

## รายการอ้างอิง

1. อภิชาติ พิทยรัตน์โสภณ. 2537. การพัฒนาโครงสร้างข้อมูลพจนานุกรมด้วยทรีแวกซ์. วิทยานิพนธ์ปริญญาโทบริหารธุรกิจ จุฬาลงกรณ์มหาวิทยาลัย.
2. Apostolico, A., and W. Szpankowski. 1987. "Self-alignments in Words and their Applications." Technical Report CSD-TR-732, Department of Computer Science, Purdue University, West Lafayette, Ind. 47907.
3. Flajolet, P. and R. Sedgewick. 1986. "Digital Search Trees Revisited." *SIAM J Computing*, 15;748-67.
4. Cashing, D.L. 1991. A Programmer's guide to file processing. St. Bonaventure University.:PWS-KENT.
5. Dewire, D.T. 1994. Text Management. America: McGraw-Hill.
6. Gonnet, G. 1983. "Unstructured Data Bases or Very Efficient Text Searching," in *ACM PODS*, vol. 2, pp. 117-24, Atlanta, Ga.
7. Gonnet, G. 1984. Handbook of Algorithms and Data Structures. London: Addison-Wesley.
8. Gonnet, G. 1988. "Efficient Searching of Text and Pictures (extended abstract)." Technical Report OED-88-02, Centre for the New OED., University of Waterloo.
9. Gonnet, G. and Baeza, R. 1991. Handbook of Algorithms and Data Structures in Pascal and C. Santiago: Addison-Wesley.
10. Knuth, D. 1973. The Art of Computer Programming: Sorting and Searching, vol. 3. Reading, Mass.: Addison-Wesley.
11. Manber, U., and G. Myers. 1990. "Suffix Arrays: A New Method for On-line String Searches," in *1st ACM-SIAM Symposium on Discrete Algorithms*, pp. 319-27, San Francisco.
12. Morrison, D. 1968. "PATRICIA-Practical Algorithm to Retrieve Information Coded in Alphanumeric." *JACM*, 15;514-34.
13. Pittel, B. 1985. "Asymptotical Growth of a Class of Random Trees." *The Annals of Probability*, 13;414-27.

14. Sedgewick, R. 1988. Algorithms. 2nd ed. :Addison-Wesley.
15. Sedgewick, R. 1990. Algorithms in C.:Addison-Wesley.
16. William,B.F., and Baeza-Yates, R. 1992. Information Retrieval Data Structures & Algorithms. New Jersey:Prentice Hall.
17. Wood, D. 1993. Data Structures, Algorithms, and Performance. University of Waterloo:Addison-Wesley.



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย



ภาคผนวก ก.

ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย

ข้อมูลสถิติของแฟ้มตำแหน่งซีสตริงฉบับที่ 1 (Example.b01)

STATISTIC OF PREFORMAT LARGE TEXT

00001.) Header of Small Text :

Large Text File : example.htm (ชื่อแฟ้มต้นฉบับ)  
 Length Text : 30 (ขนาดแฟ้มต้นฉบับย่อย)  
 Flag Pad : 0 (ตัวระบุแฟ้มต้นฉบับย่อยฉบับสุดท้าย)  
 Base Large Text : 0 (ตำแหน่งเริ่มต้นแฟ้มต้นฉบับ)  
 Index Point Element : 20 (จำนวนซีสตริงทั้งหมด)  
 Maxbytes : 2 (ความยาวซีสตริง)  
 Small Text Element : 2 (จำนวนแฟ้มต้นฉบับย่อย)  
 Pat Array File : NONE (ชื่อแฟ้มแพ็ทอะเรย์ย่อย)

ข้อมูลซีสตริงของแฟ้มตำแหน่งซีสตริงฉบับที่ 1 (Example.d01)

SISTRING OF PREFORMAT LARGE TEXT

00001.) Sistring of Small Text :

0000000001.) at 0000000000 'กา'	0000000002.) at 0000000002 'รพ'
0000000003.) at 0000000003 'พั'	0000000004.) at 0000000005 'ฒน'
0000000005.) at 0000000006 'นา'	0000000006.) at 0000000008 'ระ'
0000000007.) at 0000000010 'บบ'	0000000008.) at 0000000011 'บก'
0000000009.) at 0000000012 'กา'	0000000010.) at 0000000014 'รค'
0000000011.) at 0000000015 'ค'	0000000012.) at 0000000017 'นค'
0000000013.) at 0000000018 'ค'	0000000014.) at 0000000020 'นข'
0000000015.) at 0000000021 'ช'	0000000016.) at 0000000023 'อค'
0000000017.) at 0000000024 'คว'	0000000018.) at 0000000025 'วา'
0000000019.) at 0000000027 'มภ'	0000000020.) at 0000000028 'ภา'

ข้อมูลสถิติของแฟ้มตำแหน่งซีสตริงฉบับที่ 2 (Example.b02)

STATISTIC OF PREFORMAT LARGE TEXT

00002.) Header of Small Text :

Large Text File : example.htm (ชื่อแฟ้มต้นฉบับ)  
 Length Text : 24 (ขนาดแฟ้มต้นฉบับย่อย)  
 Flag Pad : 1 (ตัวระบุแฟ้มต้นฉบับย่อยฉบับสุดท้าย)

Base Large Text : 29 (ตำแหน่งเริ่มต้นเพิ่มเติมฉบับ)  
 Index Point Element : 16 (จำนวนซิสตริงทั้งหมด)  
 Maxbytes : 2 (ความยาวซิสตริง)  
 Small Text Element : 2 (จำนวนเพิ่มเติมฉบับย่อย)  
 Pat Array File : NONE (ชื่อเพิ่มแพ็ทอะเรย์)

ข้อมูลซิสตริงของเพิ่มเติมตำแหน่งซิสตริงฉบับที่ 2 (Example.d02)

SISTRING OF PREFORMAT LARGE TEXT

00002.) Sistring of Small Text :

000000001.) at 0000000030	'ษา'	000000002.) at 0000000032	'ไท'
000000003.) at 0000000033	'ทย'	000000004.) at 0000000034	'ยโ'
000000005.) at 0000000035	'โค'	000000006.) at 0000000036	'ดย'
000000007.) at 0000000037	'ยไ'	000000008.) at 0000000038	'ไซ'
000000009.) at 0000000039	'ชี'	000000010.) at 0000000041	'ดี'
000000011.) at 0000000043	'นไ'	000000012.) at 0000000044	'ไม'
000000013.) at 0000000045	'มี'	000000014.) at 0000000047	'แพ'
000000015.) at 0000000048	'พี'	000000016.) at 0000000050	'ทง'

ข้อมูลสถิติการสร้างแพ็ทอะเรย์ (Example.c01)

This Patricia has statistic :

1. Maxbytes : 2 bytes = 16 bits
2. File TEXT : example.i01 has size 180 bytes
3. File PAT TREE : example.t01 has size 384 bytes  
 File PAT ARRAY : example.p01 has size 252 bytes
4. Duplicate Key : 1 times
5. Height of Trees : 8
6. Internal Separate Node : 5
  - Internal Left Node : 6
  - Internal Right Node : 2
  - Internal Node : 13
  - External Node : 6
  - Duplicate Node : 1



Root Node : 1

Total Node : 21

This PAT ARRAY has HEADER :

Large Text File : example.htm

Length Text : 30

Flag Pad : 0

Base Large Text : 0

Index Point Element : 20

Maxbytes : 2

Index Point Name : example.i01

Duplicate Small Name : example.s01

Pat Array Element : 19

ข้อมูลต้นไม้แพต (Example.i01)

List Index File :

```

000000001 00016 000000000 000000001 000000030 000000000
000000002 00015 000000001 000000002 000000000 000000009
000000003 00014 000000003 000000016 000000002 000000000
000000004 00012 000000011 000000004 000000003 000000000
000000005 00011 000000004 000000005 000000005 000000000
000000006 00010 000000007 000000003 000000006 000000000
000000007 00006 000000010 000000006 000000008 000000000
000000008 00009 000000012 000000008 000000010 000000000
000000009 00004 000000008 000000007 000000011 000000000
000000010 00015 000000001 000000002 000000012 000000000
000000011 00004 000000010 000000002 000000014 000000000
000000012 00010 000000015 000000013 000000015 000000000
000000013 00006 000000014 000000005 000000017 000000000
000000014 00005 000000017 000000011 000000018 000000000
000000015 00002 000000014 000000012 000000020 000000000
000000016 00009 000000001 000000015 000000021 000000000

```

0000000017 00011 0000000018 0000000016 0000000023 0000000000  
 0000000018 00004 0000000017 0000000013 0000000024 0000000000  
 0000000019 00010 0000000019 0000000018 0000000025 0000000000  
 0000000020 00009 0000000020 0000000006 0000000027 0000000000  
 0000000021 00008 0000000020 0000000019 0000000028 0000000000

ข้อมูลชนิดเรียงตามค่าในแพ็คเกจอะเรย์ (Example.z01)

This Patricia : Small pat array has sistring :

0000000001:0000000000 'กา'

0000000002:0000000012 (D)'กา'

0000000003:0000000021 'ขี'

0000000004:0000000024 'กว'

0000000005:0000000018 'กี'

0000000006:0000000015 'คี่'

0000000007:0000000005 'ฉน'

0000000008:0000000020 'นข'

0000000009:0000000017 'นค'

0000000010:0000000006 'นา'

0000000011:0000000011 'บก'

0000000012:0000000010 'บบ'

0000000013:0000000003 'พี'

0000000014:0000000028 'ภา'

0000000015:0000000027 'มภ'

0000000016:0000000014 'รค'

0000000017:0000000002 'รพ'

0000000018:0000000008 'ระ'

0000000019:0000000025 'วา'

0000000020:0000000023 'อค'

ข้อมูลสถิติการสร้างแพ็คเกจอะเรย์ (Example.c02)

This Patricia has statistic :

1. Maxbytes : 2 bytes = 16 bits

2. File TEXT : example.i02 has size 164 bytes  
 3. File PAT TREE : example.t02 has size 312 bytes  
 File PAT ARRAY : example.p02 has size 228 bytes  
 4. Duplicate Key : 0 times  
 5. Height of Trees : 6  
 6. Internal Separate Node : 5  
 Internal Left Node : 1  
 Internal Right Node : 4  
 Internal Node : 10  
 External Node : 6  
 Duplicate Node : 0  
 Root Node : 1  
 Total Node : 17

This PAT ARRAY has HEADER :

- Large Text File : example.htm  
 Length Text : 24  
 Flag Pad : 1  
 Base Large Text : 29  
 Index Point Element : 16  
 Maxbytes : 2  
 Index Point Name : example.i02  
 Duplicate Small Name : example.s02  
 Pat Array Element : 16

ข้อมูลต้นไม้ม้วน (Example.i02)

List Index File :

```
000000001 00016 000000000 000000001 000000030 000000000
000000002 00015 000000001 000000003 000000001 000000000
000000003 00013 000000004 000000005 000000003 000000000
000000004 00014 000000009 000000002 000000004 000000000
000000005 00011 000000013 000000001 000000005 000000000
```



000000006 00010 000000014 000000012 000000006 000000000  
 000000007 00009 000000010 000000016 000000007 000000000  
 000000008 00000 000000004 000000007 000000008 000000000  
 000000009 00008 000000005 000000008 000000009 000000000  
 000000010 00012 000000009 000000011 000000010 000000000  
 000000011 00008 000000006 000000010 000000012 000000000  
 000000012 00011 000000006 000000015 000000014 000000000  
 000000013 00006 000000002 000000012 000000015 000000000  
 000000014 00009 000000013 000000007 000000016 000000000  
 000000015 00009 000000014 000000008 000000018 000000000  
 000000016 00010 000000011 000000015 000000019 000000000  
 000000017 00007 000000016 000000003 000000021 000000000

ข้อมูลซิสตริงตามค่าในแพ็คเกจอะเรย์ (Example.z02)

This Patricia : Small pat array has sistring :

000000001:000000039 'ซึ'  
 000000002:000000036 'คย'  
 000000003:000000041 'ดี'  
 000000004:000000050 'ทง'  
 000000005:000000033 'ทย'  
 000000006:000000043 'นไ'  
 000000007:000000048 'พี'  
 000000008:000000045 'มี'  
 000000009:000000034 'ยโ'  
 000000010:000000037 'ยใ'  
 000000011:000000030 'ษา'  
 000000012:000000047 'แพ'  
 000000013:000000035 'โค'  
 000000014:000000038 'ไซ'  
 000000015:000000032 'ไท'  
 000000016:000000044 'ไม'

ข้อมูลสถิติการผสมผสานแฟ้มต่อระเบียบ (Merge.sta)

This Merge Small to Large file has statistic :

1. File SMALL TEXT : example.p02 has size 53 bytes  
File PAT ARRAY SMALL : example.p02 has size 228 bytes
2. File LARGE DIRECTORY : merge.dir has size 40 bytes  
File PAT ARRAY LARGE : merge.pat has size 380 bytes  
File STAT OF MERGE : merge.sta

3. Directory of Large PAT :

00001.) 0000000030 example.i01

00002.) 0000000054 example.i02

4. Duplicate Elements : 0
5. Duplicate Totals : 0
6. Small Pat Piece : 2

This LARGE PAT ARRAY has HEADER :

Duplicate Large Name : merge.dup

Directory Large Name : merge.dir

Maxbytes : 2

Pat Array Element : 35

Pat Duplicate : 0

Pat Dupmerge : 0

Pat Piece : 2

ข้อมูลซิสตริงตามแฟ้มต่อระเบียบฉบับใหม่ (Merge.sis)

This Large pat array has sistring :

0000000001:0000000000 'กา'

0000000002:0000000012 (D)'กา'

0000000003:0000000021 'จู้'

0000000004:0000000024 'คว'

0000000005:0000000018 'กึ'

0000000006:0000000015 'กึ'

0000000007:0000000039 'จู้'

0000000008:0000000005 'ฒน'  
 0000000009:0000000036 'ดย'  
 0000000010:0000000041 'ด'  
 0000000011:0000000050 'ทง'  
 0000000012:0000000033 'ทย'  
 0000000013:0000000020 'นข'  
 0000000014:0000000017 'นค'  
 0000000015:0000000006 'นา'  
 0000000016:0000000043 'นไ'  
 0000000017:0000000011 'บก'  
 0000000018:0000000010 'บบ'  
 0000000019:0000000003 'พี'  
 0000000020:0000000048 'พี'  
 0000000021:0000000028 'ภา'  
 0000000022:0000000027 'มภ'  
 0000000023:0000000045 'มี'  
 0000000024:0000000034 'ยโ'  
 0000000025:0000000037 'ยใ'  
 0000000026:0000000014 'รค'  
 0000000027:0000000002 'รพ'  
 0000000028:0000000008 'ระ'  
 0000000029:0000000025 'วา'  
 0000000030:0000000030 'ษา'  
 0000000031:0000000023 'อค'  
 0000000032:0000000047 'แพ'  
 0000000033:0000000035 'โด'  
 0000000034:0000000038 'ไซ'  
 0000000035:0000000032 'ไท'  
 0000000036:0000000044 'ไม'

ข้อมูลสถิติกับแพ็ทอะเรย์สั้น (Lschar.ls)

This Large PAT array has SPAT :

1. Memory Use : 10
2. Ls char : 2
3.  $r = (\text{Memory}/\text{Ls char})$  : 5
4. pat array# : 35
5.  $b = (\text{pat array\#/r})$  : 7
6. flag\_last\_spat : 0
7. impicit\_pointer : 0

SPAT(1) = 'ณน'

SPAT(2) = 'นา'

SPAT(3) = 'นภ'

SPAT(4) = 'วา'

SPAT(5) = 'ไม'

ข้อมูลผลการค้นหาแบบเต็มหน้า (Search.log)

Search Text:(นา) on Large Pat Array

000000001. Found นา at 000000006 (นา) of example.i01 of example.htm

Conclude Search :

Found in Document : example.i01

ข้อมูลผลการค้นหาแบบบูล (Search.log)

Search Text:(นา+ไม) on Large Pat Array

000000001. Found นา at 000000006 (นา) of example.i01 of example.htm

000000001. Found ไม at 000000044 (ไม) of example.i02 of example.htm

Conclude Search :

Found in Document : example.i01

Found in Document : example.i02



## ประวัติผู้เขียน

นายเปรมิน จินดาวิมลเลิศ เกิดเมื่อวันที่ 10 สิงหาคม ที่กรุงเทพมหานคร สำเร็จการศึกษา  
ระดับปริญญาตรีวิทยาศาสตร์บัณฑิต สาขาวิศวกรรมคอมพิวเตอร์ คณะวิทยาศาสตร์และเทคโนโลยี  
มหาวิทยาลัยธรรมศาสตร์ และเข้าศึกษาต่อในหลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิทยาศาสตร์  
คอมพิวเตอร์ ภาควิชาวิศวกรรมคอมพิวเตอร์ ที่จุฬาลงกรณ์มหาวิทยาลัย เมื่อ พ.ศ. 2537



ศูนย์วิทยทรัพยากร  
จุฬาลงกรณ์มหาวิทยาลัย