Programming Language Principles

Programming Language Theory

Topics

- What is a Computer?
- Turing Machine
- What is a good Programming Language?

What is a Computer?

- What is a computer in PL’s perspective?
- Programming languages eventually run on computers.
- To design a programing language, or develop a program with it → it is necessary to understand how a computer works.

What is a Computer?

- When you hear this question, there might be various images of computers on your mind.

- In this week’s lectures, we will explore this question more theoretically.

- After the lectures, you will have general, universal and more theoretical view of a computer.

What should we consider?

- When we run a PL on a computer, what should we consider?
- In a PL’s eyes, a computer is providing something like these,
  - Data types
  - Operators
  - Control of Execution
  - Control of Data
  - Memory Management
  - Input and Output

Data Types

- When a computer is doing a computation, the computation is often performed on data.

- There exist various data types, and ***applicable computations are dependent on data types***.

- Data types should be considered to ***verify the correctness of computation*** and also to ***choose a correct computation***.

Operators

- It looks like a computer can handle a complex computation easily.
- However, it combines various basic operations to deal with such complex computations internally.
- How can a computer process multiplication and division?
    - e.g.) using shifter and adder or subtractor.

Control of Execution

- A computer should control its execution of operations.
- e.g.) Executing some operations repeatedly, or executing only a part of operations.
- To obtain a desired outcome, we need to execute operations based on our intention.

Control of Data

- In a computer, CPU eventually processes data which are being computed.
- However, this data do not exist in CPU at first.
- Hence it is necessary to control the flow of data inside a computer.

Memory Management

- When a computer executes a program, usually the program is loaded to memory.
- What if a program itself is larger than available memory?
- Appropriate memory management is necessary to load and remove data from memory.

Input and Output

- When a user is using a computer,
  - the computer gets input from the user,
  - and it provides output to the user.
- Usually I/O takes a huge amount of the processing time, hence a computer should handle it effectively.

So What?

- We know what should be considered, but each programming language will handle these matters differently.
  - e.g.) languages w/ unconditional branch (goto) vs. languages w/o unconditional branch
- Can we define a computer in more general, theoretical ways?

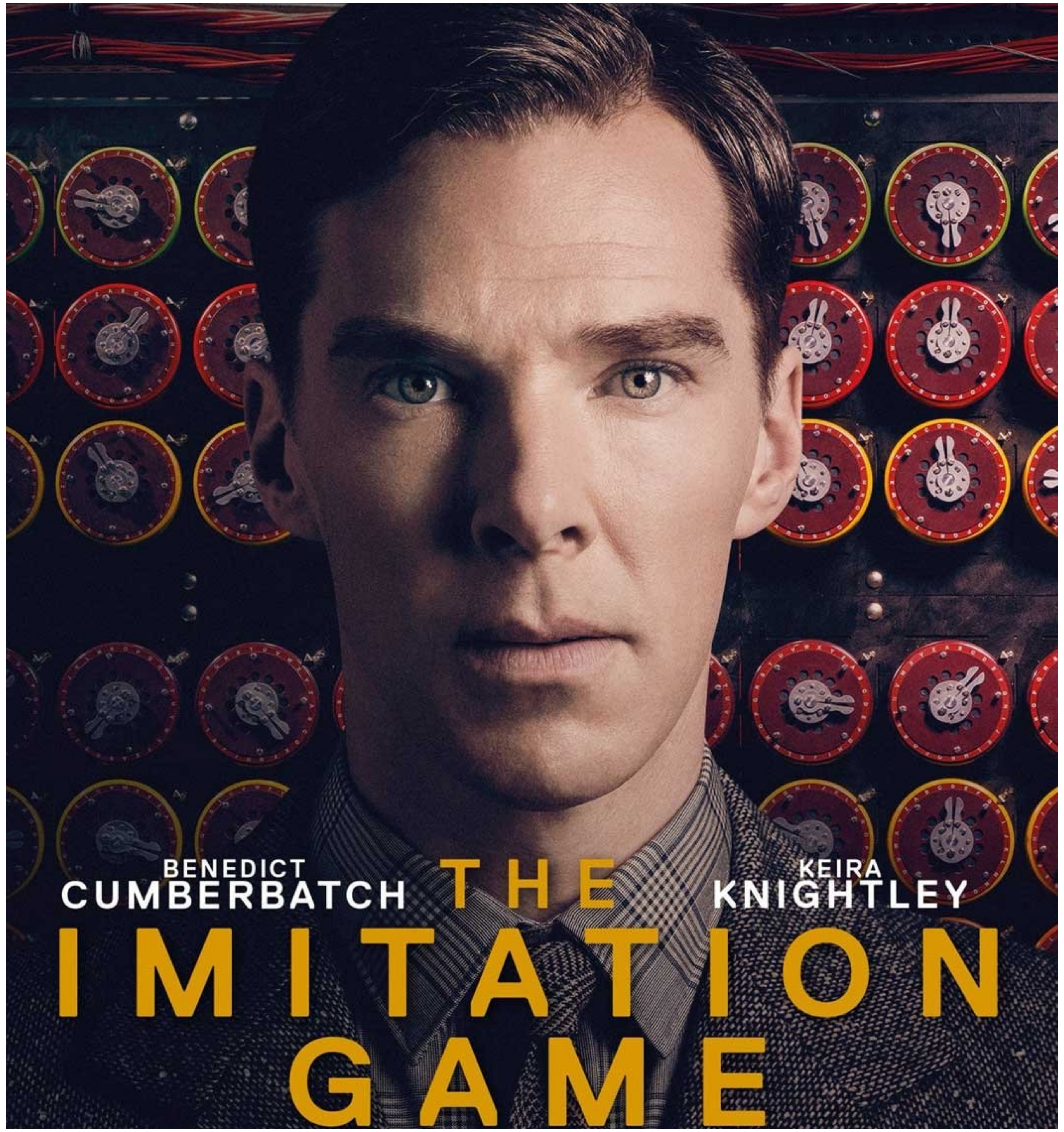
Turing Machine

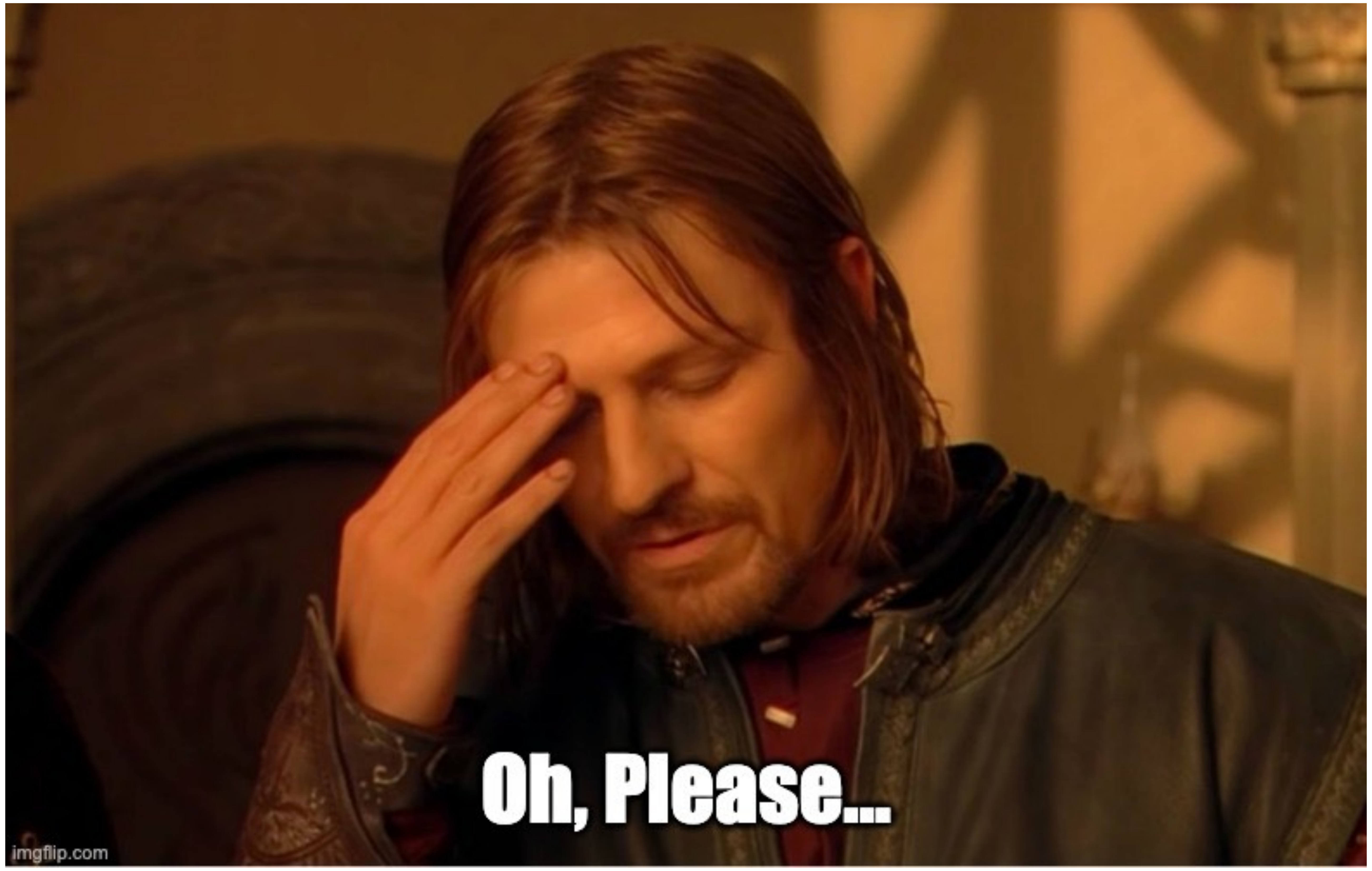
- First introduced in 1936 by Alan Turing.
- Originally it was called “a-machine”, which means automatic machine.
- It was a theoretical, imaginary machine invented to prove properties of computation in general.
- Later it became a foundation of modern computers.

Do you know
Alan Turing?

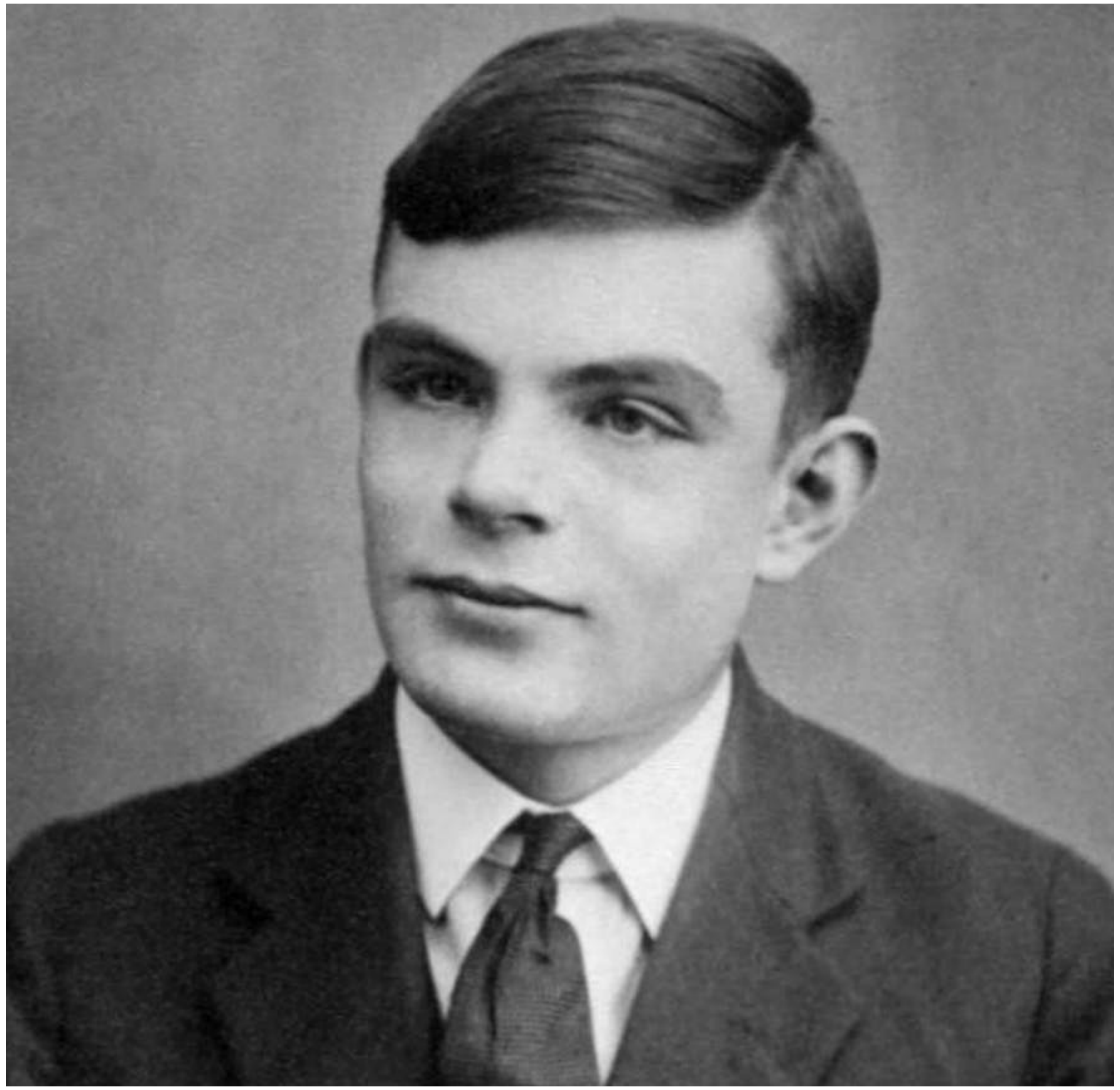
Yes, I Do!

This Guy!




Oh, Please...
imgflip.com

No! This Guy!

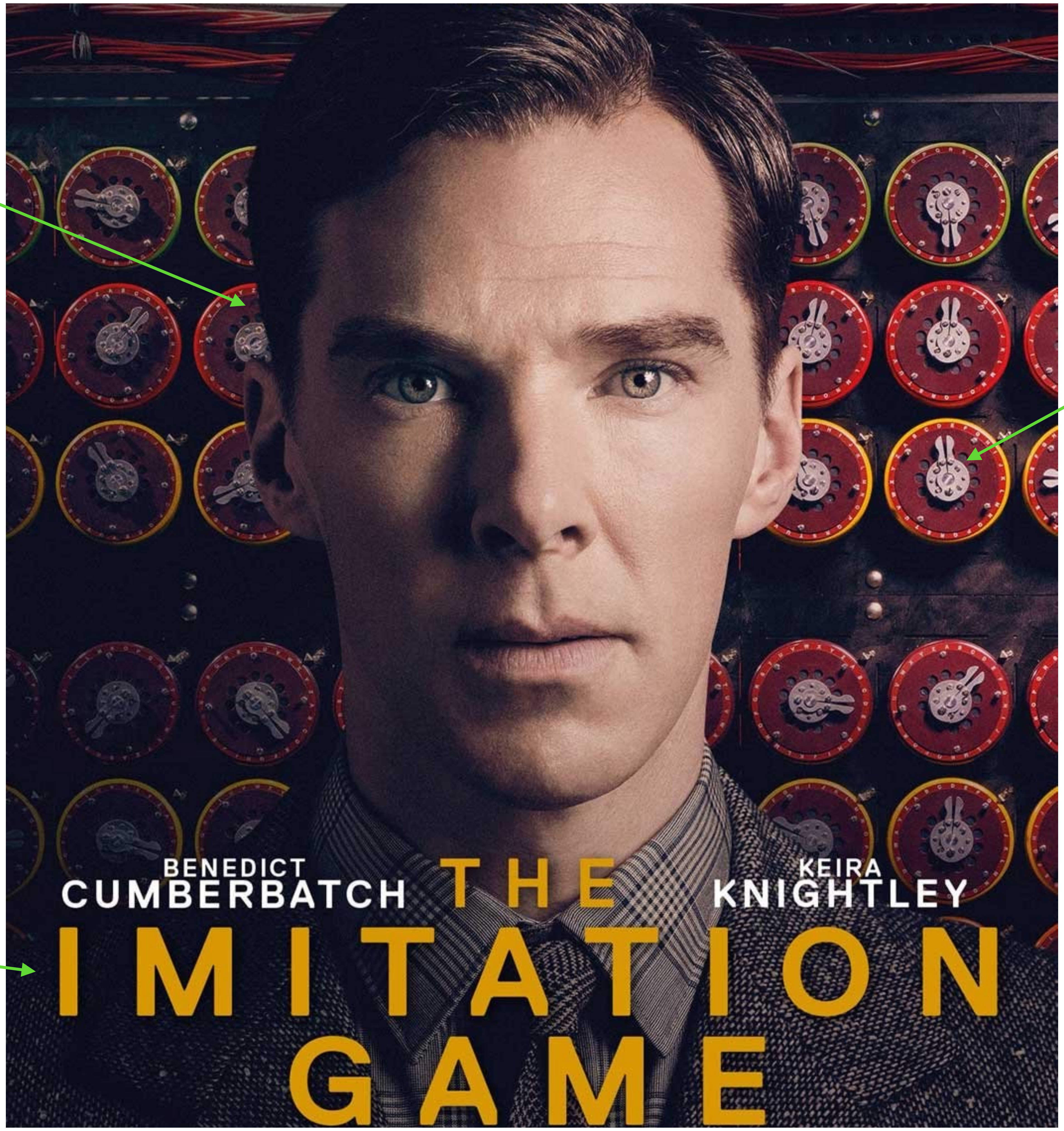


Disturbing Points

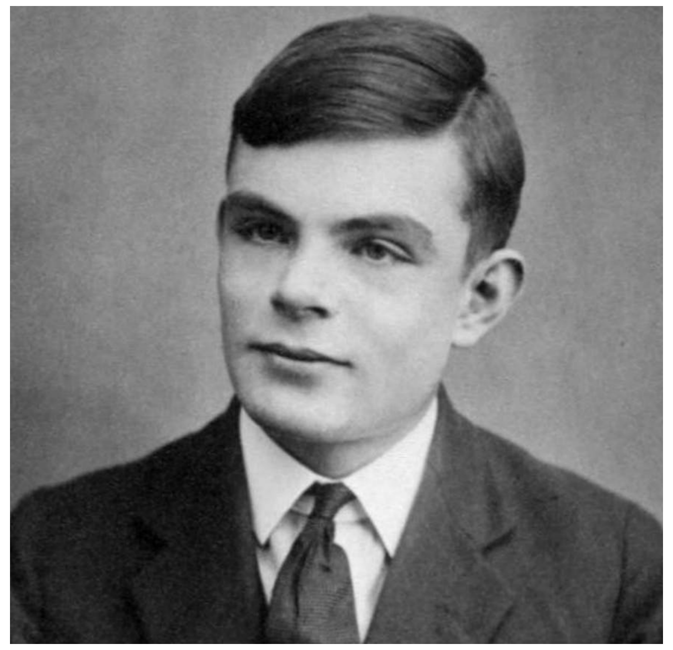
He is not Turing!

This machine is not Enigma!

The movie is not related to the imitation game!



Alan Turing
(1912~1954)

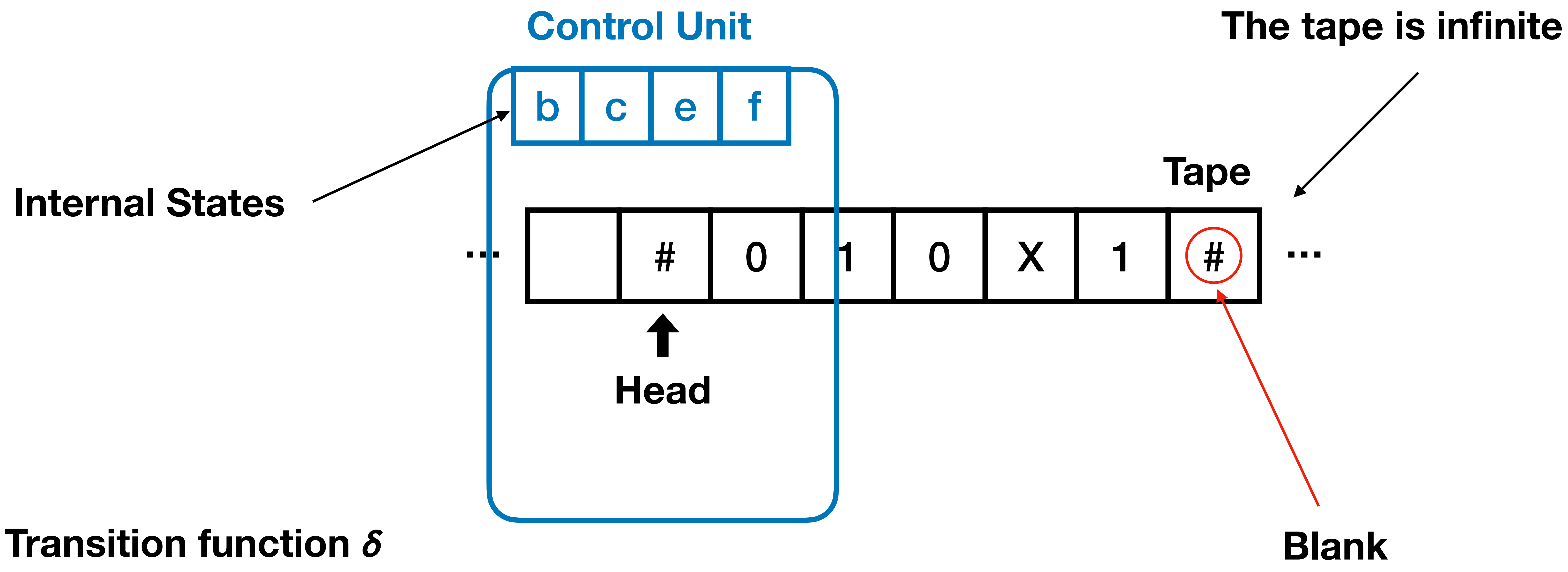


- British computer scientist, logician, cryptanalyst.
- **Imitation Game**: a.k.a Turing Test. First introduced by his paper "*Computing Machinery and Intelligence*" in 1950.
- It is about how to verify whether machines can think (or imitate human) or not.
- **Halting Problem**: Proved the existence of *undecidable* problem.
- He built a foundation of theoretical computer science.

Turing Machine (cont’d)

- Turing machine consists of a control unit and an infinite tape.
- A **tape** is divided into cells, and each **cell** contains one symbol.
- **Head** points to the current cell, and it can read or write a symbol to the cell.
- **Control unit** controls the move of the head, left or right, and performs a certain operation based on the current symbol.

Turing Machine (cont'd)



Transition function δ

current state	symbol	operations	final state
b	#	P0, R	c
c	#	R	e
e	#	P1, R	f
f	#	R	b

Turing Completeness

- So far, it is known that all ***computational problems*** can be solved by a Turing machine.
  - e.g.) Anything can be done by computers can also be done by a Turing machine.
- A system is ***Turing complete***, if it can be used to simulate a Turing machine.
- A Turing complete system has equivalent ability of computation as a Turing machine.
- Theoretical computational ability of programming languages.

To Design a New PL

- What is a purpose of the new programming language?
  - Numerical Computation, Web programming, System Programming, etc.
- Implement common concepts, which are useful to the purpose of the new programming languages.
- Minimize drawbacks brought with such concepts to fulfill the purpose.

Criteria for Language Design

- A good design of a programming language should consider various criteria.
- These criteria are affected by multiple characteristics of a language.

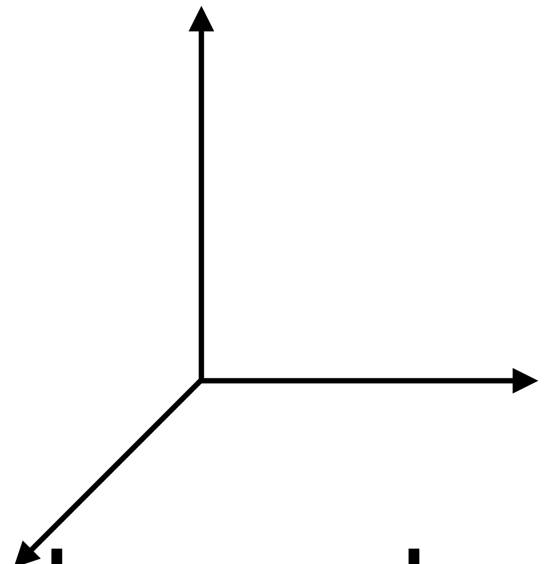
Characteristics	Criteria		
	Readability	Writability	Reliability
Simplicity	•	•	•
Orthogonality	•	•	•
Data types	•	•	•
Syntax design	•	•	•
Support for abstraction		•	•
Expressivity		•	•
Type Checking			•
Exception handling			•
Restricted aliasing			•

Table 1.1 from Sebesta, Concepts of Programming Languages, 11th Ed.

Criteria for Language Design

- ***Readability***: How easy is a language to read the code?
- ***Writability***: How easy is a language to write a program we want?
  - e.g.) assembly languages vs. C/C++ vs. Python.
- ***Reliability***: Does a language always work as expected?
  - e.g.) type checking, exception handling, aliasing.

Orthogonality



- ***Orthogonality*** means that components can be used independently.
- In PL, we can *combine a small number of primitive constructs based on a set of rules* to create a complex program.
- Such orthogonality has certain advantages.
  - We can use a language after learning a small set of constructs and rules.
- However, it may be more difficult to write a correct program.
  - A combination of constructs can be legal, but not we want.

Aliasing

- ***Aliasing*** indicates that the same object can be referred by different names.
  - e.g.) C++ reference variable.
  - `int val = 10; int& ref = val;`
- A programmer must remember all references of a variable to anticipate influences of modifying the variable value.
  - It will affect reliability of programs written by such a language.

Summary

- Computer in PL’s perspective
- Turing Machine and Turing Completeness
- Designing Programming Languages