

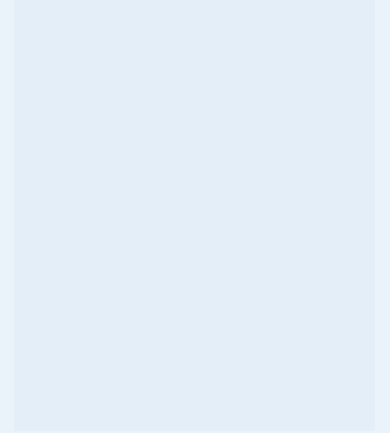


comp10002
Foundation of Algorithms
Assignment 1 – The Movie!
presented by Alistair Moffat

(c) 2022, The University of Melbourne

Starter Questions...

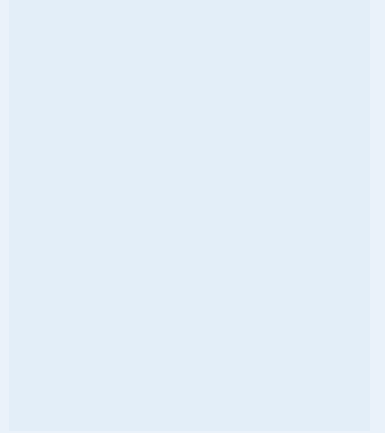
- What does the program do?
- What doesn't the program do?
- What's in the skeleton program?
- Where do I start?
- How should I ask questions?
- How will the marks be awarded?
- How do I submit? Can I submit more than once??
- What about Academic Honesty?



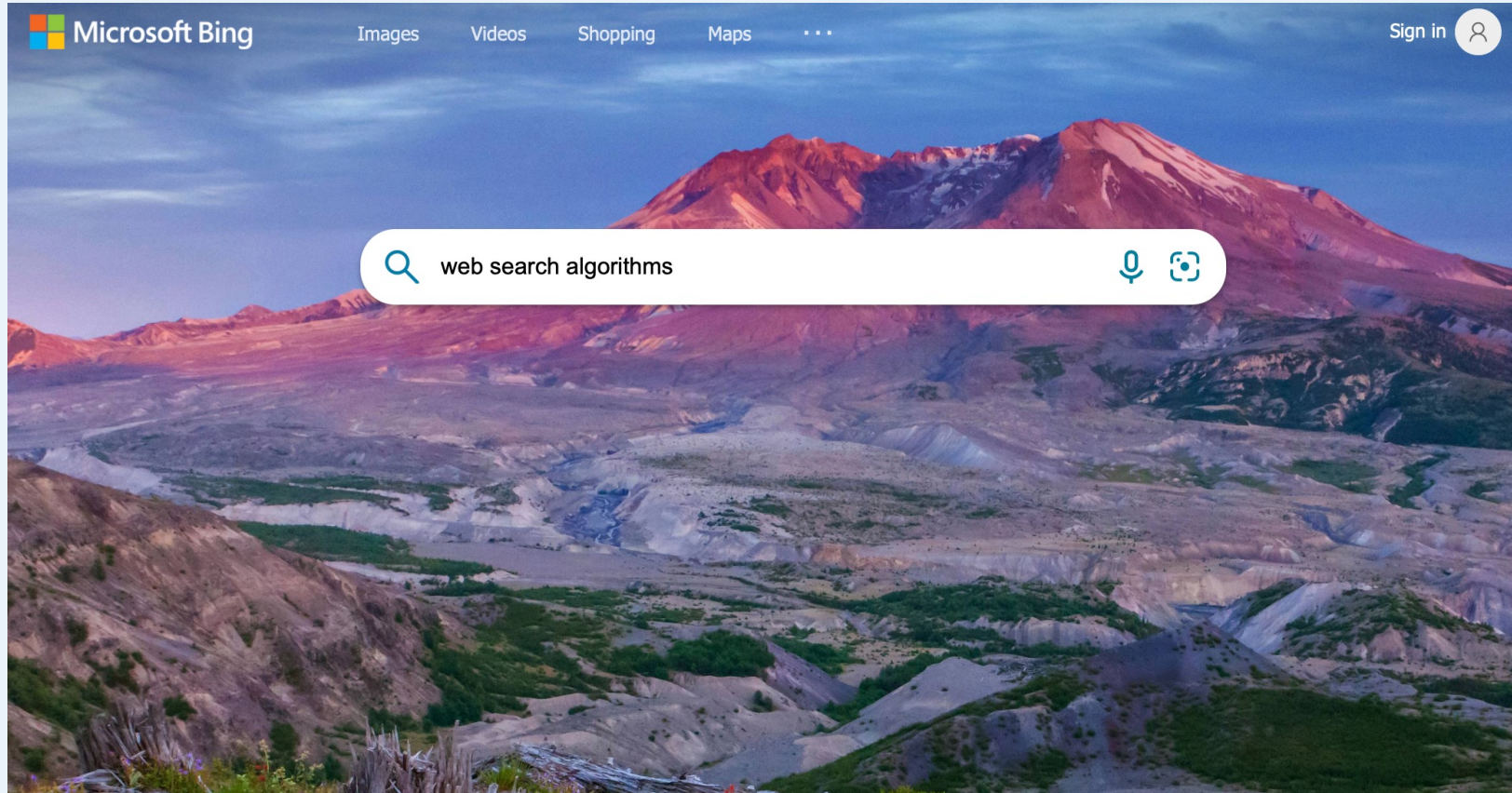
What does the program do?

Suppose you decide to learn about web search algorithms.

You go to your favorite web search service...



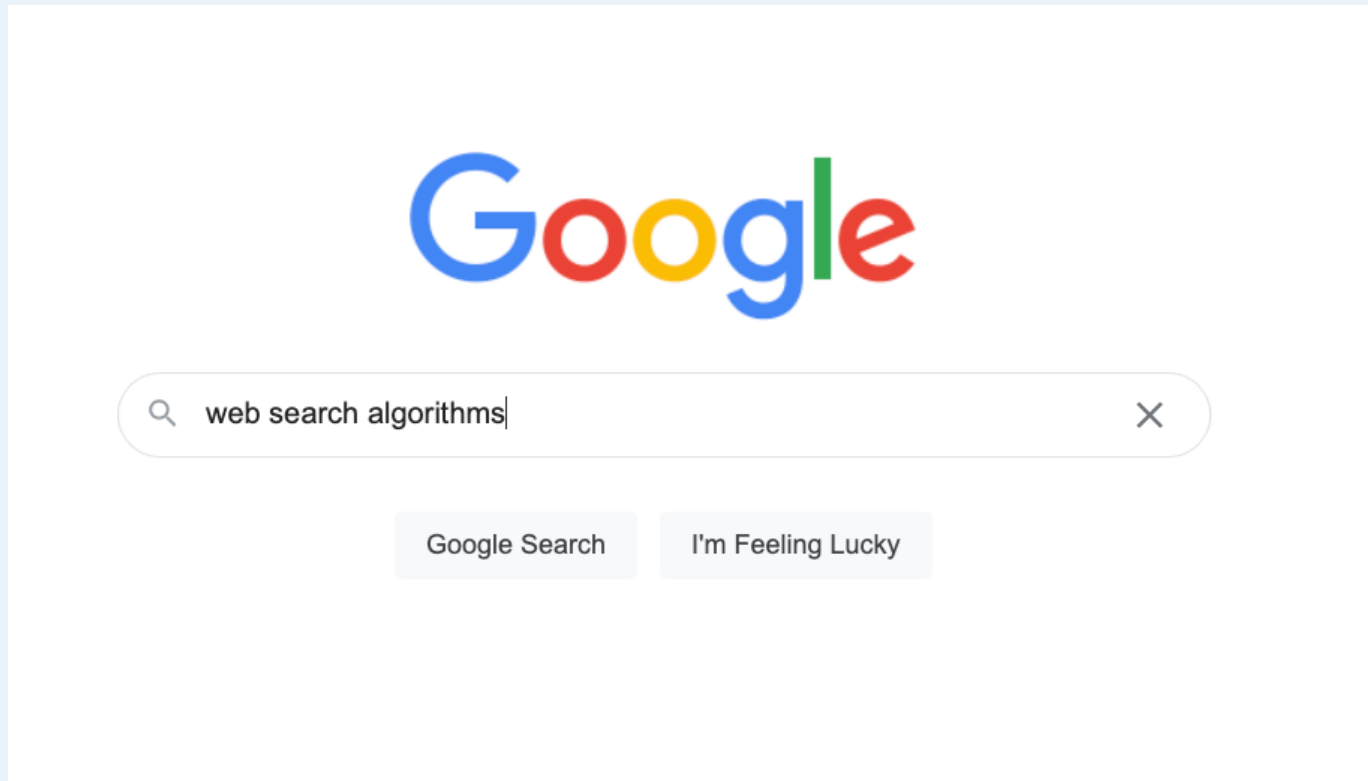
What does the program do?



What does the program do?



What does the program do?



What does the program do?

What Is an Internet Search Algorithm And How Does It Work?

<https://www.artzstudio.com/internet-search-algorithm> ▾

12/5/2020 · An Internet search algorithm is a set of instructions that describe the procedure to be followed in order to find a specific and concrete result in the network, within a larger data...

Learning Search Algorithms. Searching a page on the web ...



<https://medium.com/swlh/learning-search-algorithms-12aaaaea75df> ▾

7/7/2020 · An example of this type of search algorithm is Binary Search. Binary Search is also called half interval search because it divides the list into two-part after every step. So, let's...

Estimated Reading Time: 3 mins

Search Algorithms – Linear Search and Binary Search ...



<https://www.freecodecamp.org/news/search-algorithms-linear-and-binary-search-explained> ▾

11/1/2022 · Any algorithm which solves the search problem, namely, to retrieve information stored within some data structure, or calculated in the search space of a problem domain,...

Missing: ~~web-search~~ | Must include: [web search](#)

Everything You Must Know About Searching Algorithms



<https://codete.com/blog/everything-you-must-know-about-searching-algorithms> ▾

11/3/2021 · Here are the most common types of search algorithms in use today: linear search, binary search, jump search, interpolation search, exponential search, Fibonacci search....

What does the program do?

百度为您找到相关结果约15,500,000个

🔍 搜索工具

[Analysis of Web Search Algorithm Hits_图文 - 百度文库](#)

14页

Keywords: Analysis of **algorithms**, **Web search algorithms**. 1. Introduction It is observed that on the **Web** for a topic there are highly likely two sets of pages: An ...

百度文库  保障

[各类经典搜索算法\(Search Algorithms\)的简单介绍和C++..._...](#)

2012年4月21日 /// 斐波那契搜索法(Fbonacci **Search**) /// 定义: 利用斐波那契数列的性质(黄金分割原理)预测数据所在位置, 再以二分法方式逼近 /// 优点: 只涉及加法和减...

 CSDN技术社区 

[Deep-web search engine ranking algorithms - 百度学术](#)

Finding and fighting **search** engine spam. **Web** surfers rely on **search** engines to find information from the **web**. **Search** engine spam is the attempt to deceive **search** ...

百度学术 

[web search algorithms的中文翻译 - 百度翻译](#)

What does the program do?

<https://www.searchenginejournal.com> › search-engines ⋮

How Search Engine Algorithms Work: Everything You Need to ...

25 May 2020 — **Search algorithms** work as a large collection of other algorithms and formulas, each with its own purpose and task, to produce results a user ...

<https://shiftweb.com> › SEO ⋮

What is a Search Engine Algorithm and Why is it Important?

6 Oct 2012 — A **search engine algorithm** is a complex **algorithm** used by **search engines** such as Google, Yahoo, and Bing to determine a **web** page's ...

<https://www.brafton.com.au> › glossary › search-algorithm ⋮

Search Algorithm - Brafton

A **search algorithm** defines the process and factors used by a search company to rank websites and organize their search ranking with an accurate response to ...

<https://www.google.com> › search › ranking-results ⋮

Ranking Results – How Google Search Works

To give you the most useful information, **Search algorithms** look at many factors and signals, including the words of your query, relevance and usability of pages ...

What does the program do?

<https://www.searchenginejournal.com> › search-engines ⋮

How Search Engine Algorithms Work: Everything You Need to ...

25 May 2020 — **Search algorithms** work as a large collection of other algorithms and formulas, each with its own purpose and task, to produce results a user ...

<https://shiftweb.com> › SEO ⋮

What is a Search Engine Algorithm and Why is it Important?

6 Oct 2012 — A **search engine algorithm** is a complex **algorithm** used by **search engines** such as Google, Yahoo, and Bing to determine a **web** page's ...

<https://www.brafton.com.au> › glossary › search-algorithm ⋮

Search Algorithm - Brafton

A **search algorithm** defines the process and factors used by a search company to rank websites and organize their search ranking with an accurate response to ...

<https://www.google.com> › search › ranking-results ⋮

Ranking Results – How Google Search Works

To give you the most useful information, **Search algorithms** look at many factors and signals, including the words of your query, relevance and usability of pages ...

one snippet

What does the program do?

<https://www.searchenginejournal.com> › search-engines ⋮

How Search Engine Algorithms Work: Everything You Need to ...

25 May 2020 — **Search algorithms** work as a large collection of other algorithms and formulas, each with its own purpose and task, to produce results a user ...

<https://shiftweb.com> › SEO ⋮

What is a Search Engine Algorithm and Why is it Important?

6 Oct 2012 — A **search engine algorithm** is a complex **algorithm** used by **search engines** such as Google, Yahoo, and Bing to determine a **web** page's ...

<https://www.brafton.com.au> › glossary › search-algorithm ⋮

Search Algorithm - Brafton

A **search algorithm** defines the process and factors used by a search company to rank websites and organize their search ranking with an accurate response to ...

<https://www.google.com> › search › ranking-results ⋮

Ranking Results – How Google Search Works

To give you the most useful information, **Search algorithms** look at many factors and signals, including the words of your query, relevance and usability of pages ...

query terms
highlighted

What does the program do?

<https://www.searchenginejournal.com> › search-engines ⋮

How Search Engine Algorithms Work: Everything You Need to ...

25 May 2020 — **Search algorithms** work as a large collection of other algorithms and formulas, each with its own purpose and task, to produce results a user ...

<https://shiftweb.com> › SEO ⋮

What is a Search Engine Algorithm and Why is it Important?

6 Oct 2012 — A **search engine algorithm** is a complex **algorithm** used by **search engines** such as Google, Yahoo, and Bing to determine a **web** page's ...

<https://www.brafton.com.au> › glossary › search-algorithm ⋮

Search Algorithm - Brafton

A **search algorithm** defines the process and factors used by a search company to rank websites and organize their search ranking with an accurate response to ...

<https://www.google.com> › search › ranking-results ⋮

Ranking Results – How Google Search Works

To give you the most useful information, **Search algorithms** look at many factors and signals, including the words of your query, relevance and usability of pages ...

... and there's
more ...



What does the program do?

<https://www.searchenginejournal.com> › search-engines ⋮

How Search Engine Algorithms Work: Everything You Need to ...

25 May 2020 — **Search algorithms** work as a large collection of other algorithms and formulas, each with its own purpose and task, to produce results a user ...

<https://shiftweb.com> › SEO ⋮

What is a Search Engine Algorithm and Why is it Important?

6 Oct 2012 — A **search engine algorithm** is a complex **algorithm** used by **search engines** such as Google, Yahoo, and Bing to determine a **web** page's ...

<https://www.brafton.com.au> › glossary › search-algorithm ⋮

Search Algorithm - Brafton

A **search algorithm** defines the process and factors used by a search company to rank websites and organize their search ranking with an accurate response to ...

<https://www.google.com> › search › ranking-results ⋮

Ranking Results – How Google Search Works

To give you the most useful information, **Search algorithms** look at many factors and signals, including the words of your query, relevance and usability of pages ...

the set of
snippets is
formatted

What does the program do?

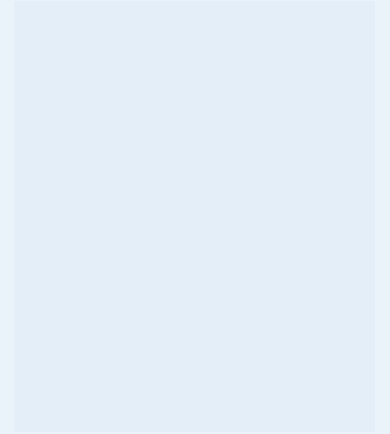
0. The skeleton code reads unformatted paragraphs of text

1. You have to count the words and paragraphs; then

2. Write each paragraph out as formatted text, with “query terms” from the command-line highlighted; then

3. Identify a best “snippet” to represent that paragraph, based on a set query terms from the command-line; then

4. At end of paragraphs, print “ta daa!”



What does the program do?

Alice was beginning to get very tired
of sitting by her sister on the
bank, and of having nothing to do: once or twice
she had peeped into
the book her sister
was reading, but it had no pictures or conversations
in it, 'and what is the use of a book,' thought
Alice
'without pictures
or conversation?'

So she was considering in her own mind (as well
as she could, for the
hot day made her feel very sleepy and stupid),
whether the
pleasure of
making a daisy-chain would be worth the trouble of getting
up and picking
the daisies, when suddenly a White Rabbit with pink eyes
ran close by
her.



Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice

What does the program do?

Alice was beginning to get very tired
of sitting by her sister on the
bank, and of having nothing to do: once or twice
she had peeped into
the book her sister
was reading, but it had no pictures or conversation
in it, 'and what is the use of a book,' thought
Alice
'without pictures
or conversation?'

So she was considering in her own mind (as well
as she could, for the
hot day made her feel very sleepy and stupid),
whether the
pleasure of
making a daisy-chain would be worth the trouble of getting
up and picking
the daisies, when suddenly a White Rabbit with pink eyes
ran close by
her.

double newline
ends paragraph



Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice

What does the program do?

Alice was beginning to get very tired
of sitting by her sister on the
bank, and of having nothing to do: once or twice
she had peeped into
the book her sister
was reading, but it had no pictures or conversations
in it, 'and what is the use of a book,' thought
Alice
'without pictures
or conversation?'

So she was considering in her own mind (as well
as she could, for the
hot day made her feel very sleepy and stupid),
whether the
pleasure of
making a daisy-chain would be worth the trouble of getting
up and picking
the daisies, when suddenly a White Rabbit with pink eyes
ran close by
her.

some trailing
punctuation
is retained

Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice

What does the program do?

Alice was beginning to get very tired
of sitting by her sister on the
bank, and of having nothing to do: once or twice
she had peeped into
the book her sister
was reading, but it had no pictures or conversations
in it, and what is the use of a book, thought
Alice
without pictures
or conversation?

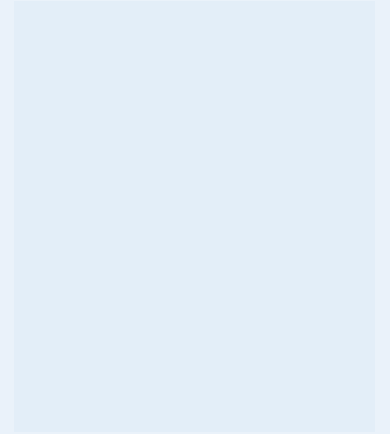
So she was considering in her own mind (as well
as she could, for the
hot day made her feel very sleepy and stupid),
whether the
pleasure of
making a daisy-chain would be worth the trouble of getting
up and picking
the daisies, when suddenly a White Rabbit with pink eyes
ran close by
her.

digits, other
punctuation, and
extra whitespace
are deleted

Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice

What does the program do?

0. The skeleton code reads unformatted paragraphs of text
1. You have to count the words and paragraphs; then
2. Write each paragraph out as formatted text, with “query terms” from the command-line highlighted; then
3. Identify a best “snippet” to represent that paragraph, based on a set query terms from the command-line; then
4. At end of paragraphs, print “ta daa!”



What does the program do?

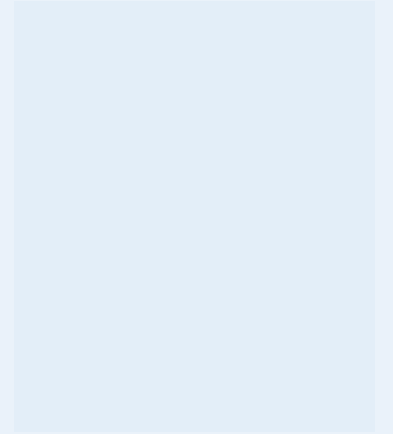
Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice



```
mac: ass1-soln < data1.txt  
  
===== Stage 1 [para 1; 57 words]  
  
===== Stage 1 [para 2; 56 words]  
  
===== Stage 1 [para 3; 144 words]  
  
===== Stage 1 [para 4; 21 words]  
  
===== Stage 1 [para 5; 40 words]
```

What does the program do?

0. The skeleton code reads unformatted paragraphs of text
1. You have to count the words and paragraphs; then
2. Write each paragraph out as formatted text, with “query terms” from the command-line highlighted; then
3. Identify a best “snippet” to represent that paragraph, based on a set query terms from the command-line; then
4. At end of paragraphs, print “ta daa!”



What does the program do?

Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice



```
mac: ass1-soln < data1.txt
```

```
=====  
Stage 1 [para 1; 57 words]
```

```
=====  
Stage 2 [para 1; 57 words; 0 matches]
```

```
Alice was beginning to get very tired of sitting by her sister on the  
bank, and of having nothing to do: once or twice she had peeped into the  
book her sister was reading, but it had no pictures or conversations in  
it, and what is the use of a book, thought Alice without pictures or  
conversation?
```



72 characters

What does the program do?

Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice



```
mac: ass1-soln pictures < data1.txt
```

```
==== Stage 1 [para 1; 57 words]
```

```
==== Stage 2 [para 1; 57 words; 2 matches]
```

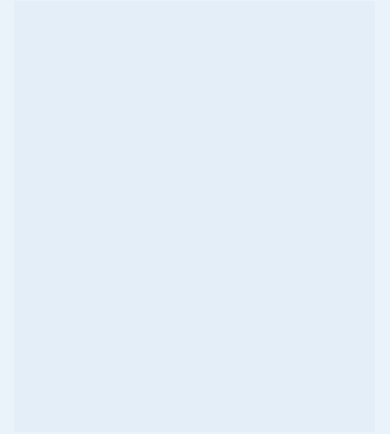
Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no ****pictures**** or conversations in it, and what is the use of a book, thought Alice without ****pictures**** or conversation?



72 characters

What does the program do?

0. The skeleton code reads unformatted paragraphs of text
1. You have to count the words and paragraphs; then
2. Write each paragraph out as formatted text, with “query terms” from the command-line highlighted; then
3. Identify a best “snippet” to represent that paragraph, based on a set query terms from the command-line; then
4. At end of paragraphs, print “ta daa!”



What does the program do?

Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice



```
mac: ass1-soln < data1.txt
```

```
=====  
Stage 1 [para 1; 57 words]
```

```
=====  
Stage 2 [para 1; 57 words; 0 matches]
```

```
Alice was beginning to get very tired of sitting by her sister on the  
bank, and of having nothing to do: once or twice she had peeped into the  
book her sister was reading, but it had no pictures or conversations in  
it, and what is the use of a book, thought Alice without pictures or  
conversation?
```

```
=====  
Stage 3 [para 1; start 0; length 21; score 2.30]
```

```
Alice was beginning to get very tired of sitting by her sister on the  
bank, and of having nothing to do:...
```

What does the program do?

Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice



```
mac: ass1-soln pictures < data1.txt

===== Stage 1 [para 1; 57 words]

===== Stage 2 [para 1; 57 words; 2 matches]
Alice was beginning to get very tired of sitting by her sister on the
bank, and of having nothing to do: once or twice she had peeped into the
book her sister was reading, but it had no **pictures** or conversations
in it, and what is the use of a book, thought Alice without **pictures**
or conversation?

===== Stage 3 [para 1; start 34; length 23; score 5.94]
but it had no **pictures** or conversations in it, and what is the use
of a book, thought Alice without **pictures** or conversation?
```

the snippet scoring rules are in the handout

What does the program do?

Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice



```
mac: ass1-soln twice peeped book < data1.txt
```

```
==== Stage 1 [para 1; 57 words]
```

```
==== Stage 2 [para 1; 57 words; 4 matches]
```

Alice was beginning to get very tired of sitting by her sister on the bank, and of having nothing to do: once or ****twice**** she had ****peeped**** into the ****book**** her sister was reading, but it had no pictures or conversations in it, and what is the use of a ****book****, thought Alice without pictures or conversation?

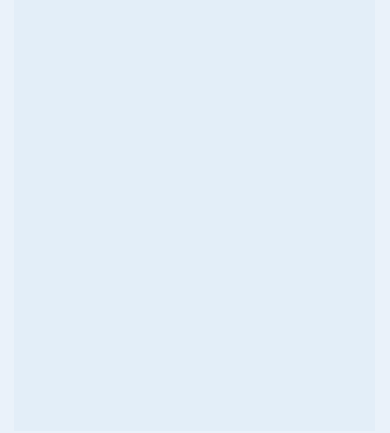
```
==== Stage 3 [para 1; start 21; length 30; score 8.88]
```

once or ****twice**** she had ****peeped**** into the ****book**** her sister was reading, but it had no pictures or conversations in it, and what is the use of a ****book****,...

you have to generate the highest-scoring
legal snippet from each input paragraph

What does the program do?

0. The skeleton code reads unformatted paragraphs of text
1. You have to count the words and paragraphs; then
2. Write each paragraph out as formatted text, with “query terms” from the command-line highlighted; then
3. Identify a best “snippet” to represent that paragraph, based on a set query terms from the command-line; then
4. At end of paragraphs, print “ta daa!”



What does the program do?

Alice
was
beginning
to
get
very
tired
of
sitting
by
her
sister
on
the
bank,
and
of
having
nothing
to
do:
once
or
twice



```
mac: ass1-soln algorithms are fun < data1.txt
```

```
==== Stage 1 [para 1; 57 words]
```

```
==== Stage 2 [para 1; 57 words; 0 matches]
```

```
Alice was beginning to get very tired of sitting by her sister on the  
bank, and of having nothing to do: once or twice she had peeped into the  
book her sister was reading, but it had no pictures or conversations in  
it, and what is the use of a book, thought Alice without pictures or  
conversation?
```

```
==== Stage 3 [para 1; start 0; length 21; score 2.30]
```

```
Alice was beginning to get very tired of sitting by her sister on the  
bank, and of having nothing to do:...
```

```
ta daa!
```

```
mac:
```


What doesn't the program do?

No need for any “defensiveness” in input processing, and you may assume that all input will always be well-formed.

No need to implement any *stemming* rules (“algorithm” = “algorithms” = “algorithmic” = “algorithmed” = ...), just use the exact terms from command-line.

What doesn't the program do?

Your program must not use files (Chapter 11), and all input is to be read from `stdin` in the normal manner.

All programs will be tested via command-line redirection:

```
mac: myass1-soln terma termb termc < data1.txt
```

If your program varies from this expectation, substantial mark penalties will apply.

What doesn't the program do?

Structs (Chapter 8) may be used; but they are not required, and there is no penalty for not using them. (They'll be required in Assignment 2.)

Your program must not use `malloc()` (Chapter 10), and all arrays are to be declared in advance. (It will be required in Assignment 2.)

What's in the skeleton program?

Two important starting elements in the skeleton program:

- An **Authorship Declaration**. You must **include it** in your submitted program, and you must **“sign” it**. There will be significant mark penalties (see the rubric) for non-compliance.
- A modified version of the function `getword()`

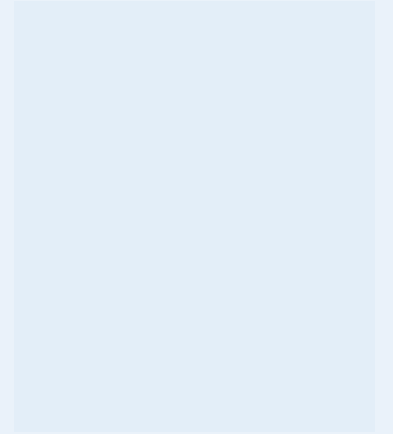
Where do I start?

Go to “Assignments -> Assignment 1”.

The handout and skeleton program are linked from there.

There is also an Assignment 1 movie, and these slides too.

Plus some test data, the marking rubric, and a link to a page of hints and tips.



How should I ask questions?

An “Assignment 1” Discussion Forum is available.

All questions should be asked there. Check back a few hours later for responses. Read other questions/answers while you wait.

Requests for Special Consideration or extensions should be emailed to Alistair, ammoffat@unimelb.edu.au, accompanied by documentary support.

How will the marks be awarded?

There is a marking rubric linked from the Assignment 1 page,
read it carefully!

Compiler warnings will be penalized; output discrepancies will be penalized.

Late submission will be penalized.

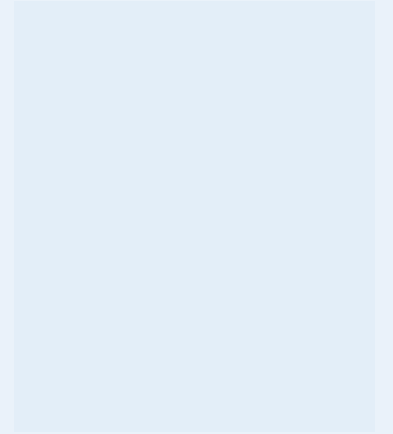
Missing Authorship Declarations (and non-compliance with what it specifies) will be **heavily** penalized.

How do I submit? Is it once only??

Submission is **via the LMS**, go to the bottom of the LMS page “Assignments -> Assignment 1”. You may submit multiple times to check the test environment.

Deadline: **6pm, 16 September**.

After that moment strict late penalties will apply.



How do I submit? Is it once only??

You may write your program using Grok or any other C programming toolkit.

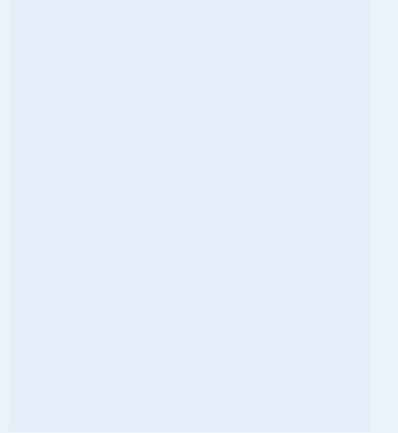
But submission must be via the LMS Gradescope tool.

Responsibility for checking that compilation was successful is with **you**.

What about Academic Honesty?

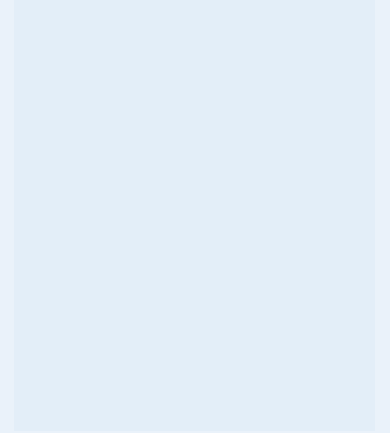
- This is an **individual** project
- You must not exchange written or electronic material, even “just to have a look”, or “just to get some ideas”
- All submissions will be compared against all other submissions. If you copy you **will** get detected, and then action **will** be taken
- If you are smart enough to copy and **not** get detected, you are smart enough to do the project by yourself
- **DON'T. EXCHANGE. ANY. MATERIAL.**
- Don't “publish” your solutions, be protective of your work

Will you show us
your program??



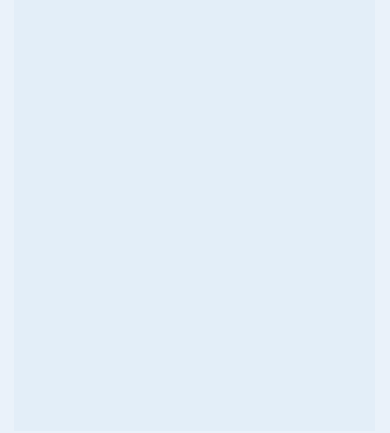
Will you show us
your program??

Please, oh kind, all-wise,
all-happiness, all-tadaa,



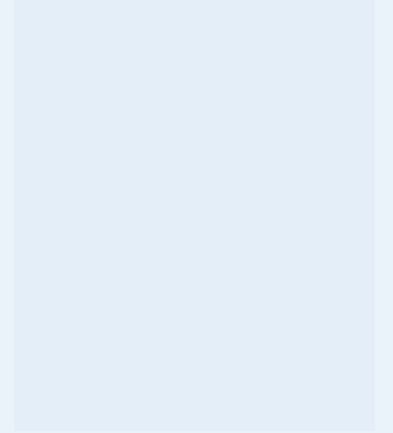
Will you show us your program??

Please, oh kind, all-wise,
all-happiness, all-tadaa,
masterful, wonderful,
benevolent, source-of-
incomprehensible fun,



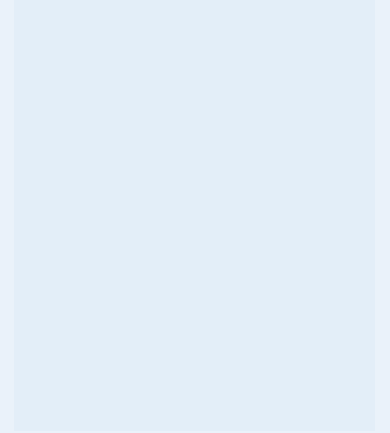
Will you show us
your program??

Please, oh kind, all-wise,
all-happiness, all-tadaa,
masterful, wonderful,
benevolent, source-of-
incomprehensible fun,
Sir?



Will you show us
your program??

Please, oh kind, all-wise,
all-happiness, all-tadaa,
masterful, wonderful,
benevolent, source-of-
incomprehensible fun,
Sir?



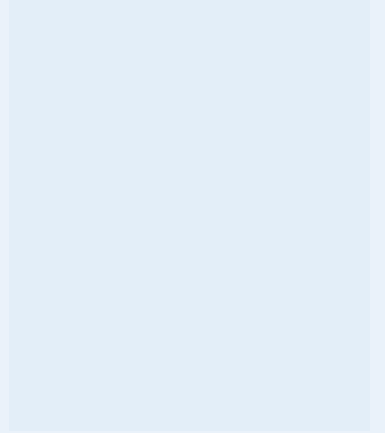
Will you show us your program??

Please, oh kind, all-wise,
all-happiness, all-tadaa,
masterful, wonderful,
benevolent, source-of-
incomprehensible fun,
Sir?

```
mac: wc -l ass1*.c
148  ass1-skel.c
532  ass1-soln.c
```

Finally, remember...

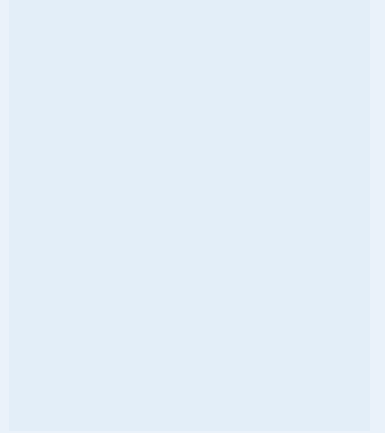
The primary objective is to learn how to manipulate arrays and strings using C.



Finally, remember...

The primary objective is to learn how to manipulate arrays and strings using C.

The secondary objective is to have ...



Finally, remember...

The primary objective is to learn how to manipulate arrays and strings using C.

The secondary objective is to have ...

FUN!



comp10002
Foundation of Algorithms
Assignment 1 – The Movie!
presented by Alistair Moffat

(c) 2022, The University of Melbourne