}{R}$, we get

$$D^{-n_i\left(\frac{R-1}{R}\right)} < \left(\frac{p_i^R}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)} \qquad (1.12)$$

Multiplying both sides of (1.12) by p_i, we get

$$P_i D^{-n_i\left(\frac{R-1}{R}\right)} < P_i\left(\frac{p_i^R}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow P_i D^{-n_i\left(\frac{R-1}{R}\right)} < P_i.(P_i)^{R-1}\left(\frac{1}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow P_i D^{-n_i\left(\frac{R-1}{R}\right)} < (P_i)^{1-1+R}\left(\frac{1}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow P_i D^{-n_i\left(\frac{R-1}{R}\right)} < .(P_i)^{R}\left(\frac{1}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)} \qquad (1.13)$$

Summing over i 1,2,3, ... ,N both sides, we get

$$\Rightarrow \sum P_i.D^{-n_i\left(\frac{R-1}{R}\right)} < \left(\sum P_i^{R}\right)\left(\frac{1}{\sum p_i^R}\right)^{\frac{R-1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow \sum P_i.D^{-n_i\left(\frac{R-1}{R}\right)} < .\left(\sum P_i^{R}\right)^{1-1+\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow \sum P_i.D^{-n_i\left(\frac{R-1}{R}\right)} < .\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow -\sum P_i.D^{-n_i\left(\frac{R-1}{R}\right)} > -.\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow 1-\sum P_i.D^{-n_i\left(\frac{R-1}{R}\right)} > 1-.\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)} \qquad (1.11)$$

We know $\frac{R}{R-1} < 0$ if 0 < R < 1

Multiplying $\frac{R}{R-1}$ by both sides of (1.11), we get

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i.D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}\left[1-\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}\right]$$

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i.D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}-\frac{R}{R-1}\left[\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}\right]$$

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i.D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}-\frac{R}{R-1}\left[\left(\sum P_i^{R}\right)^{\frac{1}{R}} D^{\left(\frac{1-R}{R}\right)}\right]$$

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i.D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}-\frac{R}{R-1}\left[1-1+\left(\sum P_i^{R}\right)^{\frac{1}{R}}\right] D^{\left(\frac{1-R}{R}\right)}$$

$$\Rightarrow \frac{R}{R-1}\left[1-\left(\sum P_i.D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right] < \frac{R}{R-1}\left[1-D^{\left(\frac{1-R}{R}\right)}\right]+\frac{R}{R-1}\left[1-\left(\sum P_i^{R}\right)^{\frac{1}{R}}\right] D^{\left(\frac{1-R}{R}\right)} \qquad (1.15)$$

But $$\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{N} P_i^{R}\right)^{\frac{1}{R}}\right)=H_R(P)$$

And $$\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{N} P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right)=L_R$$

Thus (1.15) becomes

$$L_R < H_R(P).D^{\left(\frac{R-1}{R}\right)} + \frac{R}{R-1}\left[1 - D^{\left(\frac{R-1}{R}\right)}\right] \quad \text{for } 0 < R < 1 \qquad (1.16)$$

Thus from (1.10) and (1.16), we get

$$L_R < H_R(P).D^{\left(\frac{R-1}{R}\right)} + \frac{R}{R-1}\left[1 - D^{\left(\frac{R-1}{R}\right)}\right] \quad \text{for } R \in R^+ \qquad (1.17)$$

Theorem: For all integer D > 1

$$\sum P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} = \left(\sum P_i^R\right)^{\frac{1}{R}} \quad \text{for } R \in R^+ \qquad (1.18)$$

where $n_i^- = -\log_D\left(p_i^R \Big/ \sum p_i^R\right)$

Proof: Since $\bar{n}_i = -\log_D\left(p_i^R \Big/ \sum p_i^R\right)$ (1.19)

It can be written as

$$-\log_D D^{-\bar{n}} = -\log_D\left(p_i^R \Big/ \sum p_i^R\right)$$

$$D^{-\bar{n}_i} = \left(p_i^R \Big/ \sum p_i^R\right) \qquad (1.50)$$

Raising power $\frac{R-1}{R}$ both sides of (1.50), we get

$$D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} = \left(p_i^R \Big/ \sum p_i^R\right)^{\left(\frac{R-1}{R}\right)}$$

$$D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} = \left(p_i^{R-1} \Big/ \left(\sum P_i^R\right)^{\left(\frac{R-1}{R}\right)}\right) \qquad (1.51)$$

Multiplying both sides of (1.51) by Pi, we get

$$P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} = \left(p_i^R \Big/ \left(\sum P_i^R\right)^{\left(\frac{R-1}{R}\right)}\right) \qquad (1.52)$$

Summing over i = 1,2,3, ... ,N both sides of (1.52), we get

$$\sum P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} = \left(\sum P_i^R \Big/ \left(\sum P_i^R\right)^{\left(\frac{R-1}{R}\right)} \right)$$

$$= \left(\sum P_i^R\right)^{\frac{1}{R}} \qquad (1.53)$$

Thus from (1.53), we get

$$\sum P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} = \left(\sum P_i^R\right)^{\frac{1}{R}} \quad \text{for } R \in R^+ \qquad (1.51)$$

Theorem: For every code length $\{n_i; i\} = 1,2,3,............., N.,$ L_R can be made to satisfy

$$L_R > H_R(P) + \frac{R}{R-1}\left(1 - D\right) \qquad (1.55)$$

Proof: To prove this theorem we consider the following cases:

Cases I: when R > 1

$$\text{Suppose} \qquad n_i = -\log\left(p_i^R \Big/ \sum p_i^R \right) \qquad (1.56)$$

Clearly $\bar{n}_i$ and $\bar{n}_i + 1$ satisfy the 'equality' in Holder's Inequality[12]. Moreover. $\bar{n}_i$ satisfies Kraft's Inequality (1.1)

Suppose n_i is the unique integer between $\bar{n}_i$ and $\bar{n}_i + 1$, and then obviously, n_i satisfies Kraft's Inequality (1.1), we have

$\bar{n}_i \le n_i < \bar{n}_i + 1$ Now consider $\bar{n}_i \le n_i$ It can be written as

$$-\log D^{\bar{n}_i} \le -\log D^{n_i}, \qquad D^{-n_i} \le D^{-\bar{n}_i} \qquad (1.57)$$

Raising power $\frac{R}{R-1}$ both sides of (1.57), we get

$$D^{-n_i\left(\frac{R-1}{R}\right)} \le D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} \qquad (1.58)$$

Multiply by Pi both sides of (1.58), we get

$$P_i D^{-n_i\left(\frac{R-1}{R}\right)} \le P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} < D P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} \qquad (1.59)$$

Summing over i = 1,2,3, ... ,N both sides of (1.59), we get

$$\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)} < D\sum P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} \qquad (1.60)$$

Using (1.51) in (1.60), we get

$$\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)} < D\left(\sum P_i^R\right)^{\frac{1}{R}}$$

$$\Rightarrow -\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)} > -\left(\sum P_i^R\right)^{\frac{1}{R}}.D$$

$$\Rightarrow 1-\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)} > 1-\left(\sum P_i^R\right)^{\frac{1}{R}}.D \qquad (1.61)$$

We know $\frac{R}{R-1} > 0$ if R > 1

Multiplying both sides of (1.61) by $\frac{R}{R-1}$, we get

$$\frac{R}{R-1}\left(1-\left(\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) > \frac{R}{R-1}\left(1-\left(\sum P_i^R\right)^{\frac{1}{R}}.D\right)$$

$$\Rightarrow \frac{R}{R-1}\left(1-\left(\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) > \frac{R}{R-1} - \frac{R}{R-1}\left(\sum P_i^R\right)^{\frac{1}{R}}.D$$

$$\Rightarrow \frac{R}{R-1}\left(1-\left(\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) > \frac{R}{R-1} - \frac{R}{R-1}\left(1-1+\sum P_i^R\right)^{\frac{1}{R}}.D$$

$$\Rightarrow \frac{R}{R-1}\left(1-\left(\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) > \frac{R}{R-1}(1-D) + \frac{R}{R-1}\left(1-\sum P_i^R\right)^{\frac{1}{R}}D \qquad (1.62)$$

But $\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i^R\right)^{\frac{1}{R}}\right) = H_R(P)$

And $\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) = L_R$ Thus (1.63) becomes

$$L_R > H_R(P) + \frac{R}{R-1}(1-D) \quad \text{for } R > 1 \qquad (1.63)$$

Cases II: when 0 < R < 1 Suppose $n_i = -\log\left(p_i^R \Big/ \sum p_i^R\right)$ (1.61)

Clearly $\bar{n}_i$ and $\bar{n}_i + 1$ satisfy the 'equality' in Holder's Inequality *12+. Moreover. $\bar{n}_i$ Satisfies Kraft's inequality (1.1). Suppose n_i is the unique integer b n_i and $\bar{n}_i + 1$, and then obviously, n_i satisfies (1.1), we hav $\bar{n}_i \le n_i < \bar{n}_i + 1$ low consider $\bar{n}_i \le n_i$, It can be written as

$$-\log D^{-\bar{n}_i} \le -\log D^{n_i} \quad , \quad D^{-n_i} \le D^{-\bar{n}_i} \qquad (1.65)$$

Raising power $\frac{R}{R-1}$ both sides of , we get, $D^{-n_i\left(\frac{R-1}{R}\right)} \ge D^{-\bar{n}_i\left(\frac{R-1}{R}\right)}$

Multiplying both sides of (1.66) by Pi ,we get

$$P_i D^{-n_i\left(\frac{R-1}{R}\right)} \ge P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} > D P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} \qquad (1.66)$$

Summing over i = 1,2,3, ... ,N both sides of (1.66), we get

$$\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)} > D \sum P_i D^{-\bar{n}_i\left(\frac{R-1}{R}\right)} \qquad (1.67)$$

Using (1.51) in (1.67), we get

$$\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)} > D\left(\sum P_i^R\right)^{\frac{1}{R}}$$

$$\Rightarrow -\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)} < -\left(\sum P_i^R\right)^{\frac{1}{R}}.D$$

$$\Rightarrow 1-\sum P_i D^{-n_i\left(\frac{R-1}{R}\right)} < 1-\left(\sum P_i^R\right)^{\frac{1}{R}}.D \qquad (1.68)$$

We know $\frac{R}{R-1} < 0$ if $0 < R < 1$

Multiplying both sides of (1.68) $\frac{R}{R-1}$, we get

$$\frac{R}{R-1}\left(1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right) > \frac{R}{R-1}\left(1-\left(\sum P^R\right)^{\frac{1}{R}}.D\right)$$

$$\Rightarrow \frac{R}{R-1}\left(1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right) > \frac{R}{R-1} - \frac{R}{R-1}\left(\sum P^R\right)^{\frac{1}{R}}.D$$

$$\Rightarrow \frac{R}{R-1}\left(1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right) > \frac{R}{R-1} - \frac{R}{R-1}\left(1 - 1 + \sum P^R\right)^{\frac{1}{R}}.D \qquad (1.69)$$

$$\Rightarrow \frac{R}{R-1}\left(1-\left(\sum P_i D^{-n_i\left(\frac{R}{R-1}\right)}\right)\right) > \frac{R}{R-1}\left(1-D\right) + \frac{R}{R-1}\left(1-\sum P^R\right)^{\frac{1}{R}}.D$$

But $\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i^{R}\right)^{\frac{1}{R}}\right) = H_R(P)$

And $\frac{R}{R-1}\left(1-\left(\sum_{i=1}^{n} P_i D^{-n_i\left(\frac{R-1}{R}\right)}\right)\right) = L_R$

Thus (1.69) becomes

$$L_R > H_R(P) + \frac{R}{R-1}(1-D) \quad \text{for } 0 < R < 1 \qquad (1.70)$$

Thus form (1.63) and (1.70), we get

$$L_R > H_R(P) + \frac{R}{R-1}(1-D) \quad \text{for } R \in R^{+} \qquad (1.71)$$

HENCE PROVED

REFRENCES:

[9]. HARDY, G. H., LITTLEWOOD, J. E., AND POLYA, G.(1973), "Inequalities Cambridge Univ. Press, London /New York

[10]. HAVDRA, J. AND CHARVAT, F. (1967). Quantification method of Classification processes, Concept of structural α- entropy, Kybernetika 3, 30-35.

**11+. Nath, P.(1975), On a Coding Theorem Connected with Renyi's Entropy Information and Control 29, 234-242.*

[12]. O.SHISHA, Inequalities,Academic Press, New York.

[13]. RENVI, A. (1961),On measure of entropy and information, in proceeding, Fourth Berkeley Symp. Math. Statist. Probability No-1, pp. 547-561.

**14+. R. P. Singh (2008), Some Coding Theorem for weighted Entropy of order α.. Journal of Pure and Applied Mathematics Science : New Delhi.*

[15]. SHANNON, C. E. (1948), A mathematical theoery of communication, Bell System Tech. J. 27, 379-423, 623-656.

[16]. TROUBORST , P. M., BACKER, E., BOEKKE, D. E., AND BOXMA, Y.(1974) , New families of probabilistic distance measure in "Proceeding Second Internal. Joint Conf. on Pattern Recognition, Copenhagen Denmark," pp. 3-5.

[17]. VAN DER LUBBE, J. C. A. (1977). "R- norm Information and A General Class of Measure of Certainty and Information," M.Sc. Thesis, Department of Electrical Engineering, Delft University of Technology, Delft, the Netherlands (in Dutch).

[18]. VAN DER LUBBE, J. C. A. AND BOEKKE, D. E. (1978), "Coding Theorems for Two Classes of Information Measure," Report Number IT-78-06, Delft University of Technology, Dept. E. E.

Providing Query Suggestions and Ranking for user Search History

U. Anvesh Babu, D. Khadar Hussain, C. Nagarjuna,
M.Tech Student,CSE Dept, JUTUA College of Engg, Anantapuramu, A.P.
M.Tech Student,CSE Dept, JNTUA College of Engg, Anantapuramu, A.P.
M.Tech Student,CSE Dept, JNTUA College of Engg, Anantapuramu, A.P.

A B S T R A C T

The investigation overview described focuses on the blueprint exploration record displays to support information seeking. Customers are gradually more pursuing sophisticated task-oriented aims in the net. Such as building travel plans, running funds or purchase plans. Searchers produce and use exterior reports of actions and consequent outcomes by using copy and paste capabilities, writing/ typing notes, and making printouts. The superior helps users within their extended period information quests around the net, web searchers keep tabs on their query and click on looking on-line. Within this paper, the trouble of managing a user's history inquiries in to groups in a dynamic and automated manner. In the case of different search- engines, it can be identify the query group automatically programs and components. That is query alterations, result positions, query suggestions and two-way search experimentally analyze the presentation of view their possible, practically joined goals.

INTRODUCTION

The rising no of printed electronic materials, the on-line grows into the huge recourses, the persons to get information, problem solutions and task completions using net tips. While the range and profusion of information the network grows, such that the convolution of tasks along with the selections, by providing simple navigational queries, it reduces the scope of users contents. Searchers create external memory helps to help keep an eye on improvement, plan steps, and gather information. Users are typically unwilling to expressly supply their properties due to the need of extra man effort required. Present research has concentrated on an automatic understanding of a user's priorities from the customer search database and is based on user choices, customized systems have developed.

One of information-seeking [1] tasks often performed by pupils is information gathering, which is extracting, evaluating and organizing related information for a given issue. The empowering services and properties are used by users, especially in the case of looking for complicated queries and on-line. Which recognize their capacity and connected queries combine in a group wise. Presently, some of the mashies related to the search process have introduced a new “Search record” quality. It allows auser to investigate their query in on-line by using recorded queries. Instead of tracking and maintaining the queries and the clicks in their own search history [2] better to identify the groups which are related to the given queries. This query grouping process makes the search machine to better understand. After the identification of query groups, search engines could get the best representation of search context. It also supports present query using clicks and queries within the related query group. It gives assurance to enhancing the quality of crucial elements in search engines, for example, if you take the current query “monetary assertion” related to “bank of Baroda”. Now the search engine improve the rank of page by supplying information about bank of Baroda declaration rather than monetary statement in the Wikipedia article or web pageassociated with monetarydeclarations in other banks.

This system introduced an automatic and dynamic method to arrange given customers search account into a no of query clusters. Each group is a compilation of queries by the exact customer the pertinent to each other about a general informational demand. The dynamically updating the grouping up of queries, while new queries are issued by the user, and new query groups could produce extra time. The profiling strategies of current clicks can be divided on file based and concept based techniques, by using the document based profiling tactics strive to analyze the performance of the customer's documents.

The search history broadly classified in to two categories. Like, short-term and long-term search history. The short-term search record is restricted to time duration of one search, it includes successive searchers get a logical data demand and takes with in the span of time period. Several times a user views the returned records, composes an original query, then the query modifications is not satisfied, until the research process repeats again. The above procedure to the search history throws the demanded information and get it useful search context. Long-term search history [3] includes all activities of recent, past and could is on the other hand, endless in time scope, by comparing the short-term research background, has more benefits. There isn't any need to detect session boundaries is often difficult to undertaking an arranging the query clusters within a customer's history is difficult for several reasons. First, the connected queries might not appear near each other, the search takes may be few days or even weeks. It also discovers the recent records is often considerably more useful than distant history, the overall user's history is useful to improve the accurate research of revenant queries.

The rest of data is structured as below: session 2 discourses the works. To catalogue the present user profile schemes into two classes and review the process to classes. Session 3, the personalization of our concept-based grouping method to control the relationship between uncertain queries based on the customer theoretical preference recorded with in the concept based user profile. Session 4, by using the user profiling strategies based on the concept of planning, by relating our describing schemes are present. The Session 5 conclusion of this paper.

RELATED WORK

This research takes about information retrieval; our goal would be to mechanically organize a user's database related to search into various query clusters. Each group consist single or other queries and their connected snap. Each cluster related to a unique data requires which could demand a little amount of clicks and queries associated with exactly the same search target. Let us consider an instance in the case of directional query, a cluster might consist as low as queries and clicks available. They highlight the value of planning, and outside trouble representation, and assessment in solving the problems, which is supported by research histories. Background displays have to include analytic queries and hypertext bounding in complete text techniques. Direct representation of the searchers path via a hypertext system can reduce disorientation.

User's record priorities are first retrieved in the user click process during data, and then it is used to study every performance model is basically characterized as a group of weighted structures. In the flip side, theory-based user reporting procedures aim at store user's theoretical demands. Based on user's priorities, it can create user priorities on the extracted categories.

Information Gathering [1] is a knowledge building procedure. Web learners start this process with recognizing anomalous state-of information linked to a subject. This state is the interest or concern mental state that activities the data gathering process. Ergo they make a preliminary search plan based on the prior knowledge. With every piece of new and valuable information encountered given them new ideas on the theme, they so extend or evolve their strategy to other related issues / subtopics or created the piece of information using their knowledge structure. In the end, the procedure is wound up with resolving the state. Information gathering is a very complicated information- seeking job. Students are often required to preserve many extracted results for later use and reference. However, to maintain a huge amount of information in a human's mind is troublesome since the restriction of working memory. To support the restriction of memory capability, students need to use outside memory support.

Either the information seeking methods derived some type of background techniques. They typically it includes the display of “Query Result Set” pairs. To take one example, Back in (1976) incorporated research review functions in his TIRES apparatus, the managing information retrieval structure, founded in four prior studies and techniques. Several early commercial systems had a history feature that legalized users to remember earlier search guidelines and reclaim them. This related work highlighted the importance of user boundaries to showed what type of measures have used earlier and mentioned what types of strategies (either short-term and long-term) had been followed. It also used annotation tools for customers to give feedback on the discovered tips and actions. It concludes that for observation needs the search history within the boundary of data seek and imagining a system and also stated that function are not support for the present system. Few new techniques or ideas are introduced to define and compute them. Twiddle and Nichols on 1998 introduced a toll called Ariadne, used to support collaboration between customers to visualize search session history. The system generates query results as pairs and represents them to users with thumbnails of screen shorts. Searchers share and use these histories with other users. This article reports information will be an effect on the accessing area, retrieval of needed information and in order to support the user, it suggests tools for search histories. These issues are related to co-ordination of information. Students have to frequently change alteration among them, to co-ordinate information kept in three kinds of memory aids. The frequently changed focus to get pupils easily disoriented. To locate and remember a bit of information which is previously kept in these memory aids becomes hard.

The list of ordered queries, qi, collectively with the equal set of clicked URLs, clk_i of q_i is known as the query grouping [4]. A query grouping is denoted as s= hq_1, clk_1… q_k, clk_{ki}.

Let us consider an array which consists a user query groups denoted by s consists more query groups. i.e., S= s1, s2… Sn, and their current query and its related links. Let us take one query group it is one of the current query clusters in S and is mostly connected to a latest query groups with the same queries. Suppose if they don't exist in S and is not adequately connected to query click (qc), and clicks (clkc). For this reason, to introduce one formula which defines dynamic in nature and gives some suggestion related to them. Also states that instead of proposing relevance measure based on the signal it uses time or text from search logs.

One method to identify the query ina user's search history, and query group, would first treated as each query inrecord as a singleton query cluster, after blend this singleton query group, n an iterative manner (in a k-means or collective way [7]). Nonetheless, the unreasonable with on their situation for two causes. Firstly, the process of unwanted outcomes of altering a customer's present query clusters, possibly undoing an individual's own physical efforts in arranging her record. Second, it needs more computational cost, because in every query, it can identify more no of duplicate query groups based on its similarity.

QUERY RELEVANCE USING SEARCH LOGS

The mechanism of web search logs [5] is used to explain the relevance query groups. Based on queries, our metrics capturing two important asserts. They are: First, the queries that are often performed and organized as reformulations and second, the queries can be carried out without any delay. Whenever the customer click on the same set of pages it can introduce cardinal search behavioral graphs, that uses the previously mentioned qualities following that, the graphs are useful to find query relevance and how a user's query to be able to improve our relevance metrics.

One way to classify important queries, it uses query re-formulations which are basically taken from the search query log engines. If two queries are issued to many users, consequently, more likely it uses re-formulation of one with another. For the above case, to assess the relevance among two queries it uses the metrics called time-based metric. That is, it provides some span of time for each query taken from consumers search history. A new strategy is used to provide related information about the given queries from our search logs, and it would be considers in such a way that a user will probably get related information often they click on same URLs.

Let us take an example, the queries about “iPod” and ”apple store” which don't explore text (or) its related information from the user's research history. But somewhat this information is related because it uses triggered click regarding the “iPod” artifact. In order to satisfy the properties, to develop a chart is known as query click graph. The query click graph (QCG) as well as query reformulation graph (QRG) provides two important properties for useful queries. It can combine these two charts keep on a single graph named query fusion graph (QFG) and in order to make these properties has more efficient. The relevant graph contains query click information from QCG and query reformulation sequence taken from QRG. QFG= (VQ, EQF), that submit to the query fusion graph. At a upper level, EQF enclose the no of limitssurvive in moreover EQR (or) EQC. The weight of the edge (qi, qj) in QFG, wf (qi, qj), is in used to the weighted sum of linear edges, wr (qi, qj) in EQR and wc (qi, qj) in EQC as follows.

Wf (qi, qj) = -x wr (qi, qj) + (1- α) x wc (qi, qj) algorithm [4] for scheming the query significance by replicating unsystematic walk across the query fusion graph. Relevance (q).

Input:
- Query fusion graph, QFG
- Jump vector, g
- Damping factor, d
- Total number of random walks, numRWs
- Size if neighbourhood, maxHops
- Given query, q

Output: The fusion relevance vector is q, relFq
- Initialize relFq=0
- NumWalks = 0, numVisits = 0
- While NumWalks <numRWs
- numHops = 0; v = q
- while v 6= NULL numHops < maxHops
- numHops++
- relF q (v)++; numVisits++
- v = SelwctNodeToVisit(v) NumWalks++
- for each v, normalize relF q (v) = relF, q (v)/numVisits

Above procedure to calculate the query significance by simulateun systematic walk across the query fusion graph.

By using jump vector (g) for queries and choose the unsystematic walk information. Then every outgoing edge, (v, qi), is picked with possibilitywf (v, qi), and the random walk always restarts v don't have any outgoing edges. (7) Of the algorithm for each query submission, the user defines not only included query re-formulation, but also it contains clicks in the URLs. The clicks of the user, further useful in the case of identifying the queries and query groups in an effective manner. In this paper presents a motivated example which illustrates why it is useful to compute query relevance to consider into clicked URLs of given query. Why don't consider the user's submitted a query “jaguar”. This occurs it don't understand the genuine search instead of users present issuing query “jaguar”. But all of us understanding the clicked URLs through the present customer following the question “jaguar”, according to the delegate query relevance scores and present query to behind issuing search interest to queries VQ. In this way the utilization of clicks are able to given a much superior query significance score to connected query to “animal jaguar” than linked to “auto jaguar”.

QUERY GROUPING USING THE QFG

In this paragraph, it introduces the similarity function simrel, is used in the online query group procedure outline. Their representation of relevance of one query to another query to maintain a query image, end each query group to kept context vector , to aggregator the picture of its own member of the query to form an entire representation. In our proposed representation, the crucial elements are content vector, query image, and, query relevance vector, to identify the relevance between query group to take notes on markov chain rules [6].

Context Vector: The content vector of a query group is represented is cxts, <s, the query vector (VQ) of the query group S to compare the relevance scores of every query, the singleton query cluster S includes only qs1, clks1, is defines the fusion relevance vector relv (qs1, clks1). A query cluster S= hqs1, clks1... qsk, clkski with k>1, to establish cxts by using few methods.

Query Image: The fusion relevance vector of the query q, relq, to store the amount of each query significance q0, Vq to q. The on-line query group relq for query relevance is used to successful or storage points. Typically, however, it is an extremely tiny amount that doesn't comprehensively convey the relevance of the task of query search, so don't adequate the effective relevance measure, and the robust on-line query group. Instead of storing both queries pertain to financials. On-line Query Grouping. Some programs such as query proposition may be facilitative by speedy on the fly clustering of customer's queries. The performance of unsystematic walk calculation of coalition significance vector of each new query is actual time, and instead of recomputed the query vector of our graph. The work will predominantly well for the queries. Within this situation of run time disk storage performance will be trade-off. This extra storing space is insufficient comparative to the general storing condition of the search engine. The recovery of fusion relevance vector, from the cache can be carried out in the span of time.

EXPERIMENTS

Observe the performance and appearance of the algorithms on dividing a customer's query record into single or many sets of connected queries. For instance, the series of query "Caribbean cruise"; "the bank of Baroda"; "expedia"; "monetary assertion", it could anticipate two output partitions: first "Caribbean cruise", "expedia" concerning to traveling-related queries, and second, "bank of Baroda", "monetary assertion" related to money-oriented queries.

The experiential finding on the position of search records shaped on the root of scheming search record interface. Supply continuous rising past records in the user boundary is the most common utilization of search history. The interface design recommendation for showing search record data is introduced to feed the history data returned to the customer. The first boundary prototype are described and included to represent some of the plan instructions. In addition to the straight search display, resources structure on search record information can help customers in jobs. Investigation of record based interface capabilities are describes structured around a scratchpad and result group tool. Our query group algorithm relies closely in the request of a search log in two ways: first, to assemble the query fusion graph used to compute query significance, and second, to increase the series of queries measured to compute query significance.

PERFORMANCE COMPARISON

The proposed approaches shows the performance can be categorized into five base-lines, all the base-lines are used to select the best query groups. The utilization procedure grouping the queries according to time variations for a query when compared with the latest queries in the above fixed value along with the first base-line. It is basically similar to the time metric presented in part, apart from instead of measure the comparison of the opposite of time interval. The image which is present in QFG will determine the correct estimate by using the above technique is the combination of relevance and click graphs in the query group. It actually estimated to do better afer the assessment was based on more instructions and is hence more truthful. To the other hand, these queries are infrequent within they explore logs or don't have several leaving limits in our chart to make possible the random walk, these methods may execute worse because of the required limits.

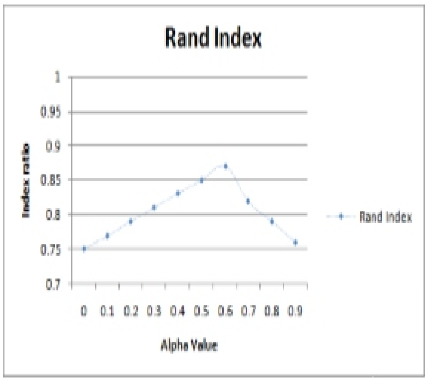


Fig.1 The unstable mix of query and click graph

To assess our algorithm over the chart to build the rising value of a. The out come is exposed in Figure 1.

CONCLUSION

The click graph and query reformulation contain valuable data on consumer behavior when looking on-line. The systematically explored the way to exploit long-term search record, it consist of previously searched queries, the result records and clicks through, the helpful search context that will get better the recovery functionality. The demonstration of search advice may be used proficiently for this task of arranging a user investigates record into the cluster. In addition to run more in-depth screening that's performed with a broad range of stuff, undertaking, and target groups are needed. It like to join the user summary using the document reposition, to offer a broader set of important outcome for the consumer instead of just rearranging the present outcomes.

REFERENCES

[1] Spink, A., Wilson, T. D., Ford, N., Foster, A., & Ellis, D.," Information seeking and mediated searching studies. Part 3.Successive searching. Journal of the American Society for Information Science and Technology", 2002.

[2] D. Beeferman and A. Berger, "Agglomerative clustering of a search engine query log," in KDD, 2000.

[3] R. Jones and K. L. Klinkner, "Beyond the session timeout: Automatic hierarchical segmentation of search topics in query logs", in CIKM, 2008.

[4] Heasoo Hwang, Hady W. Lauw, Lise Getoor and Alexandros Ntoulas," Organizing User Search Histories",in conf. IEEE Transactions on Knowledge and Data Engineering,2012.

[5] A. Border, "A taxonomy of web search," SIGIR Forum, 2002.

[6] P. Boldi, M. Santini, and S. Vigna, "Pagerank as a function of the damping factor," in WWW, 2005.

[7] J. Yi and F. Maghoul, "Query clustering using click-through graph," in WWW, 2009.

[8] Komlodi, A"Search history for user support in information seeking interfaces, University of Maryland"; College Park -2002.

[9] M. Speretta," Personalizing Search Based on User Search Histories", Master's thesis, The University of Kansas, 2004.

[10] Lee, Y. J.," Concept mapping your Web searches: a design rationale and Web-enabled application", Journal of Computer Assisted Learning, 2004.

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

-

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).

2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);

3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);

4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.

The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

Notes: