

Sub-project 'Pagoda' aka 'Schi sch' aka Shi n-Bashi ra (Heart-Pillar)

/Unfinished notes on making the "crucified/schi sch" table for a given exact word/

Building the 'on' pagoda and some notes:

First, the x-grams in use come from 'Gamera_r17_LBL' corpus, here all (with occurrences 4 and more) 1-grams/2-grams/3-grams/4-grams/5-grams.

It took 20 minutes to extract 'on' i.e. to build 'on' pagoda, ALL WORDS HAVE THEIR OWN PAGODAS, when (as here) the pagoda is 5 tiers high then the subtiers are $1+2+3+4+5=5*(1+5)/2=15$, see further below the **schi sch.bat** run for 'on'.

The (external) memory needed for that 'on' pagoda is 300MB in 15 files.

Those 15 subtiers represent the usage of that word, pagoda's heart-pillar is this very word.

Of course, limiting the height of the pagoda is memory size bound, in my view pagodas with 13 tiers should be available for each and every major English word - some 93,000 pagodas, which screams for Terabytes (roughly: $93000*100MB=9,300,000MB=9TB$).

Let's see how it looks like:

tier #01	subtier #01	on	houses 1 storey
tier #02	subtier #02	on	houses 67,917 stories
tier #02	subtier #03	on	houses 80,416 stories
tier #03	subtier #04	on	houses 376,735 stories
tier #03	subtier #05	on	houses 535,289 stories
tier #03	subtier #06	on	houses 680,282 stories
tier #04	subtier #07	on	houses 657,578 stories
tier #04	subtier #08	on	houses 1,032,827 stories
tier #04	subtier #09	on	houses 968,518 stories
tier #04	subtier #10	on	houses 980,803 stories
tier #05	subtier #11	on	houses 648,399 stories
tier #05	subtier #12	on	houses 778,578 stories
tier #05	subtier #13	on	houses 977,966 stories
tier #05	subtier #14	on	houses 974,096 stories
tier #05	subtier #15	on	houses 869,022 stories

Dots stand for any word, also the pagoda is not symmetrical - very ... asymmetric, kind of ... maverick, that's right.

And time to get scared: all tiers consist of

1+67,917+80,416+376,735+535,289+680,282+657,578+1,032,827+968,518+980,803+648,399+778,578+977,966+974,096+869,022=9,628,427 levels, that is, x-grams. Yes, with 9,628,427 stories it is a spacescraper, and what.

E:_KAZE_Gamera_r17_LBL_schi sch>**dir**

```
08/09/2013  10:49 PM                1,259 46ARC.bat
08/09/2013  10:49 PM                1,351 46GET.bat
08/09/2013  10:49 PM                1,075 46RIP.bat
08/09/2013  10:49 PM          7,112,936 4andabove_Gamera17LBL.1.txt.sorted.bsc
08/09/2013  10:49 PM          83,951,241 4andabove_Gamera17LBL.2.txt.sorted.bsc
08/09/2013  10:49 PM         271,950,657 4andabove_Gamera17LBL.3.txt.sorted.bsc
08/09/2013  10:49 PM         442,539,210 4andabove_Gamera17LBL.4.txt.sorted.bsc
08/09/2013  10:49 PM         516,645,416 4andabove_Gamera17LBL.5.txt.sorted.bsc
08/08/2013   08:26 PM          58,445,626 4andabove_Gamera17LBL.1.txt.sorted
08/08/2013   08:26 PM          938,593,835 4andabove_Gamera17LBL.2.txt.sorted
08/08/2013   08:26 PM          3,116,773,287 4andabove_Gamera17LBL.3.txt.sorted
08/08/2013   08:26 PM          4,839,075,871 4andabove_Gamera17LBL.4.txt.sorted
08/08/2013   08:26 PM          5,077,549,028 4andabove_Gamera17LBL.5.txt.sorted
08/09/2013  10:49 PM                84 DeGraffi thi ze_x-leton_MAX.bat
08/09/2013  10:49 PM                355 DEG_all.bat
08/09/2013  10:49 PM                136 Graffi thi ze_x-leton_MAX.bat
08/09/2013  10:49 PM          627,200 GRAFFITH_r2++_Graphein_2.3.0_Intel_12.1_32bit_768MB.exe
08/09/2013  10:49 PM          480,256 Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_HEXADECAD-Thread_Intel V12.exe
08/09/2013  10:49 PM          165,376 Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe
08/09/2013  10:49 PM          129,536 Leprechaun_x-leton_32bit_Intel_01_01p.exe
08/09/2013  10:49 PM           92,412 Linereporter.c
08/09/2013  10:49 PM          76,800 Linereporter.exe
08/09/2013  10:49 PM          57,527 Linereporter_r1+FI X.zip
08/09/2013  10:49 PM           1,015 MATCHKALERO.bat
08/09/2013  10:49 PM           2,695 schi sch.bat
08/09/2013  10:49 PM          49,152 sha1sum.exe
08/09/2013  10:49 PM           1,049 UNMATCHKALERO.bat
```

E:_KAZE_Gamera_r17_LBL_schi sch>**type schi sch.bat**

@echo off

if '%1'==' ' goto usage

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1" 4andabove_Gamera17LBL.1.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.1.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_" 4andabove_Gamera17LBL.2.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.2-1.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_" 4andabove_Gamera17LBL.2.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.2-2.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.3.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.3-1.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.3.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.3-2.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.3.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.3-3.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.4.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.4-1.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.4.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.4-2.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.4.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.4-3.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.4.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.4-4.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.5.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.5-1.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.5.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.5-2.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.5.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.5-3.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.5.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.5-4.txt

"Kazahana_r1-++fi x+nowai t_cri ti cal _ni xF I X_MONAD-Thread_Intel V12.exe" "%1_." 4andabove_Gamera17LBL.5.txt.sorted 1023
ren Kazahana.txt Kazahana_%1.5-5.txt

dir Kazahana_%1.*.txt/b>q

Linereporter.exe q>>schi sch.log

Leprechaun_x-leton_32bit_Intel_01_01p.exe q q.wrd 1234567 y

sort q.wrd /O Kazahana_%1.wrd

del q

del q.wrd

dir Kazahana_%1.*>>schi sch.log

```
goto quit
:usage
echo Usage: schisch.bat word
:quit
```

E:\KAZE_Gamera_r17_LBL_schisch>**schisch.bat on**

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,063,452 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 2,805,843/2,805,843/1

Kazahana: Performance: 60 KB/clock

Kazahana: Performance: 2,988 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 939/95

Kazahana: Performance: I/O time, i.e. fread() time, is 10 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 215,386,292 ticks

Kazahana: Done.

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,070,008 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 37,030,186/37,030,186/67,917

Kazahana: Performance: 68 KB/clock

Kazahana: Performance: 2,758 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 13,423/1,512

Kazahana: Performance: I/O time, i.e. fread() time, is 11 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 3,543,505,460 ticks

Kazahana: Done.

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,040,706 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 37,030,186/37,030,186/80,416

Kazahana: Performance: 39 KB/clock

Kazahana: Performance: 1,604 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 23,079/845

Kazahana: Performance: I/O time, i.e. fread() time, is 3 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 1,898,832,837 ticks

Kazahana: Done.

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,075,690 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 106,154,564/106,154,564/376,735

Kazahana: Performance: 73 KB/clock

Kazahana: Performance: 2,577 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 41,188/4,817

Kazahana: Performance: I/O time, i.e. fread() time, is 11 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 11,388,659,326 ticks

Kazahana: Done.

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,045,083 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 106,154,564/106,154,564/535,289

Kazahana: Performance: 44 KB/clock

Kazahana: Performance: 1,535 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 69,142/3,739

Kazahana: Performance: I/O time, i.e. fread() time, is 5 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 9,389,312,306 ticks

Kazahana: Done.

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,036,448 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 106,154,564/106,154,564/680,282

Kazahana: Performance: 35 KB/clock

Kazahana: Performance: 1,241 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 85,532/4,418

Kazahana: Performance: I/O time, i.e. fread() time, is 5 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 10,486,508,813 ticks

Kazahana: Done.

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,080,057 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 144,793,474/144,793,474/657,578

Kazahana: Performance: 78 KB/clock

Kazahana: Performance: 2,395 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 60,454/8,070

Kazahana: Performance: I/O time, i.e. fread() time, is 13 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 18,520,995,713 ticks

Kazahana: Done.

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,049,467 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 144,793,474/144,793,474/1,032,827

Kazahana: Performance: 48 KB/clock

Kazahana: Performance: 1,480 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 97,829/7,127

Kazahana: Performance: I/O time, i.e. fread() time, is 7 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 16,481,279,089 ticks

Kazahana: Done.

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fix+nowait_critical_nixFIX, copyleft Kaze 2013-Apr-07.

Enforcing MONAD i.e. single-thread ...

Allocating Master-Buffer 1023KB ... OK

-; 00,000,040,988 bytes/clock

Kazahana: Total /Checked/Dumped xgrams: 144,793,474/144,793,474/968,518

Kazahana: Performance: 40 KB/clock

Kazahana: Performance: 1,226 xgrams/clock

Kazahana: Performance: Total /fread() clocks: 118,095/7,141

Kazahana: Performance: I/O time, i.e. fread() time, is 6 percents

Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 15,355,443,939 ticks

Kazahana: Done.

```

Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fi x+nowai t_cri tical _ni xFfX,
copyleft Kaze 2013-Apr-07.
Enforcing MONAD i.e. single-thread ...
Allocating Master-Buffer 1023KB ... OK
-; 00,000,034,899 bytes/clock
Kazahana: Total/Checked/Dumped xgrams: 144,793,474/144,793,474/980,803
Kazahana: Performance: 34 KB/clock
Kazahana: Performance: 1,044 xgrams/clock
Kazahana: Performance: Total/fread() clocks: 138,673/6,716
Kazahana: Performance: I/O time, i.e. fread() time, is 4 percents
Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 16,942,129,446 ticks
Kazahana: Done.
Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fi x+nowai t_cri tical _ni xFfX,
copyleft Kaze 2013-Apr-07.
Enforcing MONAD i.e. single-thread ...
Allocating Master-Buffer 1023KB ... OK
-; 00,000,079,276 bytes/clock
Kazahana: Total/Checked/Dumped xgrams: 133,553,846/133,553,846/648,399
Kazahana: Performance: 77 KB/clock
Kazahana: Performance: 2,084 xgrams/clock
Kazahana: Performance: Total/fread() clocks: 64,063/11,176
Kazahana: Performance: I/O time, i.e. fread() time, is 17 percents
Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 24,552,439,417 ticks
Kazahana: Done.
Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fi x+nowai t_cri tical _ni xFfX,
copyleft Kaze 2013-Apr-07.
Enforcing MONAD i.e. single-thread ...
Allocating Master-Buffer 1023KB ... OK
-; 00,000,050,520 bytes/clock
Kazahana: Total/Checked/Dumped xgrams: 133,553,846/133,553,846/778,578
Kazahana: Performance: 49 KB/clock
Kazahana: Performance: 1,328 xgrams/clock
Kazahana: Performance: Total/fread() clocks: 100,516/12,949
Kazahana: Performance: I/O time, i.e. fread() time, is 12 percents
Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 27,642,647,747 ticks
Kazahana: Done.
Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fi x+nowai t_cri tical _ni xFfX,
copyleft Kaze 2013-Apr-07.
Enforcing MONAD i.e. single-thread ...
Allocating Master-Buffer 1023KB ... OK
-; 00,000,044,771 bytes/clock
Kazahana: Total/Checked/Dumped xgrams: 133,553,846/133,553,846/977,966
Kazahana: Performance: 43 KB/clock
Kazahana: Performance: 1,177 xgrams/clock
Kazahana: Performance: Total/fread() clocks: 113,439/8,953
Kazahana: Performance: I/O time, i.e. fread() time, is 7 percents
Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 20,352,124,056 ticks
Kazahana: Done.
Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fi x+nowai t_cri tical _ni xFfX,
copyleft Kaze 2013-Apr-07.
Enforcing MONAD i.e. single-thread ...
Allocating Master-Buffer 1023KB ... OK
-; 00,000,038,786 bytes/clock
Kazahana: Total/Checked/Dumped xgrams: 133,553,846/133,553,846/974,096
Kazahana: Performance: 37 KB/clock
Kazahana: Performance: 1,019 xgrams/clock
Kazahana: Performance: Total/fread() clocks: 130,938/8,184
Kazahana: Performance: I/O time, i.e. fread() time, is 6 percents
Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 18,774,428,970 ticks
Kazahana: Done.
Kazahana, a superFast exact & wildcards & Levenshtein Distance (Wagner-Fischer) searcher, r. 1-++fi x+nowai t_cri tical _ni xFfX,
copyleft Kaze 2013-Apr-07.
Enforcing MONAD i.e. single-thread ...
Allocating Master-Buffer 1023KB ... OK
-; 00,000,034,646 bytes/clock
Kazahana: Total/Checked/Dumped xgrams: 133,553,846/133,553,846/869,022
Kazahana: Performance: 33 KB/clock
Kazahana: Performance: 911 xgrams/clock
Kazahana: Performance: Total/fread() clocks: 146,594/6,378
Kazahana: Performance: I/O time, i.e. fread() time, is 4 percents
Kazahana: Performance: RDTSC I/O time, i.e. fread() time, is 16,867,552,196 ticks
Kazahana: Done.

```

E:_KAZE_Gamera_r17_LBL_schisch>dl r Kazahana_on.*

```

08/08/2013 08:27 PM          14 Kazahana_on.1.txt {houses          1 x-gram}
08/08/2013 08:27 PM      1,526,091 Kazahana_on.2-1.txt {houses      67,917 x-grams}
08/08/2013 08:28 PM      1,814,333 Kazahana_on.2-2.txt {houses      80,416 x-grams}
08/08/2013 08:29 PM     10,214,496 Kazahana_on.3-1.txt {houses     376,735 x-grams}
08/08/2013 08:30 PM     14,707,122 Kazahana_on.3-2.txt {houses     535,289 x-grams}
08/08/2013 08:31 PM     18,836,032 Kazahana_on.3-3.txt {houses     680,282 x-grams}
08/08/2013 08:32 PM     20,525,010 Kazahana_on.4-1.txt {houses     657,578 x-grams}
08/08/2013 08:34 PM     33,067,735 Kazahana_on.4-2.txt {houses    1,032,827 x-grams}
08/08/2013 08:36 PM     30,589,460 Kazahana_on.4-3.txt {houses     968,518 x-grams}
08/08/2013 08:38 PM     31,099,819 Kazahana_on.4-4.txt {houses     980,803 x-grams}
08/08/2013 08:39 PM     23,167,283 Kazahana_on.5-1.txt {houses     648,399 x-grams}
08/08/2013 08:41 PM     28,639,569 Kazahana_on.5-2.txt {houses     778,578 x-grams}
08/08/2013 08:43 PM     35,951,048 Kazahana_on.5-3.txt {houses     977,966 x-grams}
08/08/2013 08:45 PM     35,105,695 Kazahana_on.5-4.txt {houses     974,096 x-grams}
08/08/2013 08:47 PM     31,656,723 Kazahana_on.5-5.txt {houses     869,022 x-grams}

```

E:_KAZE_Gamera_r17_LBL_schwork>type Kazahana_on.1.txt
9,999,999 on

E:_KAZE_Gamera_r17_LBL_schwork>type Kazahana_on.2-1.txt|more
9,087,698 on_the
0,939,934 on_his
0,444,778 on_thi s
...

E:_KAZE_Gamera_r17_LBL_schwork>type Kazahana_on.2-2.txt|more
0,529,688 based_on
0,372,517 and_on
0,302,085 go_on
...

E:_KAZE_Gamera_r17_LBL_schwork>type Kazahana_on.3-1.txt|more
0,396,236 on_the_other
0,143,762 on_account_of
0,132,696 on_the_ground
...

E:_KAZE_Gamera_r17_LBL_schwork>type Kazahana_on.3-2.txt|more
0,176,560 based_on_the

```

0, 143, 498      and_on_the
0, 088, 361      was_on_the
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 3-3. txt|more
0, 154, 080      and_so_on
0, 083, 764      s_books_on
0, 079, 567      i_s_based_on
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 4-1. txt|more
0, 227, 093      on_the_other_hand
0, 086, 049      on_the_other_side
0, 083, 062      on_the_part_of
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 4-2. txt|more
0, 083, 586      books_on_cd_rom
0, 016, 966      and_on_the_other
0, 012, 961      encycl opedi a_on_cd_rom
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 4-3. txt|more
0, 083, 575      s_books_on_cd
0, 032, 257      i_s_based_on_the
0, 014, 713      sat_down_on_the
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 4-4. txt|more
0, 083, 575      osho_s_books_on
0, 027, 369      for_more_i nformation_on
0, 022, 680      what_s_goi ng_on
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 5-1. txt|more
0, 044, 133      on_the_other_si de_of
0, 042, 008      on_the_part_of_the
0, 019, 850      on_the_edge_of_the
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 5-2. txt|more
0, 012, 957      encycl opedi a_on_cd_rom_contai ns
0, 012, 485      based_on_separate_sources_get
0, 012, 335      provi de_on_request_at_no
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 5-3. txt|more
0, 083, 575      s_books_on_cd_rom
0, 012, 957      cathol i c_encycl opedi a_on_cd_rom
0, 012, 512      versi ons_based_on_separate_sources
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 5-4. txt|more
0, 083, 575      osho_s_books_on_cd
0, 012, 313      you_recei ved_i t_on_a
0, 012, 107      they_may_be_on_may
...
E: \_KAZE_Gamera_r17_LBL_schi_schwork>type Kazahana_on. 5-5. txt|more
0, 083, 564      foundati on_osho_s_books_on
0, 012, 363      medi um_i t_may_be_on
0, 012, 313      i f_you_recei ved_i t_on
...

```

The first/top 3 x-grams are given from each subtier, in other words the TOP 45:

```

9, 999, 999      on
9, 087, 698      on_the
0, 939, 934      on_his
0, 444, 778      on_thi s
0, 529, 688      based_on
0, 372, 517      and_on
0, 302, 085      go_on
0, 396, 236      on_the_other
0, 143, 762      on_account_of
0, 132, 696      on_the_ground
0, 176, 560      based_on_the
0, 143, 498      and_on_the
0, 088, 361      was_on_the
0, 154, 080      and_so_on
0, 083, 764      s_books_on
0, 079, 567      i_s_based_on
0, 227, 093      on_the_other_hand
0, 086, 049      on_the_other_si de
0, 083, 062      on_the_part_of
0, 083, 586      books_on_cd_rom
0, 016, 966      and_on_the_other
0, 012, 961      encycl opedi a_on_cd_rom
0, 083, 575      s_books_on_cd
0, 032, 257      i_s_based_on_the
0, 014, 713      sat_down_on_the
0, 083, 575      osho_s_books_on
0, 027, 369      for_more_i nformati on_on
0, 022, 680      what_s_goi ng_on
0, 044, 133      on_the_other_si de_of
0, 042, 008      on_the_part_of_the
0, 019, 850      on_the_edge_of_the
0, 012, 957      encycl opedi a_on_cd_rom_contai ns
0, 012, 485      based_on_separate_sources_get
0, 012, 335      provi de_on_request_at_no
0, 083, 575      s_books_on_cd_rom
0, 012, 957      cathol i c_encycl opedi a_on_cd_rom
0, 012, 512      versi ons_based_on_separate_sources
0, 083, 575      osho_s_books_on_cd
0, 012, 313      you_recei ved_i t_on_a
0, 012, 107      they_may_be_on_may
0, 083, 564      foundati on_osho_s_books_on
0, 012, 363      medi um_i t_may_be_on
0, 012, 313      i f_you_recei ved_i t_on

```

The so-needed precomputed 'on'-like BINDING-WORDS are a MUST-HAVE. Here I did rip 46 utmost major BINDING-WORDS, they span 8,565,872,967 bytes across 46x(15+1)=736 files:

Pagoda dimensions of these [about, above, across, after, against, along, among, apart, around, as, aside, at, away, before, behind, below, between, beyond, back, but, by, down, for, forward, forth, from, in, inside, into, near, of, off, on, onto, or,

out, over, through, to, under, unto, up, upon, with, within, without] 46 words are given below:

For instance, 'on' pagoda has width/height/weight: 72/9, 628, 427/318, 449, 813 measured in bytes/lines/bytes:

...

Another usage: *Do 'disguised' and 'humanism' coexist within one pagoda order 5?*

The answer lies in pagoda's distinct words list 'Kazahana_humanism.wrd' which, shocker, contains not 'disguised'.

A glimpse at what constitutes the process of x-gram utilizing, or of ... pagodifying.

Note: As my English is forever crippled and I didn't know what preposition (or rather postposition) to use with 'glimpse' I had to build a pagoda for the latter, however I chose the backward approach to inquire into 'at' pagoda, conveniently I precomputed 46 utmost major BINDING-WORDS, naturally conjunctions plus adverbs plus prepositions fall in one category: BINDING-WORDS. Initially I thought that 'glimpse on/over/in/of' were plausible, let's see how 'at' fares statistically:

```
0,050,309      glimpse_of
0,000,983      glimpse_at
0,000,340      glimpse_in
0,000,085      glimpse_over
0,000,065      glimpse_on
```

Heritage holds for 'glimpse':

v. tr.

To obtain a brief, incomplete view of.

v. intr.

To look briefly; glance: glimpsed at the headlines

Well I was blind for this nuance.

It seems to me that 'of' emphasizes the subject more than 'at'.

I believe (no time now for thorough walkthrough) that exploring those 50,309 occurrences will reveal and confirm the concise definitions provided by HERITAGE. Okay, found the time: narrowing those 50,309 down to 3-grams reduces them to 33,530.

The key thing, all along, is that I am careless what roles plays 'glimpse', my attention is locked on the statistics for 'a_glimpse_' the very phrase my sentence is to begin.

Looking into 'Kazahana_of.3-3.txt' i.e. '._._of' x-grams:

Usage as a noun:

```
0,033,530      a_glimpse_of
```

Usage as a verb:

NONE

Looking into 'Kazahana_at.3-3.txt' i.e. '._._at' x-grams:

Usage as a noun:

```
0,000,548      a_glimpse_at
0,000,041      one_glimpse_at
0,000,033      first_glimpse_at
0,000,028      quick_glimpse_at
0,000,022      last_glimpse_at
0,000,019      another_glimpse_at
0,000,015      backward_glimpse_at
0,000,013      few_glimpsed_at
```

...

Usage as a verb:

```
0,000,012      to_glimpse_at
```

...

Looking into 'Kazahana_in.3-3.txt' i.e. '._._in' x-grams:

Usage as a noun:

```
0,000,215      a_glimpse_in
```

Usage as a verb:

NONE

Looking into 'Kazahana_over.3-3.txt' i.e. '._._over' x-grams:

Usage as a noun:

```
0,000,055      a_glimpse_over
0,000,005      first_glimpse_over
0,000,005      barest_glimpse_over
```

...

Usage as a verb:

NONE

Looking into 'Kazahana_on.3-3.txt' i.e. '._._on' x-grams:

Usage as a noun:

```
0,000,034      a_glimpse_on
0,000,006      faint_glimpsed_on
0,000,004      gave_glimpsed_on
```

...

Usage as a verb:

```
0,000,038      had_glimpsed_on
0,000,007      to_glimpse_on
```

...

So, I stand corrected, my second choice becomes 'a_glimpse_of' replacing my first - the second-in-queue 'a_glimpse_at'. Similar phrases as 'look/glance at' deserve similar investigation.

Let's build one more pagoda.

Another usage, say I forgot what was the name of a book written by Grimmelshausen I have once read, or want to see words having something to do with 'grimmelshausen':

E:_KAZE_Gamera_r17_LBL_schisch>**schisch.bat grimmelshausen**

...

E:_KAZE_Gamera_r17_LBL_schisch>**type schisch.log**

Linereporter: Encountered lines in all files: 21

Linereporter: Longest line: 52

```
08/09/2013 04:32 AM      26 Kazahana_grimmelshausen.1.txt
08/09/2013 04:32 AM      86 Kazahana_grimmelshausen.2-1.txt
08/09/2013 04:32 AM     117 Kazahana_grimmelshausen.2-2.txt
08/09/2013 04:33 AM      43 Kazahana_grimmelshausen.3-1.txt
08/09/2013 04:34 AM      94 Kazahana_grimmelshausen.3-2.txt
08/09/2013 04:36 AM     155 Kazahana_grimmelshausen.3-3.txt
08/09/2013 04:36 AM        0 Kazahana_grimmelshausen.4-1.txt
08/09/2013 04:37 AM        0 Kazahana_grimmelshausen.4-2.txt
08/09/2013 04:38 AM        0 Kazahana_grimmelshausen.4-3.txt
08/09/2013 04:43 AM     137 Kazahana_grimmelshausen.4-4.txt
08/09/2013 04:43 AM        0 Kazahana_grimmelshausen.5-1.txt
08/09/2013 04:44 AM        0 Kazahana_grimmelshausen.5-2.txt
08/09/2013 04:45 AM        0 Kazahana_grimmelshausen.5-3.txt
08/09/2013 04:47 AM        0 Kazahana_grimmelshausen.5-4.txt
08/09/2013 04:52 AM     106 Kazahana_grimmelshausen.5-5.txt
08/09/2013 04:52 AM     113 Kazahana_grimmelshausen.wrd
```

E:_KAZE_Gamera_r17_LBL_schisch>type Kazahana_grimmelshausen.1.txt

```
0,000,106      grimmelshausen
```

```

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 2-1. txt
0,000,019      gri mmel shausen_s
0,000,005      gri mmel shausen_c
0,000,004      gri mmel shausen_had

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 2-2. txt
0,000,034      von_gri mmel shausen
0,000,006      by_gri mmel shausen
0,000,005      of_gri mmel shausen
0,000,005      i n_gri mmel shausen

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 3-1. txt
0,000,007      gri mmel shausen_s_simpl i ci ssi mus

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 3-2. txt
0,000,005      von_gri mmel shausen_c
0,000,004      of_gri mmel shausen_s
0,000,004      i n_gri mmel shausen_s

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 3-3. txt
0,000,023      chri stoffel _von_gri mmel shausen
0,000,005      chri stoph_von_gri mmel shausen
0,000,004      wri tten_by_gri mmel shausen
0,000,004      j akob_von_gri mmel shausen

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 4-1. txt

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 4-2. txt

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 4-3. txt

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 4-4. txt
0,000,014      j akob_chri stoffel _von_gri mmel shausen
0,000,008      j akob_chri stoffel _von_gri mmel shausen
0,000,004      h ans_j akob_von_gri mmel shausen

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 5-1. txt

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 5-2. txt

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 5-3. txt

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 5-4. txt

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. 5-5. txt
0,000,013      h ans_j akob_chri stoffel _von_gri mmel shausen
0,000,006      h ans_j akob_chri stoffel _von_gri mmel shausen

E: \_KAZE_Gamera_r17_LBL_schi sch>type Kazahana_gri mmel shausen. wrd
by
c
chri stoffel
chri stoph
gri mmel shausen
had
hans
i n
j akob
j akob
of
s
simpl i ci ssi mus
von
wri tten

E: \_KAZE_Gamera_r17_LBL_schi sch>

```

The number '21' (from schisch.log) tells how tall the pagoda is i.e. the total number of x-grams with 'gri mmel shausen'.
Oh, looking into 'Kazahana_gri mmel shausen. wrd' (it contains all distinct words occurring in x-grams) and seeing 'simpl i ci ssi mus' I recalled the needed book.
This search technique is so basic that not having it in your disposal makes the lameness worse.

Finally the 'gri mmel shausen' pagoda:

```

0,000,106      gri mmel shausen
0,000,019      gri mmel shausen_s
0,000,005      gri mmel shausen_c
0,000,004      gri mmel shausen_had
0,000,034      von_gri mmel shausen
0,000,006      by_gri mmel shausen
0,000,005      of_gri mmel shausen
0,000,005      i n_gri mmel shausen
0,000,007      gri mmel shausen_s_simpl i ci ssi mus
0,000,005      von_gri mmel shausen_c
0,000,004      of_gri mmel shausen_s
0,000,004      i n_gri mmel shausen_s
0,000,023      chri stoffel _von_gri mmel shausen
0,000,005      chri stoph_von_gri mmel shausen
0,000,004      wri tten_by_gri mmel shausen
0,000,004      j akob_von_gri mmel shausen
0,000,014      j akob_chri stoffel _von_gri mmel shausen
0,000,008      j akob_chri stoffel _von_gri mmel shausen
0,000,004      h ans_j akob_von_gri mmel shausen
0,000,013      h ans_j akob_chri stoffel _von_gri mmel shausen
0,000,006      h ans_j akob_chri stoffel _von_gri mmel shausen

```

It is true that pagoda houses treasures, moreover, they can be found relatively fast.
How two words are used together - a question arising every minute.

One of questions that is crying to be answered:

Are word A and word B in one neighborhood i.e. are they adjacent i.e. do they appear within an x-gram?

By answering it one could determine how 'appropriately' is to mix these two words within a phrase!
Not appearing together into one pagoda (i.e. under one roof) is a sign 'watch out', think twice then - you are in 'coining' mode.