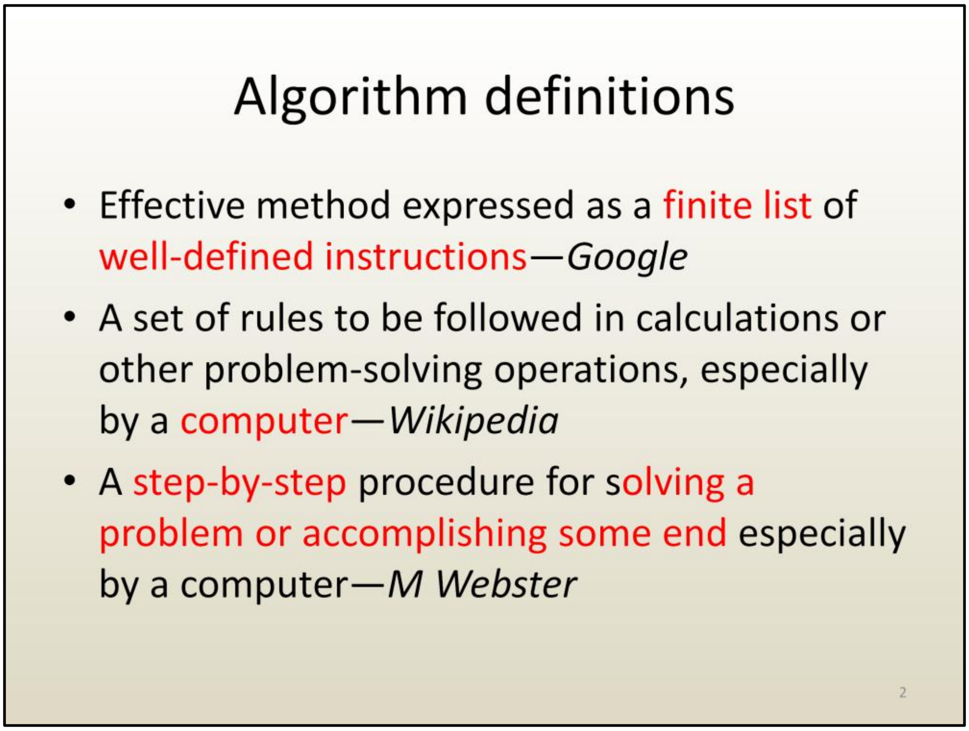
Algorithms

How data is processed

In the first lecture, we called computers malleable tools for processing data. We looked at data, we looked at how it is structured to be stored on the computer, let us now look how to describe the actual data processing.



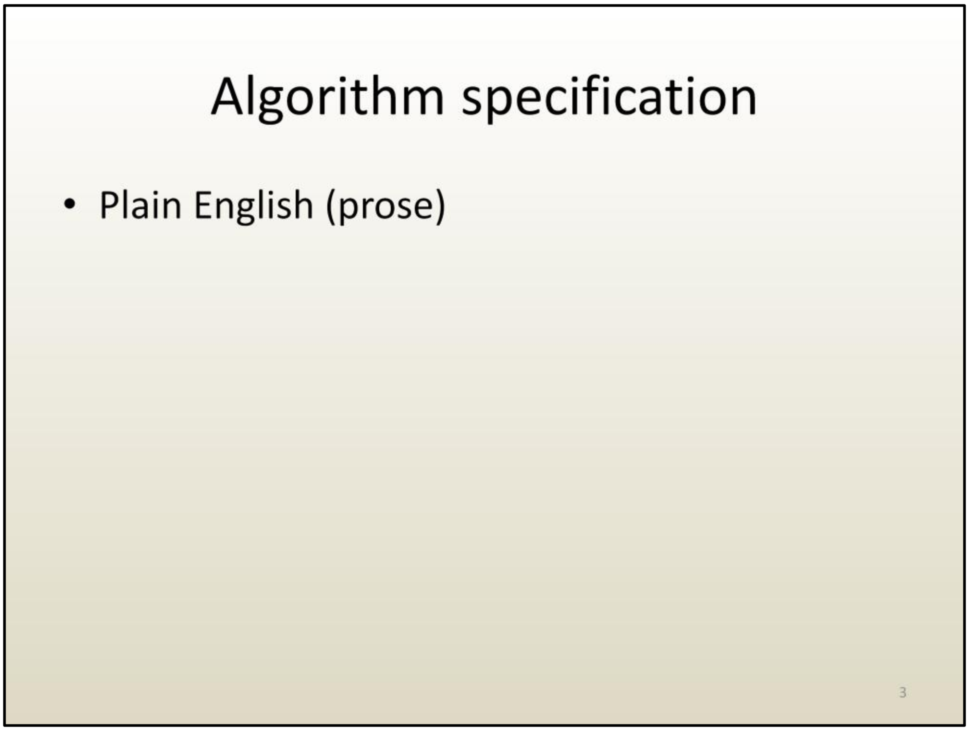
Here are various definitions of the word algorithm. As you can see there is no strict agreement for a single definition. Highlighted in red are important features of algorithms.

The number of instructions has to be finite, meaning that an algorithm cannot involve an infinite number of steps.

It is also the case that the instructions have to be well defined. There can be no ambiguity. The algorithm has to prescribe precisely how data is to be processed.

Algorithms are a set of step instructions for processing data. The instructions are sorted. The first instruction is executed first, the second one after that and so on.

In computer science and certainly in our case, algorithms prescribe how computers are to process data. But remember, algorithms can only be executed by computers once they are encoded into programs. We will worry about programming later, for now we will worry about designing algorithms. For this, we need to find a way of specifying algorithms.



One option would be to use the English language. It is a powerful language, it certainly allows us to express all instructions we need, and it is familiar to us.

Apple pie algorithm

Ingredients:
Pastry for 2 crusts (recipe below)
8 cups sliced, peeled assorted baking apples - about 3 lbs. (Granny Smith, Cortland, Jonathan)
2 Tablespoons lemon juice
3/4 cup white sugar
1/4 cup brown sugar
1/4 cup all-purpose flour
1 teaspoon ground cinnamon
1/4 teaspoon ground nutmeg
2 Tablespoons butter
1 egg yolk
1 Tablespoon milk

Directions:
1. In a large bowl, toss the sliced apples with lemon juice.
2. Combine sugars, flour, cinnamon and nutmeg; add to apples and toss well to coat.
3. Fill pastry lined 9 inch pie pan with apple mixture. Dot with butter.
4. Place second crust on top of pie filling, cut slits in top of crust to vent. Seal the edges of the crust with a fork or by hand.
5. In a small bowl, beat the egg yolk and milk. Brush mixture over top crust.
6. Bake at 425 degrees F for 15 minutes.
7. Reduce heat to 350 degrees F and bake 40-45 minutes more or until crust is golden and filling is bubbly.

www.momswhothink.com

4

And in fact, the English language is used to specify algorithms, i.e. step by step instructions. Here is an algorithm for baking an apple pie.

Apple pie algorithm

Ingredients:
Pastry for 2 crusts (recipe below)
8 cups sliced, peeled assorted baking apples - about 3 lbs. (Granny Smith, Cortland, Jonathan)
2 Tablespoons lemon juice
3/4 cup white sugar
1/4 cup brown sugar
1/4 cup all-purpose flour
1 teaspoon ground cinnamon
1/4 teaspoon ground nutmeg
2 Tablespoons butter
1 egg yolk
1 Tablespoon milk

Input

Directions:
1. In a large bowl, toss the sliced apples with lemon juice.
2. Combine sugars, flour, cinnamon and nutmeg; add to apples and toss well to coat.
3. Fill pastry lined 9 inch pie pan with apple mixture. Dot with butter.
4. Place second crust on top of pie filling, cut slits in top of crust to vent. Seal the edges of the crust with a fork or by hand.
5. In a small bowl, beat the egg yolk and milk. Brush mixture over top crust.
6. Bake at 425 degrees F for 15 minutes.
7. Reduce heat to 350 degrees F and bake 40-45 minutes more or until crust is golden and filling is bubbly.

www.momswhothink.com

5

The algorithm has an input, which in this case (i.e. for a recipe) are called ingredients.

Apple pie algorithm

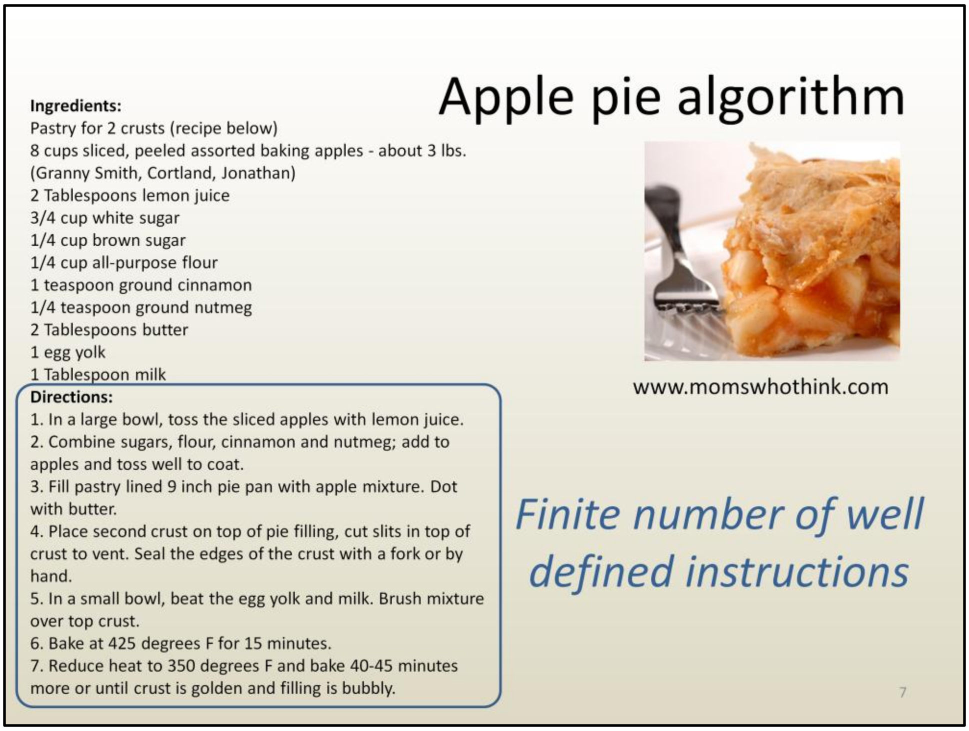
Ingredients:
Pastry for 2 crusts (recipe below)
8 cups sliced, peeled assorted baking apples - about 3 lbs. (Granny Smith, Cortland, Jonathan)
2 Tablespoons lemon juice
3/4 cup white sugar
1/4 cup brown sugar
1/4 cup all-purpose flour
1 teaspoon ground cinnamon
1/4 teaspoon ground nutmeg
2 Tablespoons butter
1 egg yolk
1 Tablespoon milk

Directions:
1. In a large bowl, toss the sliced apples with lemon juice.
2. Combine sugars, flour, cinnamon and nutmeg; add to apples and toss well to coat.
3. Fill pastry lined 9 inch pie pan with apple mixture. Dot with butter.
4. Place second crust on top of pie filling, cut slits in top of crust to vent. Seal the edges of the crust with a fork or by hand.
5. In a small bowl, beat the egg yolk and milk. Brush mixture over top crust.
6. Bake at 425 degrees F for 15 minutes.
7. Reduce heat to 350 degrees F and bake 40-45 minutes more or until crust is golden and filling is bubbly.

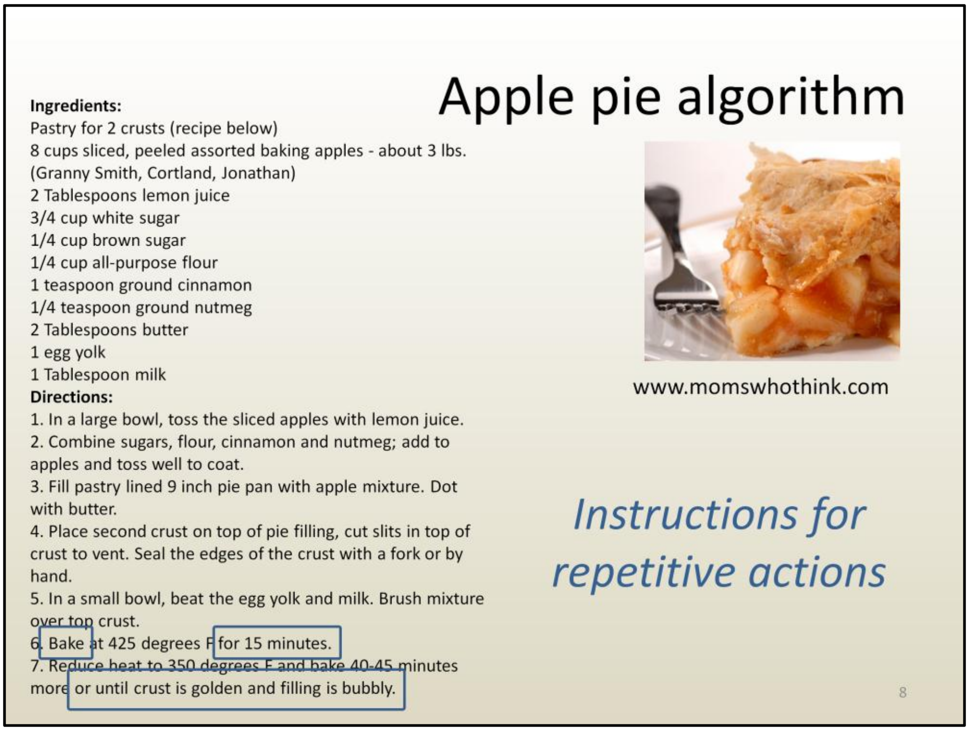
Output

www.momswhothink.com

The algorithm also has an output. Here the output is represented with a photograph, easy to guess why.



And here are the step by step instructions. You see that there is a finite number of instructions, they are well defined, and they transform the ingredients in an apple pie.



Note that some instructions require the repetition of an action. There is a loop in the set of instructions, to repeat some instructions.

Bake at 425 degrees for 15 minutes, means repeat the baking from minute 0 to minute 15. We will see that for loops are a very popular way of specifying instruction repetition in algorithms. It is much easier to have a bake for 15 minutes loop as opposed to saying bake for 1 minute, bake for another minute, bake for another minute, …, bake for another minute.

Another way of asking for an action to be repeated is to say that the action be repeated until a condition is satisfied. Bake until crust is golden and filling is bubbly. This “repeat until” construct is another very popular way of specifying instruction repetition or loops in algorithms.

Repeat until is used when you do not know exactly how many times the repetition should occur, i.e. the body of the loop should be executed. For example here, due to differences in the ingredients, in oven performance, it’s not clear how many minutes the pie needs to be in the oven. Just bake until the conditions are satisfied.

The for loop is used when you can specify the number of times the loop should be executed, i.e. the number of iterations.

Apple pie algorithm

Ingredients:
Pastry for 2 crusts (recipe below)
8 cups sliced, peeled assorted baking apples - about 3 lbs. (Granny Smith, Cortland, Jonathan)
2 Tablespoons lemon juice
3/4 cup white sugar
1/4 cup brown sugar
1/4 cup all-purpose flour
1 teaspoon ground cinnamon
1/4 teaspoon ground nutmeg
2 Tablespoons butter
1 egg yolk
1 Tablespoon milk

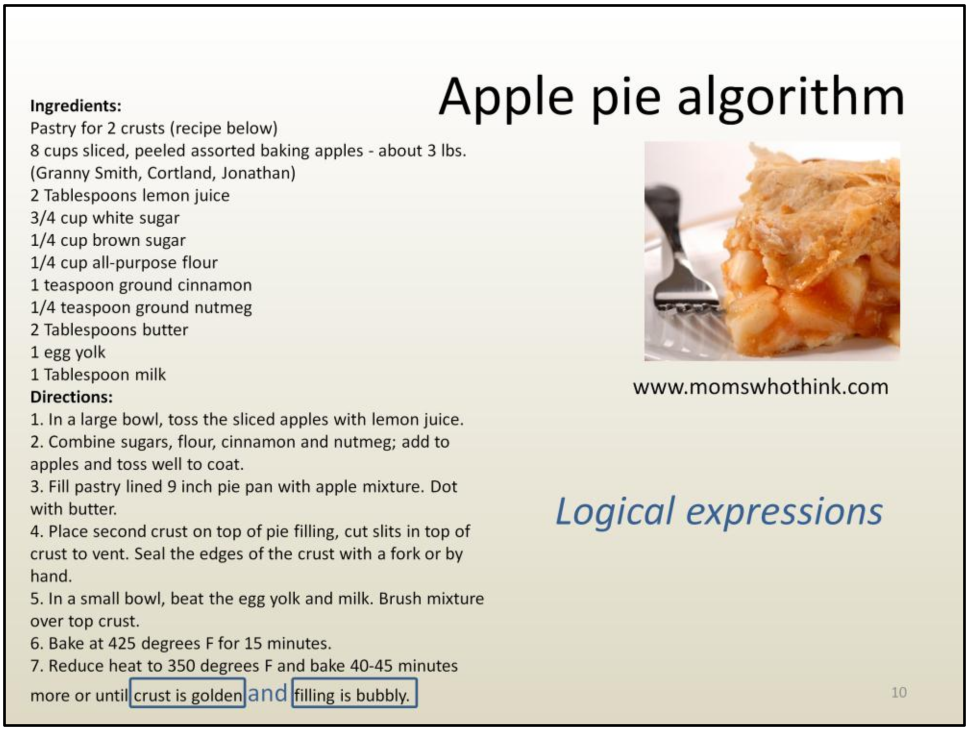
Directions:
1. In a large bowl, toss the sliced apples with lemon juice.
2. Combine sugars, flour, cinnamon and nutmeg; add to apples and toss well to coat.
3. Fill pastry lined 9 inch pie pan with apple mixture. Dot with butter.
4. Place second crust on top of pie filling, cut slits in top of crust to vent. Seal the edges of the crust with a fork or by hand.
5. In a small bowl, beat the egg yolk and milk. Brush mixture over top crust.
6. Bake at 425 degrees F for 15 minutes.
7. Reduce heat to 350 degrees F and bake 40-45 minutes more or until crust is golden and filling is bubbly.

www.momswhothink.com

Instructions for (implicitly) repetitive actions

9

Now because this algorithm is intended for a human, some repetitions are implied and not stated explicitly. Toss well means keep tossing until there are no apple pieces that are not coated. Dot with butter means put a lot of butter dots, use common sense, I am not going to tell you how many, i.e. 17 or 19.



Another important component of the algorithm, and we touched on this somewhat already, are logical expressions. A logical expression evaluates to true or false. Here the termination condition is the logical expression that the crust be golden and the filling be bubbly. Both have to be satisfied, i.e. true, for the compound conditional be true.

True and true is true.
True and false is false.
False and true is false.
False and false is false.

Logical and is true only when both operands are true.

For example the sentence “our university is called Purdue and it is located in Australia” is false, even though the first operand “our university is called Purdue” is true.

Logical or is true when any of the operands are true.

True or true is true.
True or false is true.

False or true is true.
False or false is false.

“Our university is called Purdue or our university is called Stanford” is true because the first operand is true, and regardless of the fact that the second operand is false.

Apple pie algorithm

Ingredients:
Pastry for 2 crusts (recipe below)
8 cups sliced, peeled assorted baking apples - about 3 lbs. (Granny Smith, Cortland, Jonathan)
2 Tablespoons lemon juice
3/4 cup white sugar
1/4 cup brown sugar
1/4 cup all-purpose flour
1 teaspoon ground cinnamon
1/4 teaspoon ground nutmeg
2 Tablespoons butter
1 egg yolk
1 Tablespoon milk

Directions:
1. In a large bowl, toss the sliced apples with lemon juice.
2. Combine sugars, flour, cinnamon and nutmeg; add to apples and toss well to coat.
3. Fill pastry lined 9 inch pie pan with apple mixture. Dot with butter.
4. Place second crust on top of pie filling, cut slits in top of crust to vent. Seal the edges of the crust with a fork or by hand.
5. In a small bowl, beat the egg yolk and milk. Brush mixture over top crust.
6. Bake at 425 degrees F for 15 minutes.
7. Reduce heat to 350 degrees F and bake 40-45 minutes more or until crust is golden and filling is bubbly.

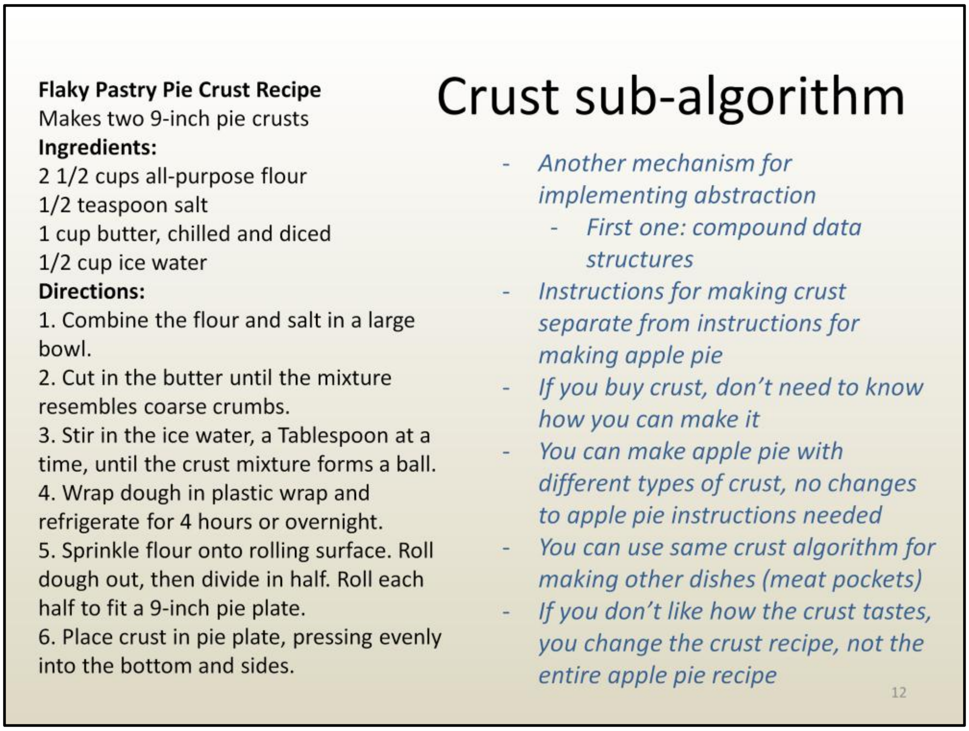
www.momswhothink.com

Use of sub-algorithms

Another important tool in specifying algorithms is to use sub-algorithms.

A sub-algorithm is a stand alone algorithm, which is used by the main algorithm.

For example here the recipe (i.e. sub-algorithm) for making the crust is specified separately.



The sub-algorithm looks just like an algorithm. You can't tell it's a sub-algorithm by looking at it. What makes the crust algorithm a sub-algorithm is the fact that it is called by the pie algorithm.

This is again an example of the use of abstraction. Remember abstraction was also used to specify complex data hierarchically: a car has a chassis, a power train, and a body, a power train has an engine, a gearbox etc. Here specifying the crust recipe separately has the following advantages:

- It makes the overall recipe more readable.
- It decreases the probability of making mistakes: worry about crusts first, then use crusts to make pie and don't worry about how crusts are made anymore.
- It allows to change the crust recipe w/o changing the pie recipe.
- It allows to change the pie recipe w/o changing the crust recipe. Say there are 5 pie recipes and 4 crust recipes. You can combine them in 20 ways if the crust recipe is a sub-algorithm, w/o needing to spell out all combinations.

Apple pie algorithm

Ingredients:
Pastry for 2 crusts (recipe below)
8 cups sliced, peeled assorted baking apples - about 3 lbs. (Granny Smith, Cortland, Jonathan)
2 Tablespoons lemon juice
3/4 cup white sugar
1/4 cup brown sugar
1/4 cup all-purpose flour
1 teaspoon ground cinnamon
1/4 teaspoon ground nutmeg
2 Tablespoons butter
1 egg yolk
1 Tablespoon milk

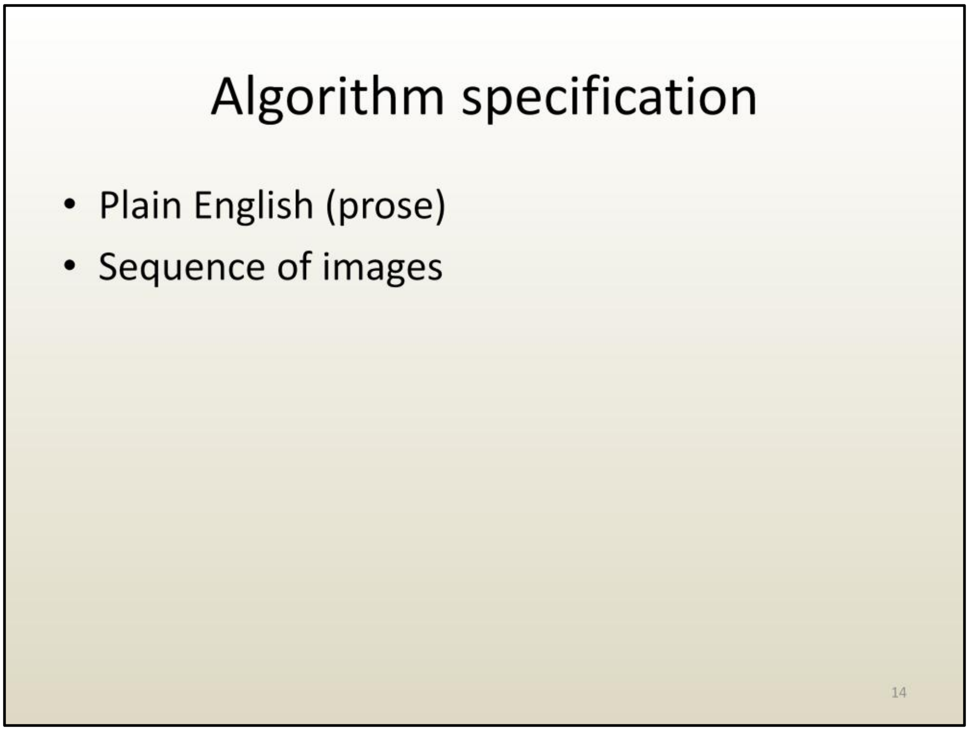
Directions:
1. In a large bowl, toss the sliced apples with lemon juice.
2. Combine sugars, flour, cinnamon and nutmeg; add to apples and toss well to coat.
3. Fill pastry lined 9 inch pie pan with apple mixture. Dot with butter.
4. Place second crust on top of pie filling, cut slits in top of crust to vent. Seal the edges of the crust with a fork or by hand.
5. In a small bowl, beat the egg yolk and milk. Brush mixture over top crust.
6. Bake at 425 degrees F for 15 minutes.
7. Reduce heat to 350 degrees F and bake 40-45 minutes more or until crust is golden and filling is bubbly.

www.momswhothink.com

Generality
- Multiple apple types
- Any brand flour
- Any oven that can bake at 425

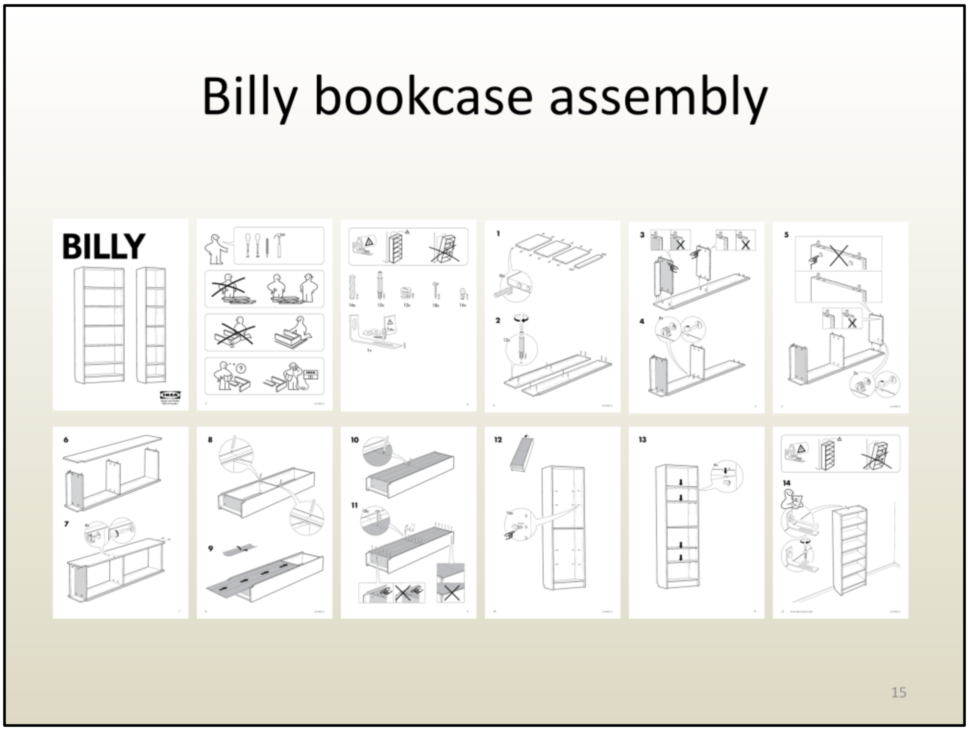
13

Another important trait of algorithms is that they aim for maximum generality. The more the algorithm can accept many types of input, the more useful the algorithm is.



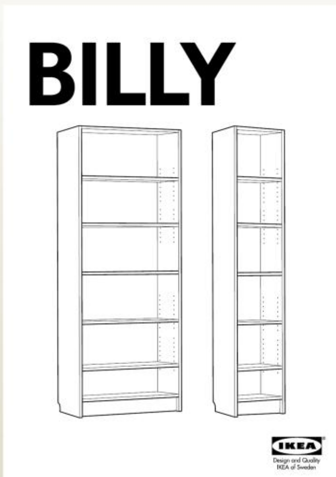
So we have seen algorithms being specified using the English language.

Another popular way of specifying algorithms is using image sequences.



IKEA made a fortune making their clients assemble the furniture they buy. Clear visual instructions are instrumental to pull this off. A long set of written instructions would be harder to follow.

Output



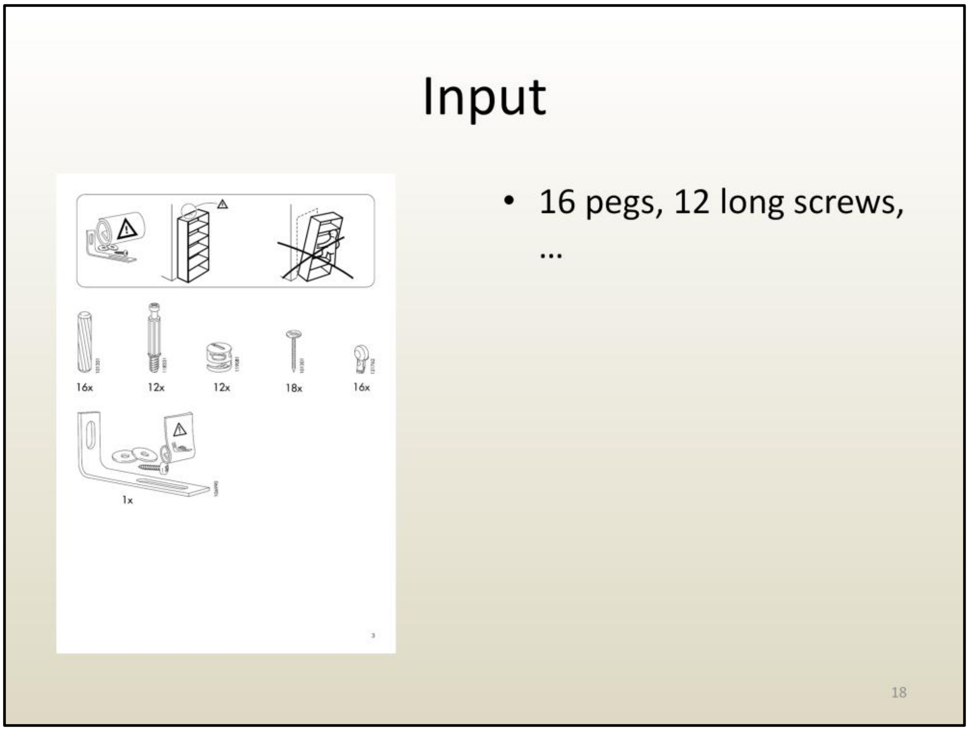
16

Like before, the algorithm has an output.

General rules

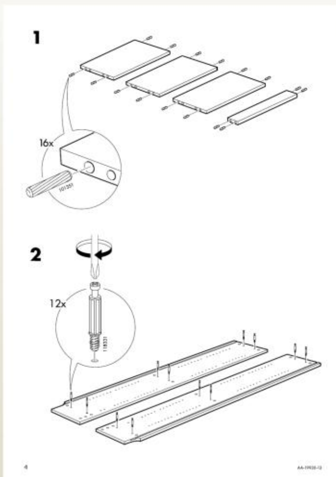
- These are the tools you need
- Don’t work alone, work with someone else
- Don’t work on a hard surface or else you’ll damage the corners, work on a soft surface like a carpet
- If after reading instruction booklet you have questions call IKEA

There are also general rules, not specific to the Billy bookcase. Coming up with images that convey these rules is not straightforward. They spent a lot of time to make these images: clear, minimalistic, universal.



The algorithm also has input.

Step-by-step instructions

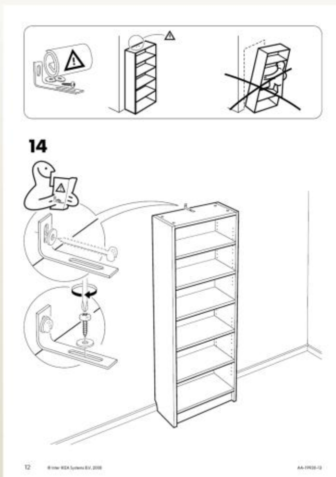


- Instruction 1
- Instruction 2
- …

19

And then there are step by step instructions. Notice that there is also repetition or loops. Instruction 2 is to be repeated 12 times, once for each of the screws. This is indicated with the 12x and with the visual representation of all the screws. Details are provided only once.

Step-by-step instructions

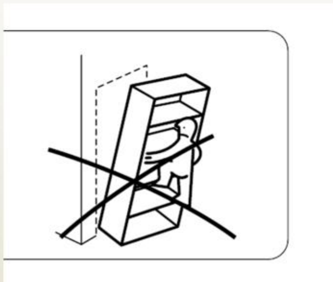


- Instruction 1
- Instruction 2
- …
- Instruction 14
  - Last one, done!

20

And after the final instruction, the bookcase is done. Note that the instructions booklet not only teaches how to built the bookcase, but also how to assemble and use furniture safely.

iClicker questions



What does the image most likely mean?

A. Do not allow children to climb on the bookcase as it might tip over.
B. Children can climb on the bookcase safely as it'll never tip over.
C. Place the bookcase in a spot where it does not case a shadow (shown by dotted line).
D. Never place the bookcase close to a corner.
E. Hard to say, it's probably Swedish

21

Algorithm specification

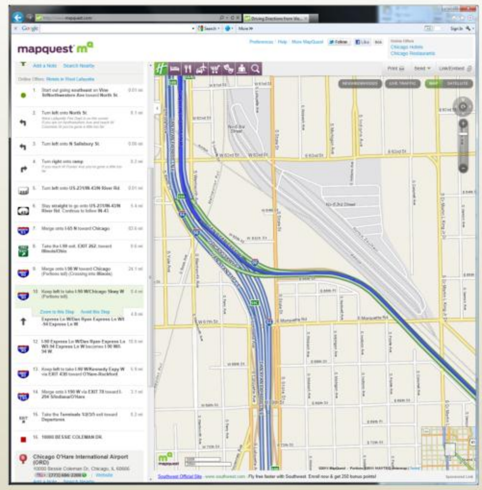
- Plain English (prose)
- Sequence of images
- Video sequence
  - E.g. cooking videos

22

Algorithms can also be specified using videos. As you know there are a myriad of instructional videos and TV shows on an infinite number of topics.

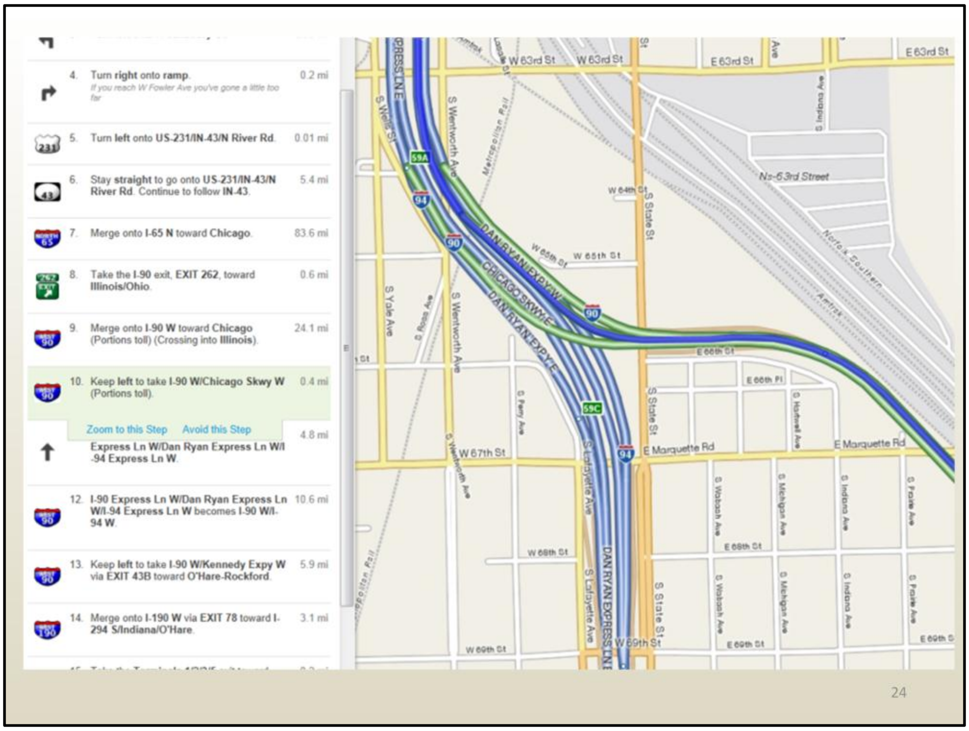
Algorithm specification

- Plain English (prose)
- Sequence of images
- Video sequence
- Hybrid
  - E.g. text and images



23

In some cases a hybrid representation is preferred: text and images (and even spoken instructions).



Each textual instruction has a corresponding image.

Algorithm building blocks

- Input
- Instructions
- Sub-algorithms
  - “Make crust according to instructions below”
  - “Get to I-65 according to instructions below.”
  - “Mount bracket according to instructions on page 3.”
- Repetition
  - For
    - “For each of the four corners, attach one long screw”
    - “Mix for 15 minutes”
  - While
    - “While not golden brown bake”
  - Until
    - “Until golden brown”
    - “Tap in until flush with surface”
    - “Stay on I-65 until Merrillville”
- Conditionals
  - If then do this, else do that
    - “If tolls are acceptable, then take Skyway, else take I-80”
    - “If apples are large, then use only 2, else use 3”
    - “If 1000W microwave, then heat on high for 10 minutes, else if 500W microwave, then heat on high for 17 minutes, else use oven.”
- Output

25

Here are the components of the algorithm.

Input is transformed into output.

The transformation is prescribed by instructions.

Solutions to sub-problems, or sub-tasks are specified in sub-algorithms. The sub-algorithms are invoked as needed from the main algorithm.

Some instructions have to be repeated, in a loop. There are for loops, until loops and while loops. A while loop is similar to an until loop, with the difference that the condition is listed first (“while not golden brown bake” vs. “bake until golden brown”).

The algorithm provides options, to be taken based on the actual input. Not all instructions are executed all the time. For some data, some instructions are chosen over others. This is achieved with conditionals. If this then do that else do something else.

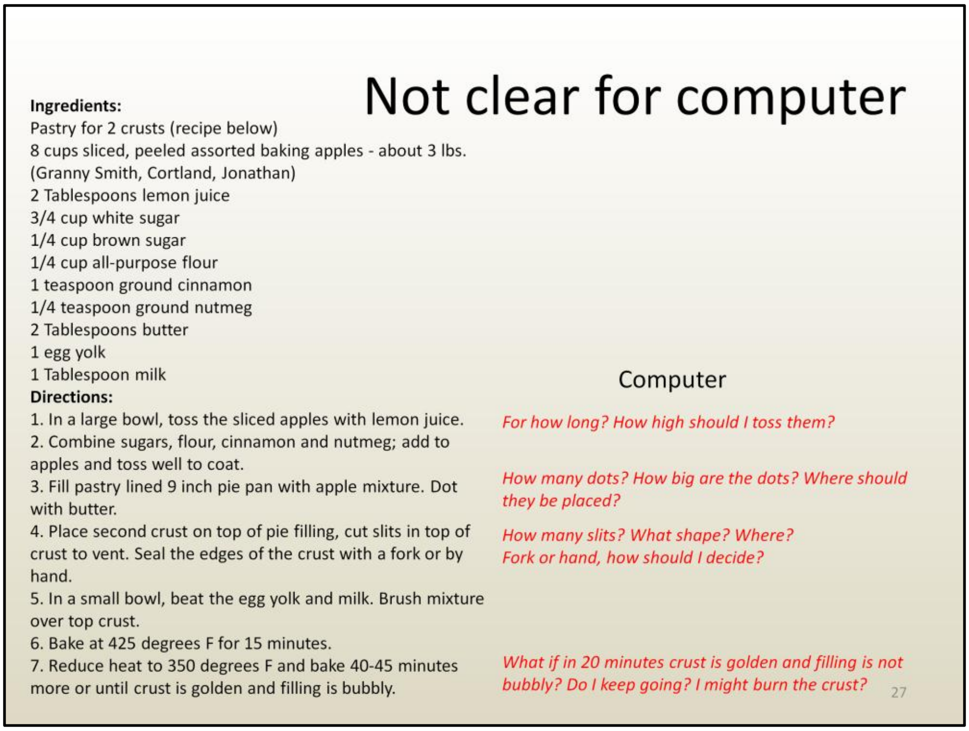
Algorithm specification

- Plain English (prose)
- Sequence of images
- Video sequence
- Hybrid

- These algorithm specifications are intended for humans
- They are not sufficiently clear for computers

26

The algorithm specifications we used so far are intended for humans and not for computers.



The English language specification of the recipe relies on the human's ability to fill in the blanks where needed, to make obvious choices, or to make choices based on personal preference.

Computers don't know how to do anything unless precisely instructed to do so. Computers do not have common sense. Computers do not have a personality. (Yes, artificial intelligence research tries to simulate these human traits, but the point remains that computers need clear instructions.)

iClicker question

7. Reduce heat to 350 degrees F and bake 40-45 minutes more or until crust is golden and filling is bubbly.

What does this recipe fragment mean?

A. If crust is golden and filling is not bubbly, keep baking until filling is bubbly, even if it means burning the crust.

B. If crust is not golden and filling is bubbly, keep baking until crust is golden, even if it means drying out the filling.

C. If crust is golden, stop baking regardless of whether filling is bubbly, to avoid burning the crust.

D. If filling is bubbly, stop baking even if crust is not golden, to avoid drying out the filling.

E. If crust is not golden or filling is not bubbly, keep baking, but use common sense, avoid burning crust or drying out filling. Stop as soon as crust is golden and filling is bubbly.

Not clear for computer

- For the computer some pixels are white some pixels are black…

29

Images are even farther from what computers can understand. You can scan these images into a computer, by digitizing them, but don’t expect the computer to understand what instructions they encode. Yes, there is a computer science research direction that develops algorithms that understand drawings, even sketching. It remains that images is not the way one should specify algorithms for execution on a computer.

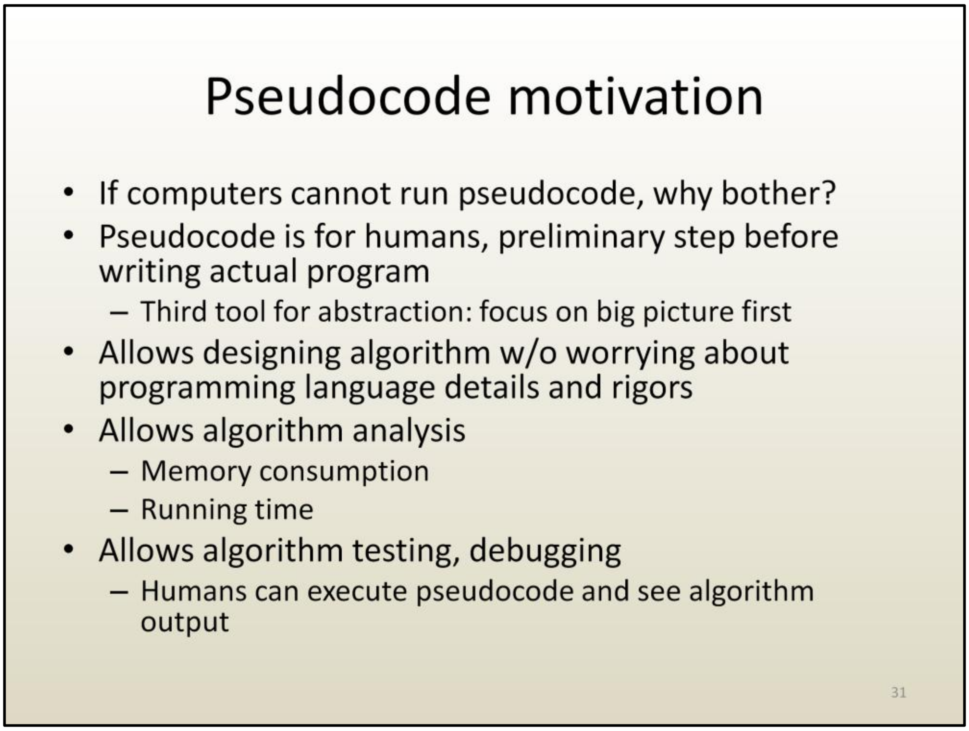
2 steps for specifying algorithms for computers

- (1) Pseudocode
  - "Fake code", "almost code", "quasi-code"
  - English language and mathematical notations for specifying algorithms
  - Formatting and indentation rules
  - Concise, precise
  - Some syntax flexibility
  - Computers cannot execute it
- (2) Program
  - Algorithms specified using programming language
  - Strict syntax rules
  - Computers can execute programs
    - Compilers create machine code
    - Hardware executes machine code

30

We will take two steps for specifying algorithms for execution on computers.

First we will design the algorithm in pseudocode. Then we will implement the pseudocode algorithm using a programming language (Python).



Pseudocode is for humans. It allows us to worry about the algorithm first, and then about making the computer understand and run the algorithm. Using pseudocode is another example of using abstraction.

Do not think that only beginner programmers use pseudocode. I have been programming for over 25 years, and whenever I develop a new algorithm, I always start in pseudocode. If you can write algorithms in pseudocode, not only do you have a great start writing Python programs, you have a great start writing programs in any language.

This coming week we will be writing and analyzing algorithms involving arrays, using pseudocode.