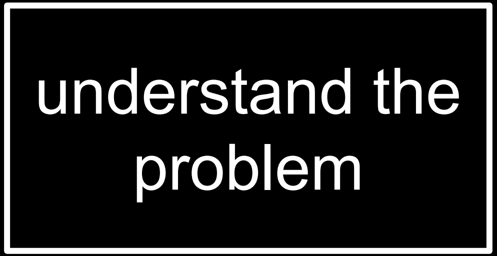
PROBLEM SOLVING

BACK

Polya's approach to problem-solving

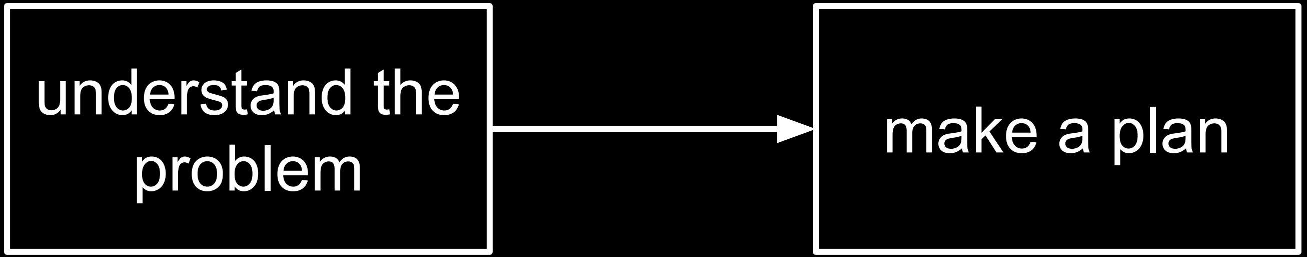
image source: http://doi.org/10.3932/ethz-a-000099441

Polya's approach to problem-solving



understand the problem

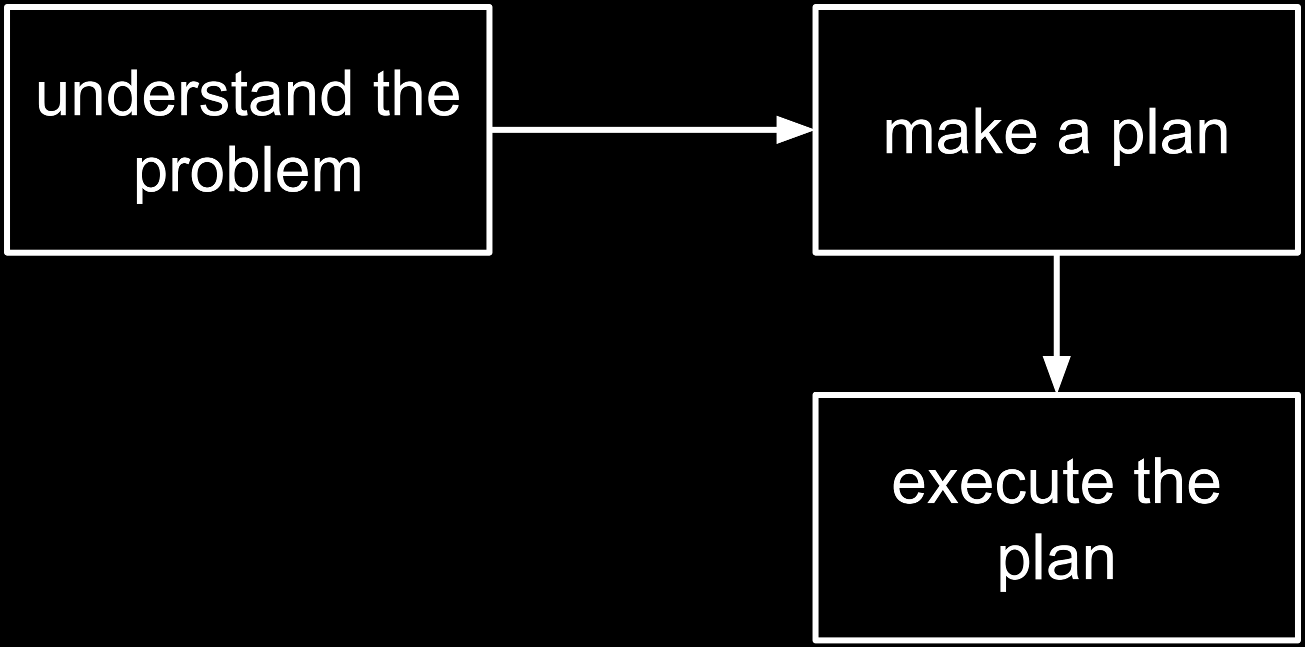
Polya's approach to problem-solving



understand the problem

make a plan

Polya's approach to problem-solving

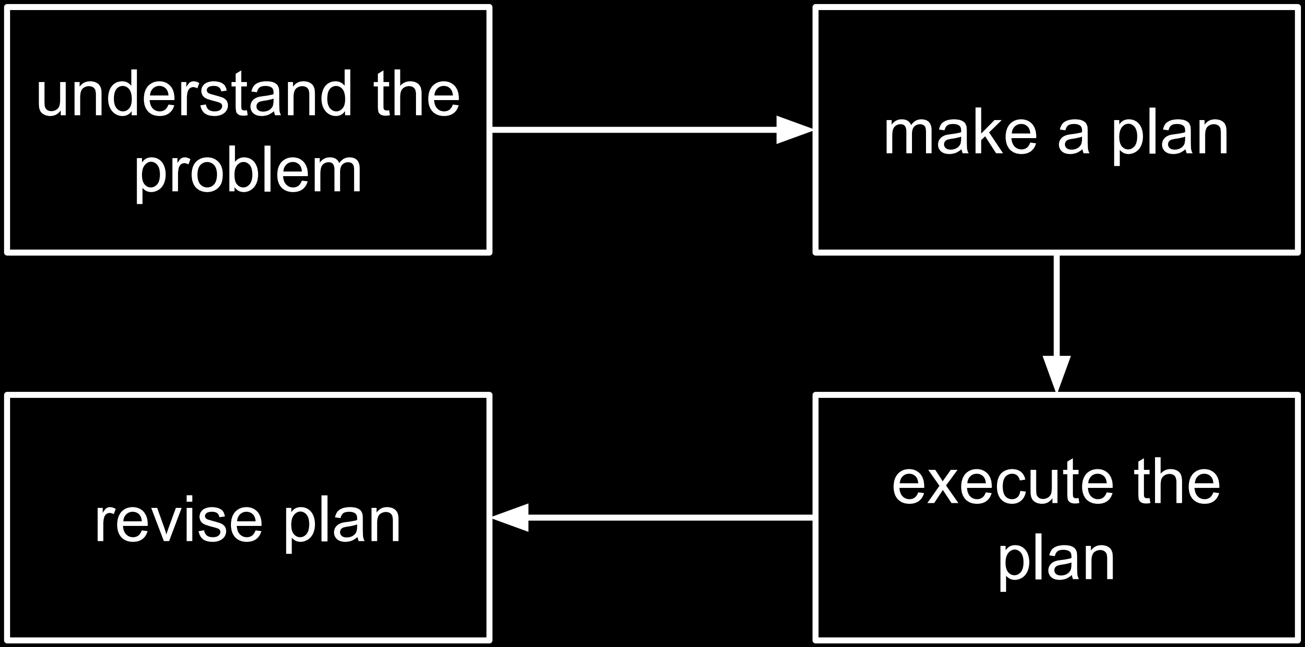


understand the problem

make a plan

execute the plan

Polya's approach to problem-solving



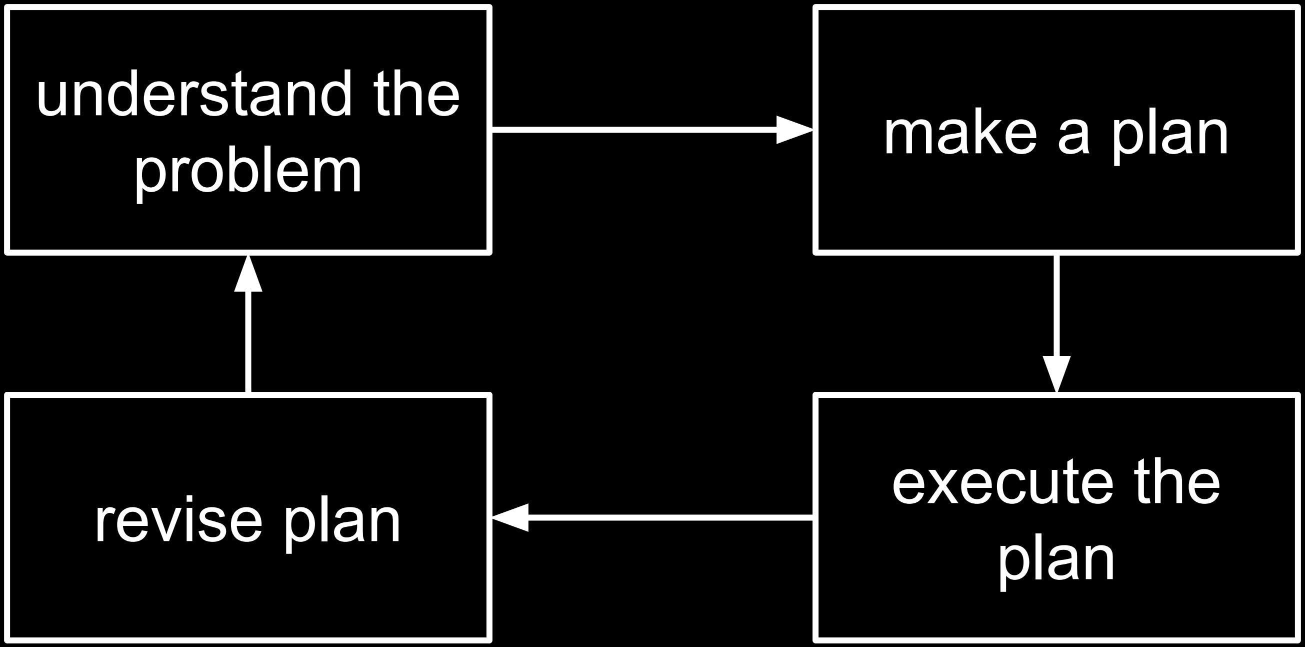
understand the problem

make a plan

execute the plan

revise plan

Polya's approach to problem-solving



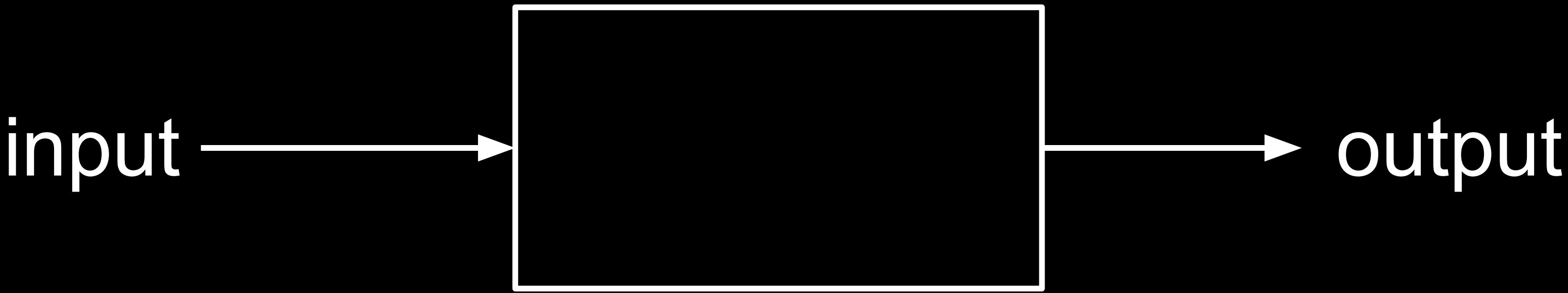
understand the problem

make a plan

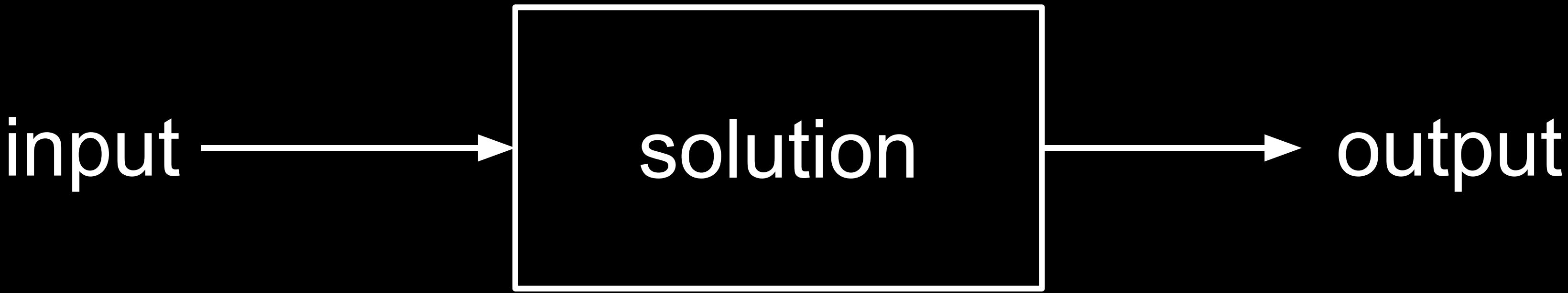
execute the plan

revise plan

a model to represent problems



a model to represent problems

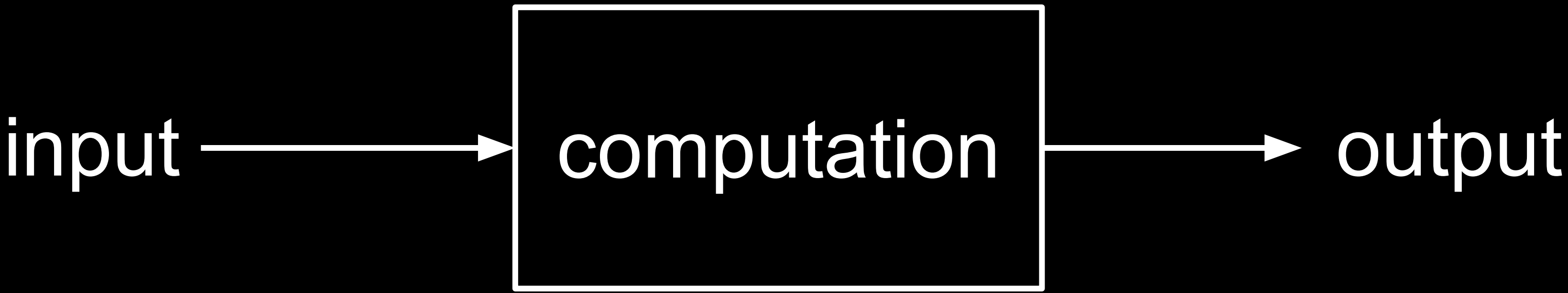


input

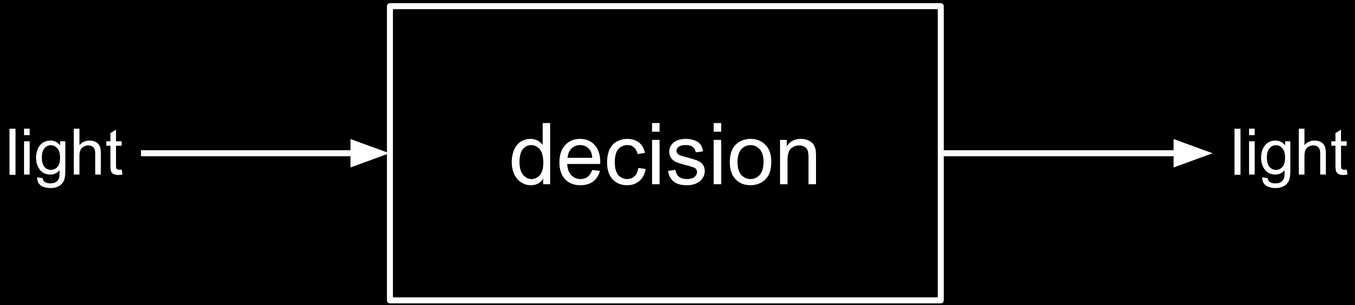
solution

output

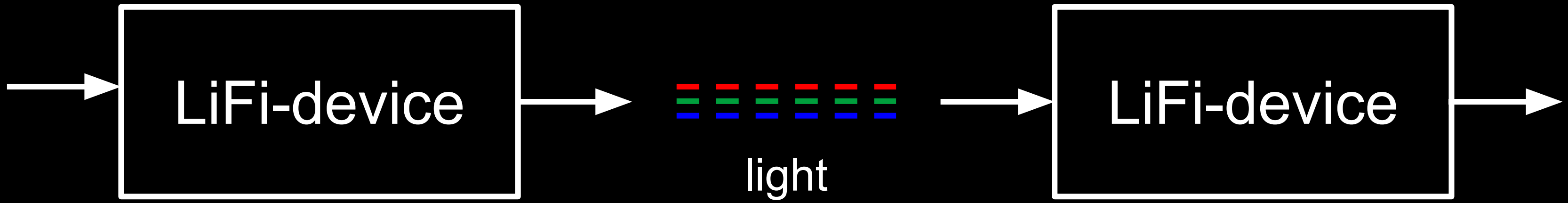
a model to represent problems

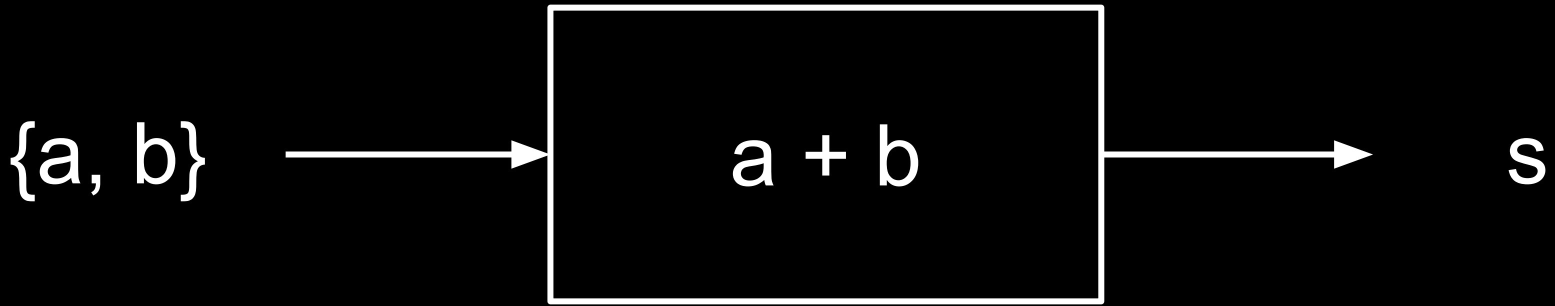


the LiFi-project as an input - processing - output - problem



the LiFi-project as a communication problem

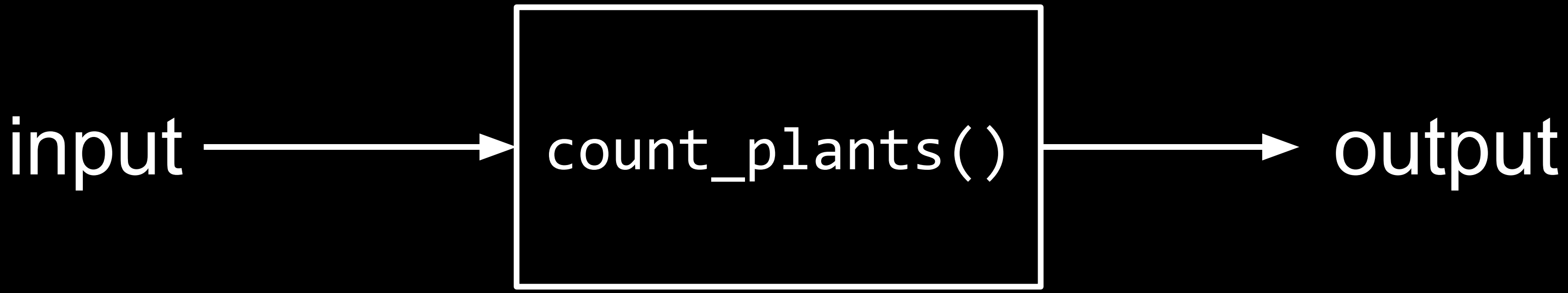




{a, b}

a + b

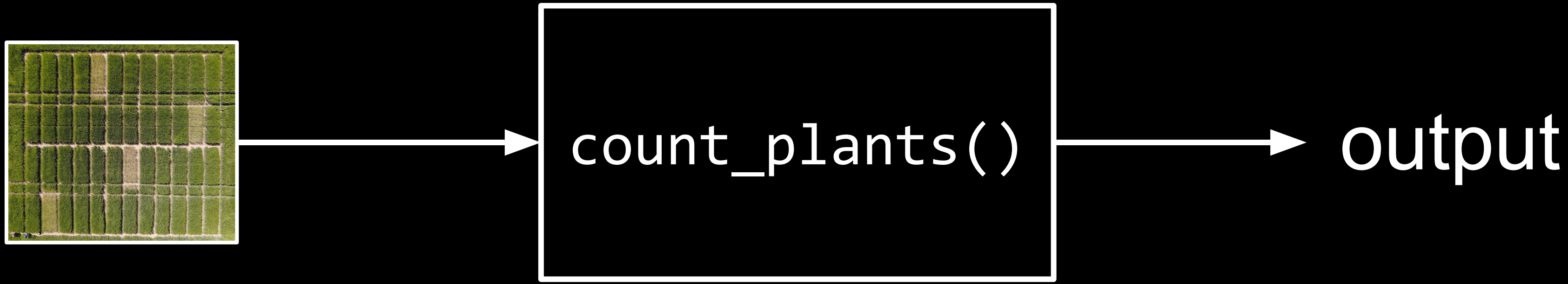
s



input

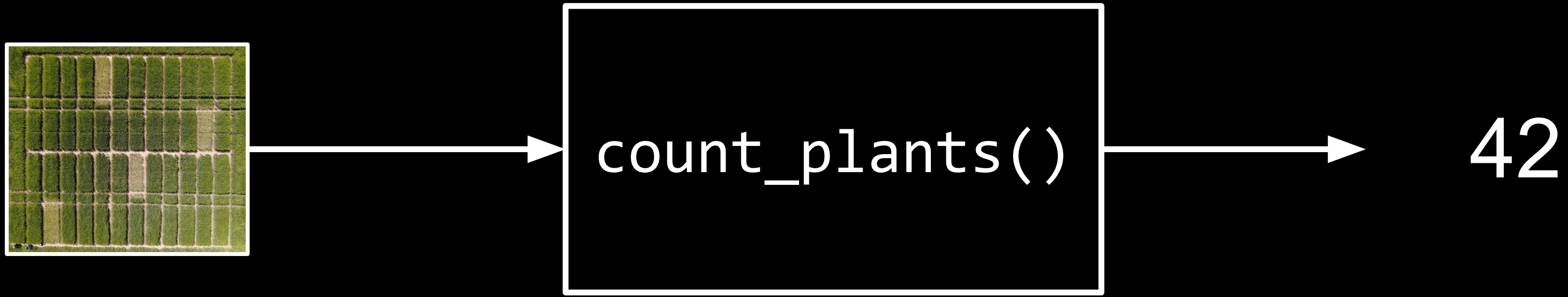
`count_plants()`

output



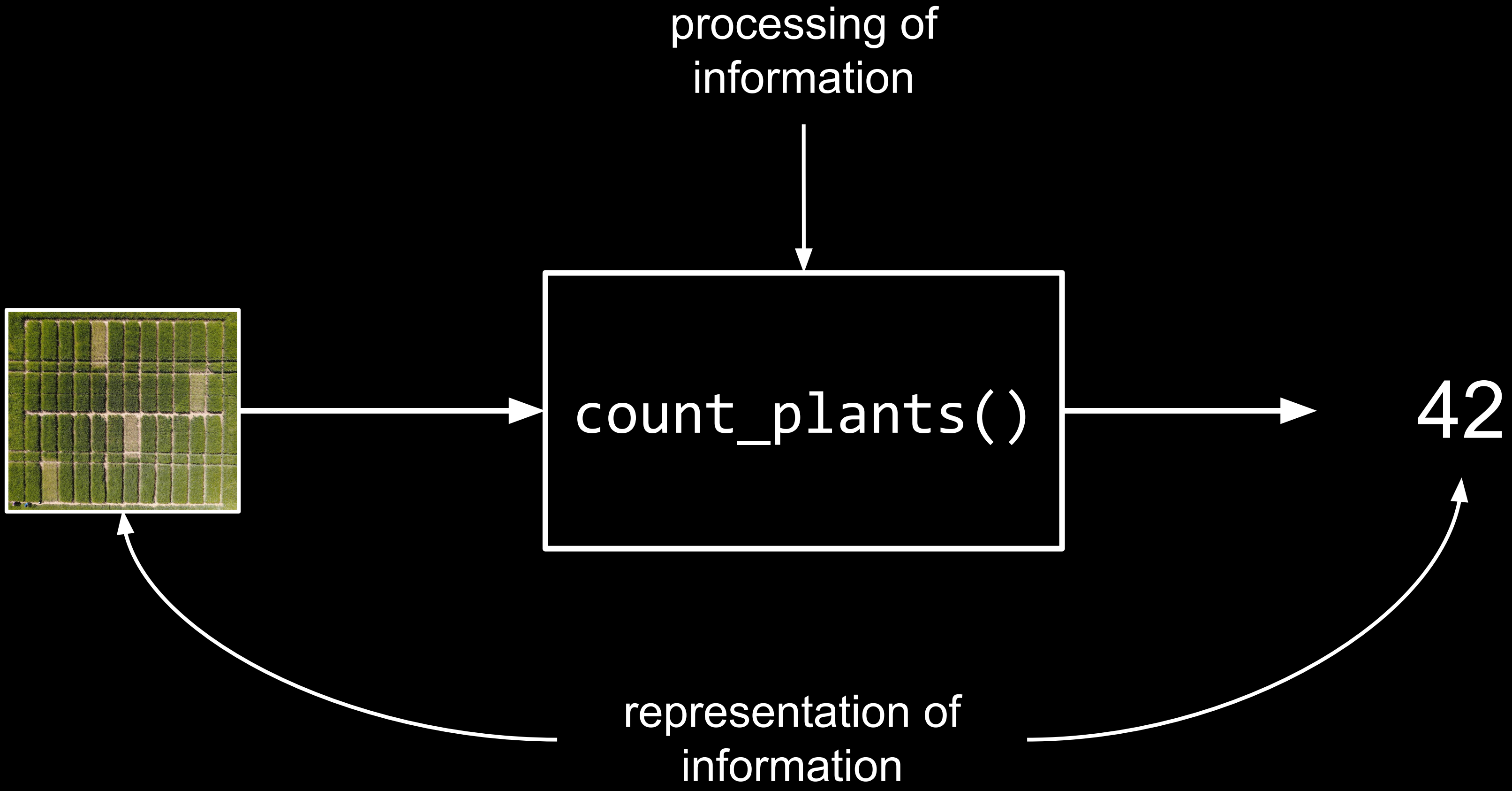
count_plants()

output



count_plants()

42

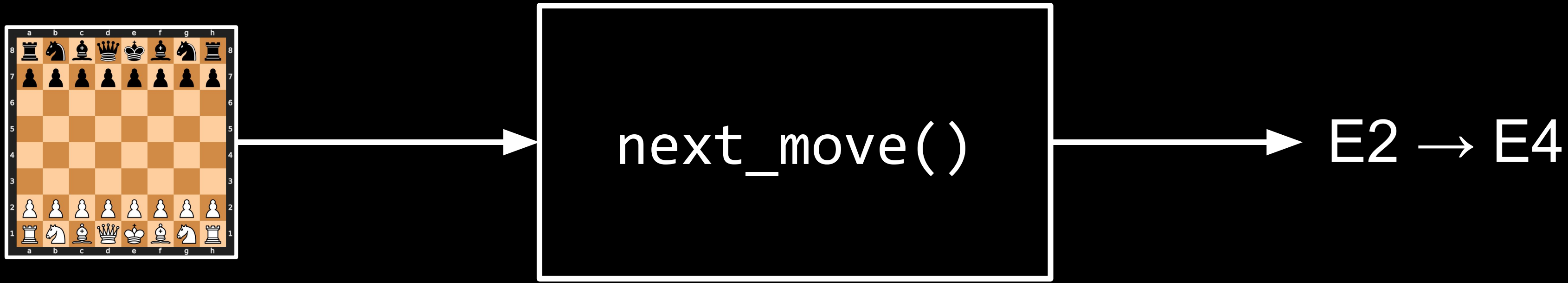


processing of information

count_plants()

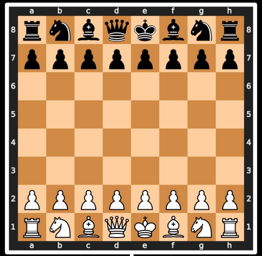
42

representation of information



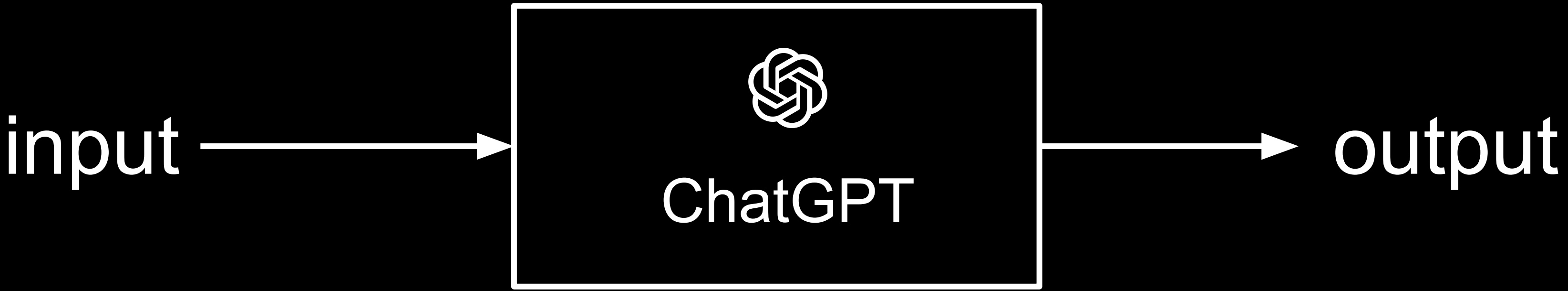
next_move()

E2 → E4



```
1: 0R          9: 0P            57: 1R
2: 0N         10: 0P            58: 1N
3: 0B         11: 0P            59: 1B
4: 0K         12: 0P            60: 1K
5: 0Q         13: 0P     ...    61: 1Q
6: 0B         14: 0P            62: 1B
7: 0N         15: 0P            63: 1N
8: 0R         16: 0P            64: 1R
```

representation of information



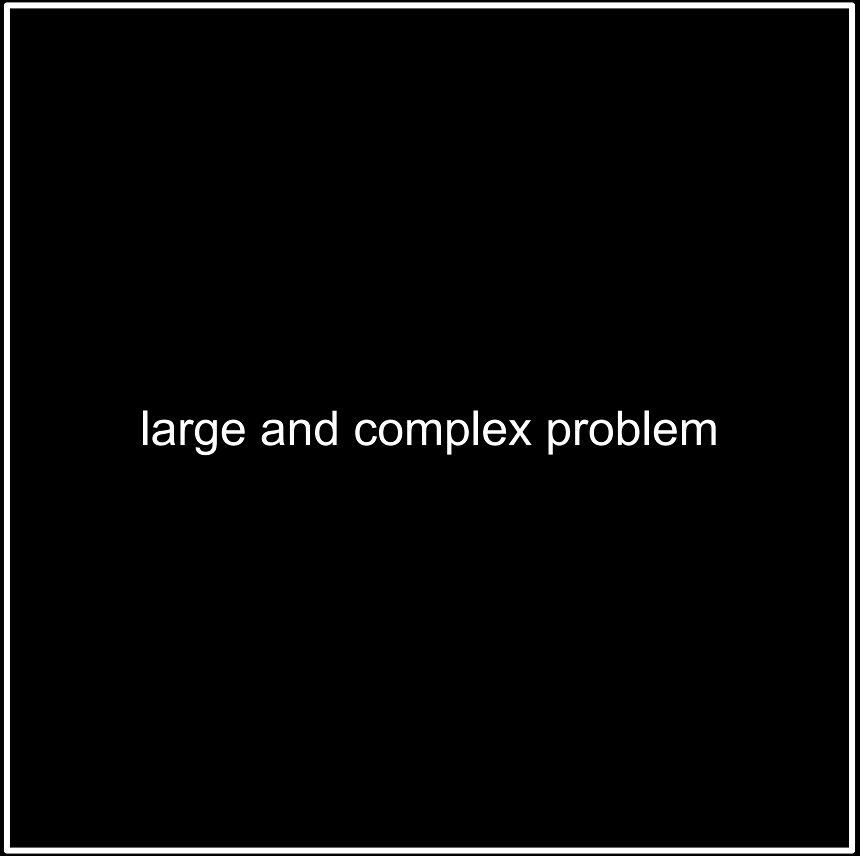
input

ChatGPT

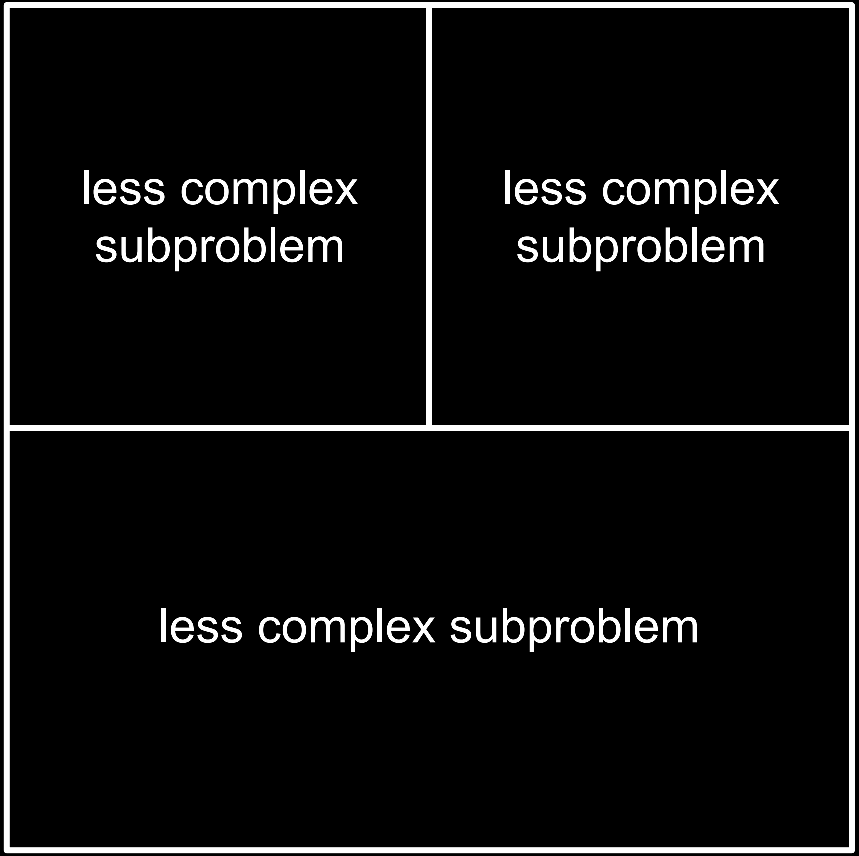
output

problem solving strategies

problem decomposition



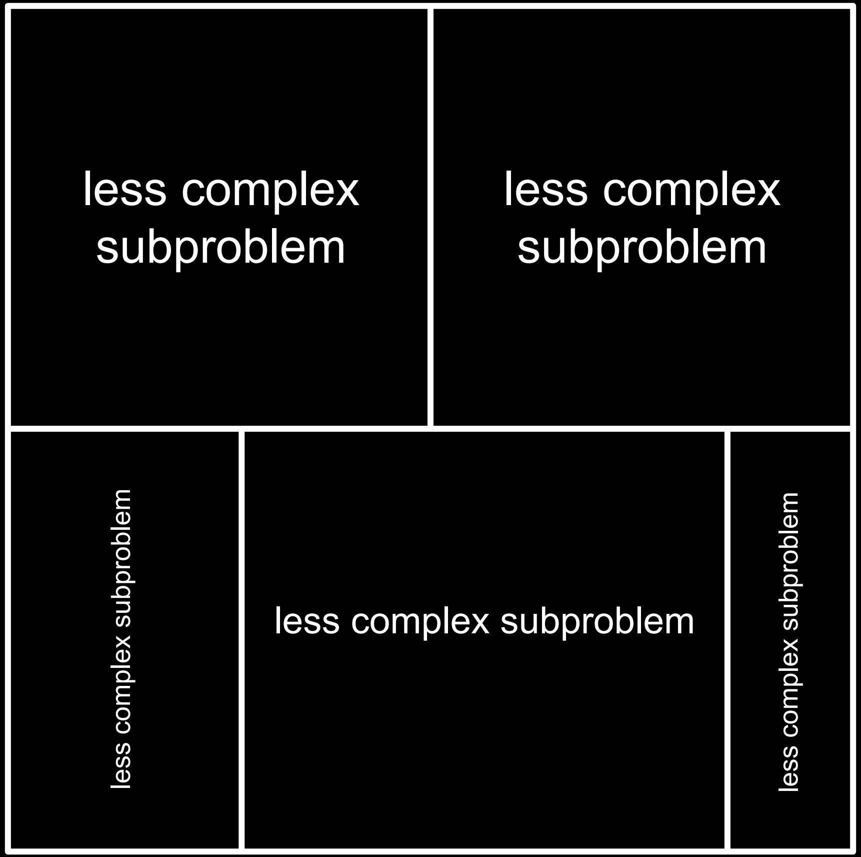
large and complex problem



less complex subproblem

less complex subproblem

less complex subproblem



less complex subproblem

less complex subproblem

less complex subproblem

less complex subproblem

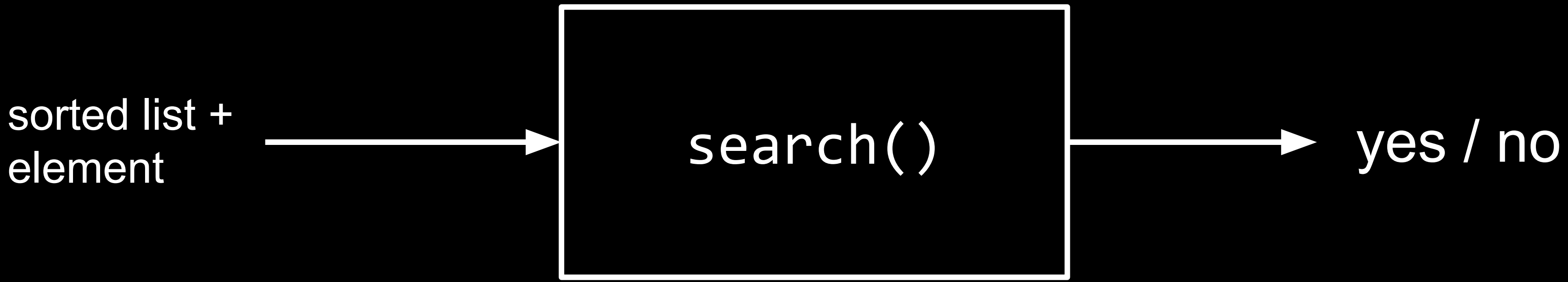
less complex subproblem

divide and conquer

large and complex problem of type A

smaller problem of type A	smaller problem of type A
smaller problem of type A	smaller problem of type A

even smaller problem of type A	even smaller problem of type A
even smaller problem of type A	even smaller problem of type A
even smaller problem of type A	even smaller problem of type A
even smaller problem of type A	even smaller problem of type A



sorted list +
element

`search()`

yes / no

is 67 a prime number?

```
2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97
```

linear search

⬇

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

linear search

↓

~~2~~, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

linear search

⬇

~~2~~, ~~3~~, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

linear search

~~2~~, ~~3~~, ~~5~~, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

linear search

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
~~43~~, ~~47~~, ~~53~~, ~~59~~, ~~61~~, 67, 71, 73, 79, 83, 89, 97

⬆

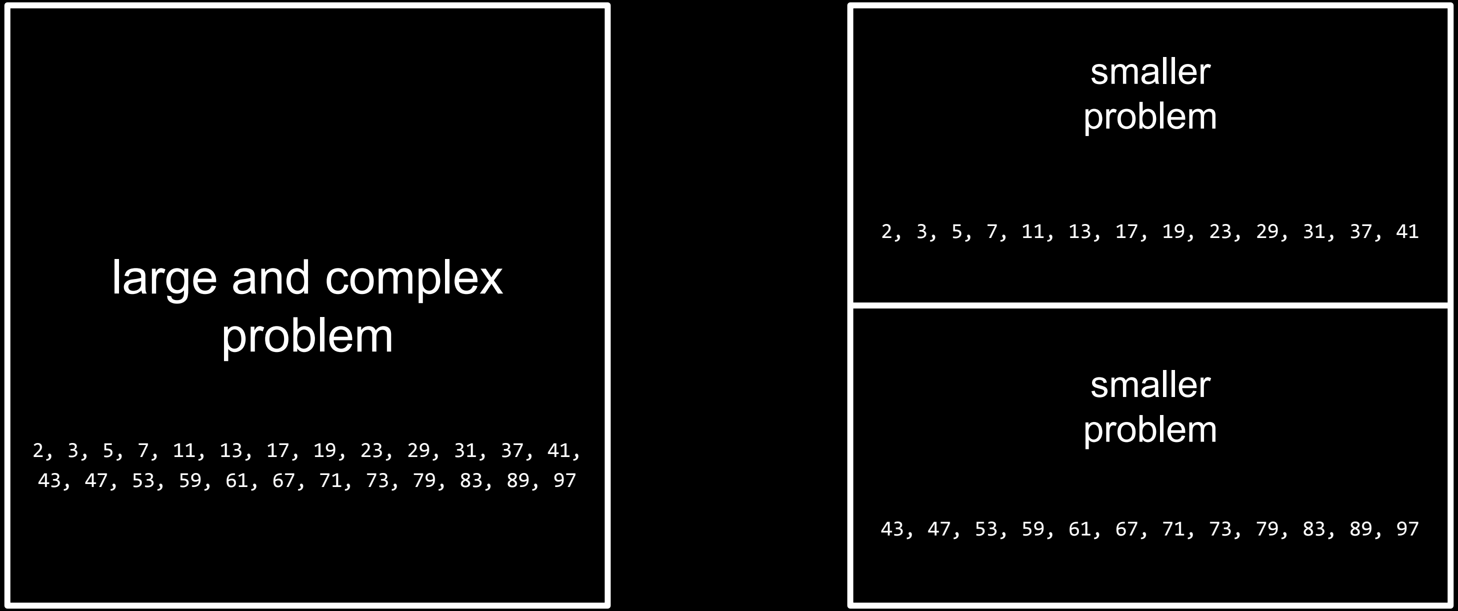
19 steps… can’t we do better?

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
~~43~~, ~~47~~, ~~53~~, ~~59~~, ~~61~~, 67, 71, 73, 79, 83, 89, 97

⬆

large and complex problem

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97



large and complex problem

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

smaller problem

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41

smaller problem

43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

binary search

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

binary search

67 != 41

⬇

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, ~~41~~,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

binary search

67 > 41

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, ~~41~~,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

binary search

67 > 41

⬇

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

binary search

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97

⬆

67 != 71

binary search

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
43, 47, 53, 59, 61, 67, ~~71~~, 73, 79, 83, 89, 97

⬆

67 != 71

binary search

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
43, 47, 53, 59, 61, 67, ~~71~~, ~~73~~, ~~79~~, ~~83~~, ~~89~~, ~~97~~

⬆

67 < 71

binary search

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
43, 47, 53, ~~59~~, 61, 67, ~~71~~, ~~73~~, ~~79~~, ~~83~~, ~~89~~, ~~97~~

⬆

67 != 59

binary search

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
~~43~~, ~~47~~, ~~53~~, ~~59~~, 61, 67, ~~71~~, ~~73~~, ~~79~~, ~~83~~, ~~89~~, ~~97~~

⬆

67 > 59

binary search

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
~~43~~, ~~47~~, ~~53~~, ~~59~~, 61, 67, ~~71~~, ~~73~~, ~~79~~, ~~83~~, ~~89~~, ~~97~~

⬆

67 = 67

binary search

~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
~~43~~, ~~47~~, ~~53~~, ~~59~~, ~~61,~~ 67, ~~71~~, ~~73~~, ~~79~~, ~~83~~, ~~89~~, ~~97~~

↑

67 = 67

3 splits → much better

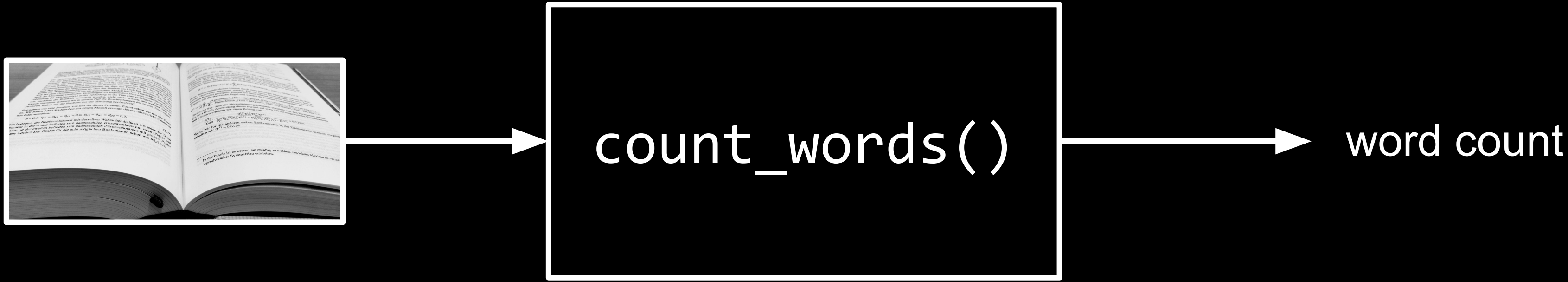
~~2~~, ~~3~~, ~~5~~, ~~7~~, ~~11~~, ~~13~~, ~~17~~, ~~19~~, ~~23~~, ~~29~~, ~~31~~, ~~37~~, ~~41~~,
~~43~~, ~~47~~, ~~53~~, ~~59~~, ~~61,~~ 67, ~~71~~, ~~73~~, ~~79~~, ~~83~~, ~~89~~, ~~97~~

⬆

67 = 67

🤔

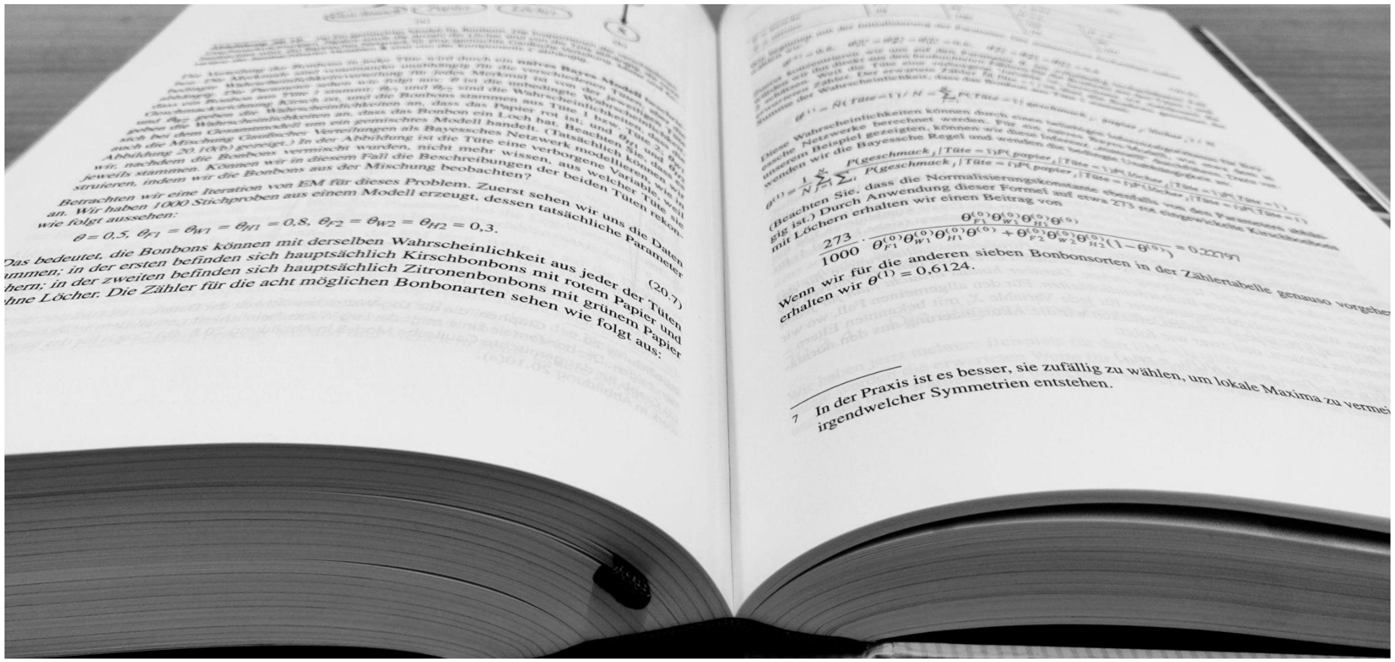
how efficient are linear and binary search in general?



count_words()

word count

how many words are in the book?



wir, nachdem die Bonbons vermischt wurden, nicht mehr wissen, aus welcher Tüte sie
jeweils stammen. Können wir in diesem Fall die Beschreibungen der beiden Tüten rekon-
struieren, indem wir die Bonbons aus der Mischung beobachten?

Betrachten wir eine Iteration von EM für dieses Problem. Zuerst sehen wir uns die Daten
an. Wir haben 1000 Stichproben aus einem Modell erzeugt, dessen tatsächliche Parameter
wie folgt aussehen:

$$\theta = 0{,}5,\ \theta_{F1} = \theta_{W1} = \theta_{H1} = 0{,}8,\ \theta_{F2} = \theta_{W2} = \theta_{H2} = 0{,}3. \quad (20.7)$$

Das bedeutet, die Bonbons können mit derselben Wahrscheinlichkeit aus jeder der Tüten
ammen; in der ersten befinden sich hauptsächlich Kirschbonbons mit rotem Papier und
hern; in der zweiten befinden sich hauptsächlich Zitronenbonbons mit grünem Papier
hne Löcher. Die Zähler für die acht möglichen Bonbonarten sehen wie folgt aus:

Diese Wahrscheinlichkeiten können durch einen
esche Netzwerke berechnet werden. Für ein
unserem Beispiel gezeigten, können wir diese
wenden wir die Bayessche Regel und wenden die

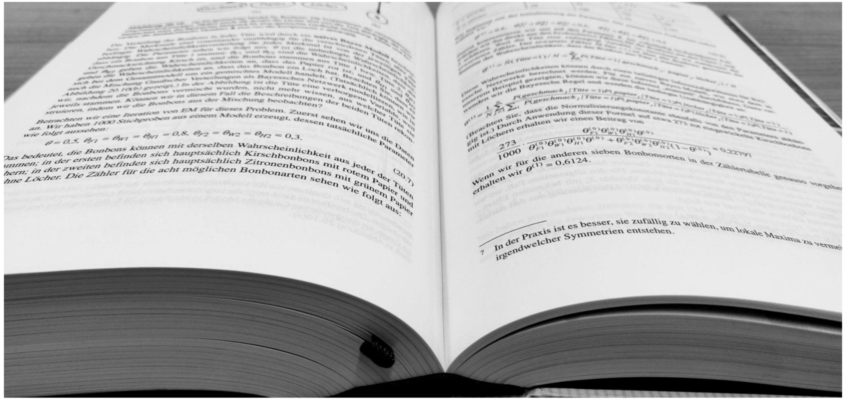
(Beachten Sie, dass die Normalisierungskonstante ebenfalls von den Parametern abhän-
gig ist.) Durch Anwendung dieser Formel auf etwa 273 rot eingewickelte Kirschbonbons
mit Löchern erhalten wir einen Beitrag von

$$\frac{273}{1000} \cdot \frac{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)}}{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)} + \theta_{F2}^{(0)}\theta_{W2}^{(0)}\theta_{H2}^{(0)}(1-\theta^{(0)})} = 0{,}22797$$

Wenn wir für die anderen sieben Bonbonsorten in der Zählertabelle genauso vorgehe
erhalten wir $\theta^{(1)} = 0{,}6124$.

[7] In der Praxis ist es besser, sie zufällig zu wählen, um lokale Maxima zu vermei
irgendwelcher Symmetrien entstehen.

strategies, anyone?



wir, nachdem die Bonbons vermischt wurden, nicht mehr wissen, aus welcher Tüte sie jeweils stammen. Können wir in diesem Fall die Beschreibungen der beiden Tüten rekonstruieren, indem wir die Bonbons aus der Mischung beobachten?

Betrachten wir eine Iteration von EM für dieses Problem. Zuerst sehen wir uns die Daten an. Wir haben 1000 Stichproben aus einem Modell erzeugt, dessen tatsächliche Parameter wie folgt aussehen:

$$\theta = 0{,}5,\ \theta_{F1} = \theta_{W1} = \theta_{H1} = 0{,}8,\ \theta_{F2} = \theta_{W2} = \theta_{H2} = 0{,}3. \qquad (20.7)$$

Das bedeutet, die Bonbons können mit derselben Wahrscheinlichkeit aus jeder der Tüten ...mmen; in der ersten befinden sich hauptsächlich Kirschbonbons mit rotem Papier und ...hern; in der zweiten befinden sich hauptsächlich Zitronenbonbons mit grünem Papier ...hne Löcher. Die Zähler für die acht möglichen Bonbonarten sehen wie folgt aus:

Diese Wahrscheinlichkeiten können durch einen beliebigen Inferenzalgorithmus für Bayessche Netzwerke berechnet werden. Für ein naives Bayes-Modell, wie etwa dem in unserem Beispiel gezeigten, können wir diese Inferenz ... wenden wir die Bayessche Regel und wenden die bedingte Unabhängigkeit an ...

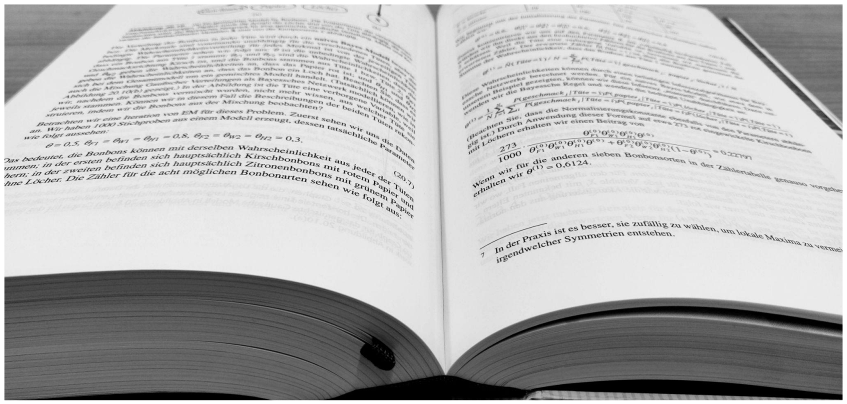
$$\theta^{(1)} = \frac{1}{N}\sum_{j=1}^{N}\frac{P(geschmack_j \mid Tüte = 1)P(papier_j \mid Tüte = 1)P(löcher_j \mid Tüte = 1)P(Tüte = 1)}{\sum_i P(geschmack_j \mid Tüte = i)P(papier_j \mid Tüte = i)P(löcher_j \mid Tüte = i)P(Tüte = i)}$$

(Beachten Sie, dass die Normalisierungskonstante ebenfalls von den Parametern abhängig ist.) Durch Anwendung dieser Formel auf etwa 273 rot eingewickelte Kirschbonbons mit Löchern erhalten wir einen Beitrag von

$$\frac{273}{1000}\cdot\frac{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)}}{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)} + \theta_{F2}^{(0)}\theta_{W2}^{(0)}\theta_{H2}^{(0)}(1-\theta^{(0)})} = 0{,}22797$$

Wenn wir für die anderen sieben Bonbonsorten in der Zählertabelle genauso vorgehen, erhalten wir $\theta^{(1)} = 0{,}6124$.

7 In der Praxis ist es besser, sie zufällig zu wählen, um lokale Maxima zu vermei... irgendwelcher Symmetrien entstehen.



sich bei dem Gesamtmodell um ein gemischtes Modell handelt. (Tatsächlich können wir
auch die Mischung Gaußscher Verteilungen als Bayessches Netzwerk modellieren, wie in
Abbildung 20.10(b) gezeigt.) In der Abbildung ist die Tüte eine verborgene Variable, weil
wir, nachdem die Bonbons vermischt wurden, nicht mehr wissen, aus welcher Tüte sie
jeweils stammen. Können wir in diesem Fall die Beschreibungen der beiden Tüten rekon-
struieren, indem wir die Bonbons aus der Mischung beobachten?

Betrachten wir eine Iteration von EM für dieses Problem. Zuerst sehen wir uns die Daten
an. Wir haben 1000 Stichproben aus einem Modell erzeugt, dessen tatsächliche Parameter
wie folgt aussehen:

$$\theta = 0{,}5,\ \theta_{F1} = \theta_{W1} = \theta_{H1} = 0{,}8,\ \theta_{F2} = \theta_{W2} = \theta_{H2} = 0{,}3. \qquad (20.7)$$

Das bedeutet, die Bonbons können mit derselben Wahrscheinlichkeit aus jeder der Tüten
ammen; in der ersten befinden sich hauptsächlich Kirschbonbons mit rotem Papier und
hern; in der zweiten befinden sich hauptsächlich Zitronenbonbons mit grünem Papier
hne Löcher. Die Zähler für die acht möglichen Bonbonarten sehen wie folgt aus:

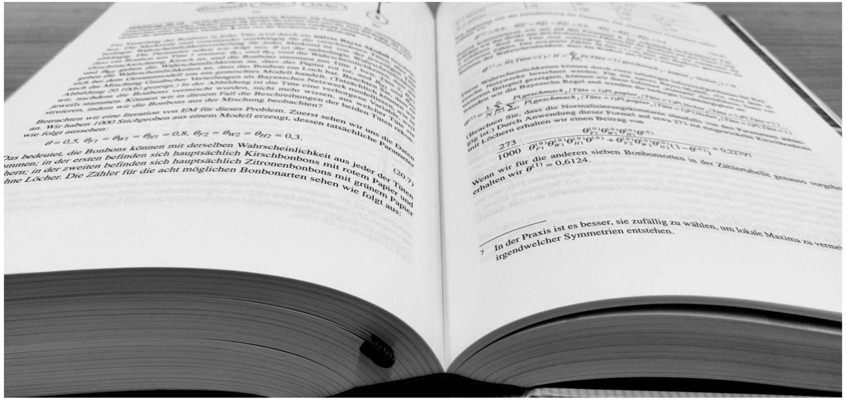
(Beachten Sie, dass die Normalisierungskonstante ebenfalls von den Parametern abhän-
gig ist.) Durch Anwendung dieser Formel auf etwa 273 rot eingewickelte Kirschbonbons
mit Löchern erhalten wir einen Beitrag von

$$\frac{273}{1000} \cdot \frac{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)}}{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)} + \theta_{F2}^{(0)}\theta_{W2}^{(0)}\theta_{H2}^{(0)}(1-\theta^{(0)})} = 0{,}22797$$

Wenn wir für die anderen sieben Bonbonsorten in der Zählertabelle genauso vorgehe
erhalten wir $\theta^{(1)} = 0{,}6124$.

[7] In der Praxis ist es besser, sie zufällig zu wählen, um lokale Maxima zu vermei
irgendwelcher Symmetrien entstehen.

page 1



wir, nachdem die Bonbons vermischt wurden, nicht mehr wissen, aus welcher Tüte sie jeweils stammen. Können wir in diesem Fall die Beschreibungen der beiden Tüten rekonstruieren, indem wir die Bonbons aus der Mischung beobachten?

Betrachten wir eine Iteration von EM für dieses Problem. Zuerst sehen wir uns die Daten an. Wir haben 1000 Stichproben aus einem Modell erzeugt, dessen tatsächliche Parameter wie folgt aussehen:

$$\theta = 0,5,\ \theta_{F1} = \theta_{W1} = \theta_{H1} = 0,8,\ \theta_{F2} = \theta_{W2} = \theta_{H2} = 0,3. \quad (20.7)$$

Das bedeutet, die Bonbons können mit derselben Wahrscheinlichkeit aus jeder der Tüten stammen; in der ersten befinden sich hauptsächlich Kirschbonbons mit rotem Papier und Löchern; in der zweiten befinden sich hauptsächlich Zitronenbonbons mit grünem Papier ohne Löcher. Die Zähler für die acht möglichen Bonbonarten sehen wie folgt aus:

Diese Wahrscheinlichkeiten können durch einen beliebigen Inferenzalgorithmus für Bayessche Netzwerke berechnet werden. Für ein naives Bayes-Modell, wie etwa dem in unserem Beispiel gezeigten, können wir diese Inferenz „manuell" durchführen. Dazu wenden wir die Bayessche Regel und wenden die bedingte Unabhängigkeit an:

$$\theta^{(1)} = \frac{1}{N}\sum_{j=1}^{N} \frac{P(geschmack_j \mid Tüte = 1)P(papier_j \mid Tüte = 1)P(löcher_j \mid Tüte = 1)P(Tüte = 1)}{\sum_i P(geschmack_j \mid Tüte = i)P(papier_j \mid Tüte = i)P(löcher_j \mid Tüte = i)P(Tüte = i)}$$

(Beachten Sie, dass die Normalisierungskonstante ebenfalls von den Parametern abhängig ist.) Durch Anwendung dieser Formel auf etwa 273 rot eingewickelte Kirschbonbons mit Löchern erhalten wir einen Beitrag von

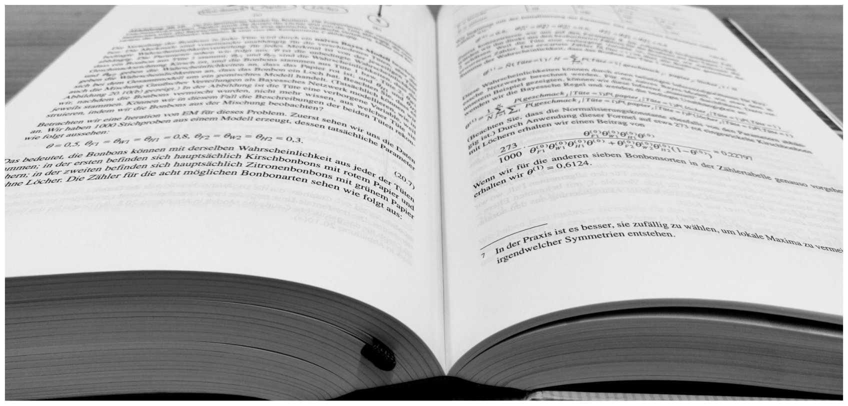
$$\frac{273}{1000} \cdot \frac{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)}}{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)} + \theta_{F2}^{(0)}\theta_{W2}^{(0)}\theta_{H2}^{(0)}(1-\theta^{(0)})} = 0,22797$$

Wenn wir für die anderen sieben Bonbonsorten in der Zählertabelle genauso vorgehen, erhalten wir $\theta^{(1)} = 0,6124$.

[7] In der Praxis ist es besser, sie zufällig zu wählen, um lokale Maxima zu vermei irgendwelcher Symmetrien entstehen.

187

page 1



[illegible] geben die Wahrscheinlichkeiten an, dass das Bonbon ein Loch hat. Beachten Sie, dass [illegible] sich bei dem Gesamtmodell um ein gemischtes Modell handelt. (Tatsächlich können [illegible] auch die Mischung Gaußscher Verteilungen als Bayessches Netzwerk modellieren, wie in Abbildung 20.10(b) gezeigt.) In der Abbildung ist die Tüte eine verborgene Variable, weil wir, nachdem die Bonbons vermischt wurden, nicht mehr wissen, aus welcher Tüte sie jeweils stammen. Können wir in diesem Fall die Beschreibungen der beiden Tüten rekonstruieren, indem wir die Bonbons aus der Mischung beobachten?

Betrachten wir eine Iteration von EM für dieses Problem. Zuerst sehen wir uns die Daten an. Wir haben 1000 Stichproben aus einem Modell erzeugt, dessen tatsächliche Parameter wie folgt aussehen:

$$\theta = 0{,}5,\ \theta_{F1} = \theta_{W1} = \theta_{H1} = 0{,}8,\ \theta_{F2} = \theta_{W2} = \theta_{H2} = 0{,}3. \quad (20.7)$$

Das bedeutet, die Bonbons können mit derselben Wahrscheinlichkeit aus jeder der Tüten ammen; in der ersten befinden sich hauptsächlich Kirschbonbons mit rotem Papier und hern; in der zweiten befinden sich hauptsächlich Zitronenbonbons mit grünem Papier hne Löcher. Die Zähler für die acht möglichen Bonbonarten sehen wie folgt aus:

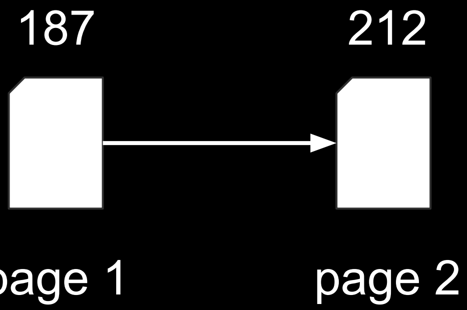
Diese Wahrscheinlichkeiten können durch einen [illegible] esche Netzwerke berechnet werden. Für ein [illegible] unserem Beispiel gezeigten, können wir diese [illegible] wenden wir die Bayessche Regel und wenden die [illegible]

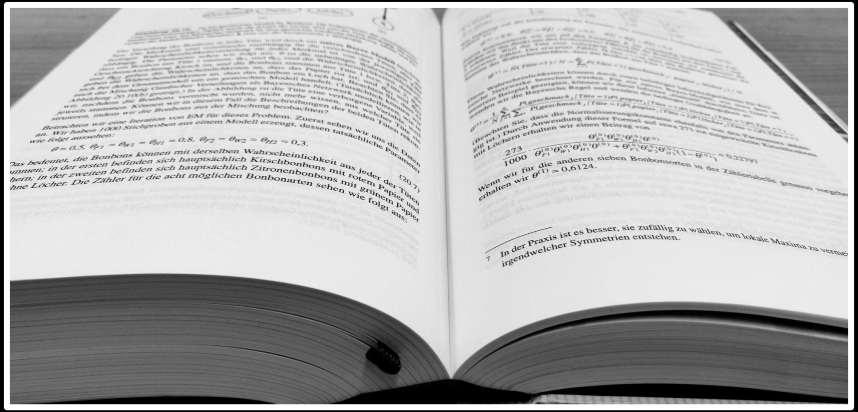
(Beachten Sie, dass die Normalisierungskonstante ebenfalls von den Parametern abhängig ist.) Durch Anwendung dieser Formel auf etwa 273 rot eingewickelte Kirschbonbons mit Löchern erhalten wir einen Beitrag von

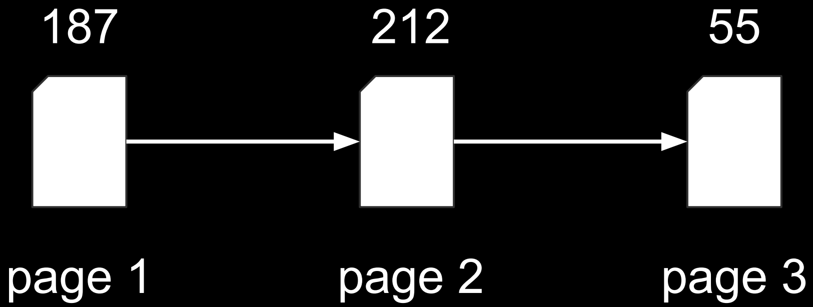
$$\frac{273}{1000} \cdot \frac{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)}}{\theta_{F1}^{(0)}\theta_{W1}^{(0)}\theta_{H1}^{(0)}\theta^{(0)} + \theta_{F2}^{(0)}\theta_{W2}^{(0)}\theta_{H2}^{(0)}(1-\theta^{(0)})} = 0{,}22797$$

Wenn wir für die anderen sieben Bonbonsorten in der Zählertabelle genauso vorgehe erhalten wir $\theta^{(1)} = 0{,}6124$.

[7] In der Praxis ist es besser, sie zufällig zu wählen, um lokale Maxima zu vermei irgendwelcher Symmetrien entstehen.







187

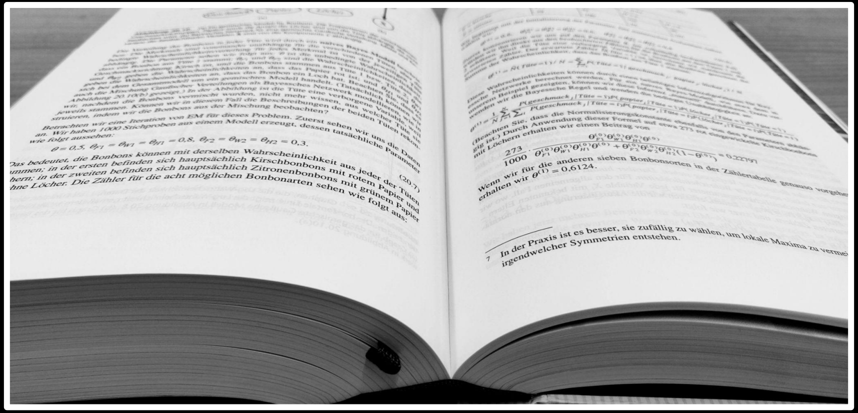
212

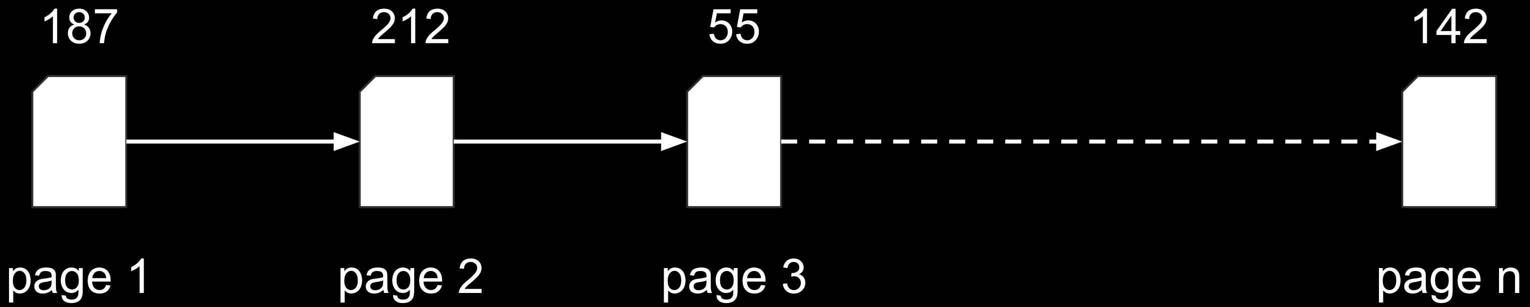
55

page 1

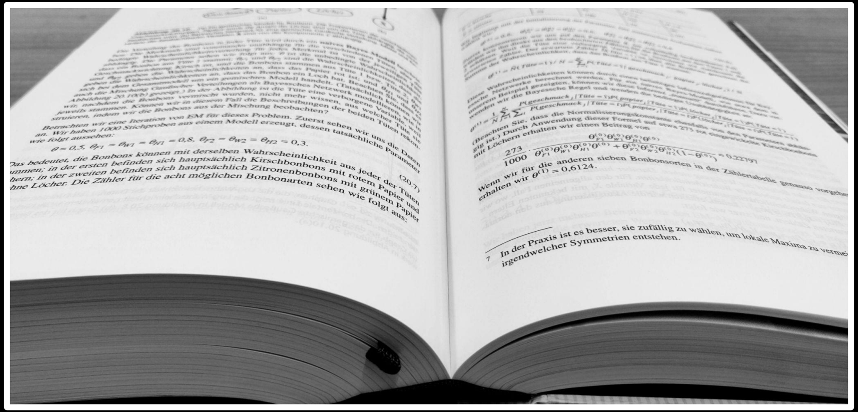
page 2

page 3





187
212
55
142
page 1
page 2
page 3
page n



😅

n = 1327 pages

Ø 2:23 minutes per page

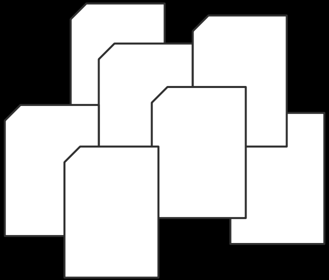
~ 52.34 hours

divide and conquer

+

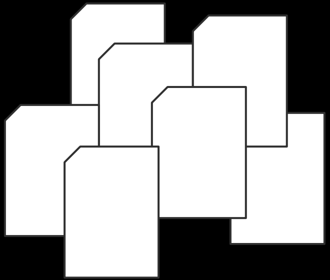
?

pages 1 - 700



student 1

pages 701 - 1327



student 2

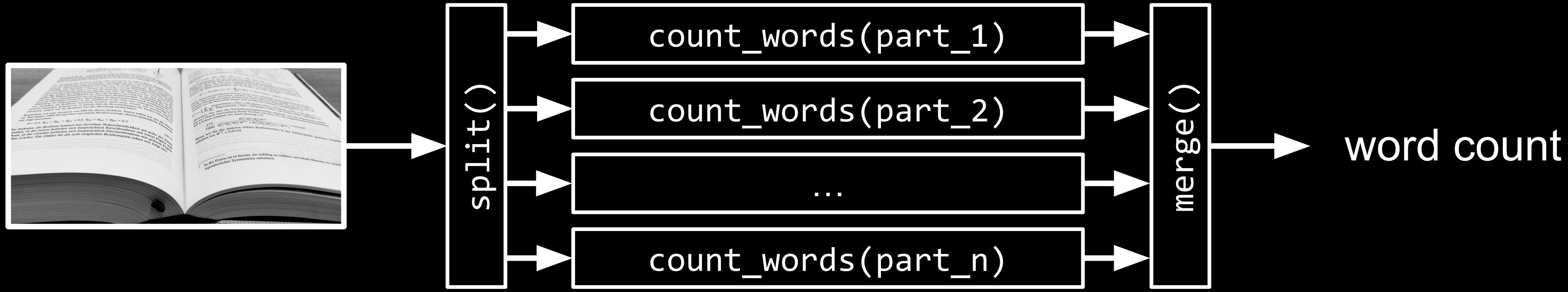
pages 1 - 350	pages 351 - 700	pages 701 - 1050	pages 1051- 1327
student 1	student 2	student 3	student 4

%%problem_solving_example_word_count_distribute_further%%

divide and conquer

+

distribution and parallelization



split()

count_words(part_1)

count_words(part_2)

...

count_words(part_n)

merge()

word count