Rovesti Gabriel

Computability simple (for real)

Summary

[Introduction 2](#_Toc148274801)

[Algorithms, effective procedures, non-computable functions 4](#_Toc148274802)

[URM Computability 10](#_Toc148274803)

[Decidable problems 18](#_Toc148274804)

# Introduction

Course reference page: <http://www.math.unipd.it/~baldan/Computability>

(This lesson is based on the only set of slides of the course, available as “Intro-en.pdf”)

We start by a simple reflection; can we give the enumeration of all numbers and store them efficiently? A suggestion might be, “rather than the phone number itself, you might store a program that generates the number”. So, instead of We can write

It isn’t convenient; there are *numbers*  such that, for all *program P* generating *n*, . These are defined as *random numbers*; we observe there are an infinite number of them. There is no program capable of determining whether a number is random or not, because such a program does not exist.

Exercise (coming from the 8th slide):

1. Prove that there are infinitely many random numbers
2. Prove there is no program able whether a number is random or not

Notes on the previous:

1. We see this as compression, as a function that takes a set of inputs and with a special property we take back the previous file, because we inject it; for this reason, we can’t compress every single file
2. We take a programming language of preference to try to prove this, but we miss the full theory course to prove that entirely; still, we can try

Solutions

My proof for the first question (not proof-corrected by any teacher, so to take with a grain of salt)

1. We define a set of functions , where there is a natural number (phone number) mapped to another phone number. This represents a set of functions.
   * We assume there can be only finitely random numbers, such as . We call this set .
   * We let as a natural number inside of
   * Let be the program generated by for
   * Let the program generated by for , where
   * Let’s extend the previous concept of here
     + For each , there exists a program such that such that generates with . This follows the definition of the compression function .
     + Because the compression should not lose data in computation, we consider a number random in such a way the compression function should be injective for
2. For each , we have:
   1. generating
3. Consider a new number not inside the set
4. If is random, there should be a program generated by for such that:
   1. generates
5. However, since is not in the set , must have a different representation from . This is because if were the same as any of the , it would generate a number from the set , which is a contradiction.
6. By this, we prove there is a new program that generates a new program that generates a number not inside the original set and contradicts the assumption there are only finitely many random numbers

Proof for the second question (again, not proof-corrected by any teacher, so to take with a grain of salt):

* To prove this problem, for the sake of contradiction, we might argument there exists a program that can determine, given an input , if this number is random or not.
* We construct a list of numbers, each corresponding to any programming language of preference, structured low level as binary strings; consider for instance:
  + Program
  + Program
  + Program
* We create a program as follows:
  + For an input , does the opposite of what does (if structured as a Turing machine problem, it would be for example the complement of the TM that takes as input the previous strings)
    - If says that is random, outputs “not random”
    - If say that is not random, outputs “random”
* Now, consider what happens when we apply B to its own description. That is, we ask whether is a random program or not: . This leads to a paradox:
  + If is "random," then by the definition of , should be "not random." But this contradicts the assumption that can correctly determine randomness.
  + If is "not random," then by the definition of , should be "random." Again, this contradicts the assumption that can correctly determine randomness.

What we do know is that not all problems are not solvable by a computer, because of power constraints and limitations of machines, e.g. the halting problem and the program correctness (it’s impossible for even simple specifications). A natural question we naturally ask: “Which problems can we solve by a computer / by an effective procedure?”. Some problems are intrinsically theoretical, so they are completely independent from the underlying computation model.

Other specific questions:

* What is an *effective procedure*?
  + Maybe the simple program can do the job, but we must prove it formally
* What does it mean that *a problem is solved by an effective procedure*?
* Characterize the problems that can and those that cannot be solved
  + Problems that are not always binary
* Relating *unsolvable* problem (degree of unsolvability)

We tend to classify *solvable/unsolvable problems without limitations on the use of resources* (memory and time). For example, the complexity theory, considering the resources and classifying solvable problems in an hierarchy according to their “difficulty”.

*Computability theory* is a branch of computer science and mathematics that explores the theoretical limits of computation, this well before its proper birth. It revolves around the concept of decidability and undecidability, focusing on what can and cannot be computed algorithmically. So, *computer science* may be described as “the ability of building and using tools, according to some (codified) procedure, is a distinctive feature of human beings”. It depends on “how we use the tools and what we find out when we do”, according to *Djikstra.*

We don’t tend to think meaningfully always, but to think *according to patterns*, because there is a general combinatoric procedure to find all truths, reasoning and deriving consequences from a set of premises. Thing is, it doesn’t depend on the language, but we can try to represent things abstractly as a set of customized symbols (creating laws or languages), compute them logically (arithmetically) without contradiction and evaluating problems with procedures, to avoid controversies of decidability and solvability as criteria (*Leibniz, Boole, Lullus and others*) using *logic* as the main foundation.

Others posed the need of an artificial language, formally with syntactic and manipulation rules that can be programmed via *variables* and *statements* general purpose. Using cases like Russell’s paradox, we can use the same tools we already have to contradict ourselves and pushing further, even finding new meanings, possibly having a *consistent* system, where it proves itself as correct solidly (*Hilbert*). Many times, this observation led to creation of special-purpose machines, able to compute a specific class of problems.

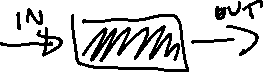
We might try to take problems considering a small set of rules, which may not be always complete or prove the consistency of the theory (*Godel*). There may be a machine which computes a problem given a computable function and the same language, given a specific input and an output (*Turing*). We may express a universal machine to make *any kind of calculation*, storing the result of operations (memory) and solving problem discretely (*Von Neumann*).

Other things:

* On Moodle there are unofficial notes
* There are the exercises with solutions (suggested the ones with no solutions)
* There will be tutoring activities for this course

# Algorithms, effective procedures, non-computable functions

An effective procedure it’s just a sequence of *elementary steps* which are describing a procedure intended to solve a problem (reaching some objective mechanically), transforming some *input* data into some *output*. We can see an algorithm as a black box of sort:



If this is deterministic, we can mathematically describe a function , where each possible input will uniquely determine the corresponding output (we will see later this happens on *partial functions*, so maps between two sets X and Y that may not be defined on the entire set X; an example might be the square root, where not all real numbers have real square roots so *we can compute it but not always solve*).

A function is computable if *there exist*s an algorithm such that the induced function is (so is the function computed if is *effectively computable*). It’s important to note the algorithm that computer must exist.

We informally expect some functions to be computable, given the definition above, such as:

* (eventually an n-th prime will be found)
  + this is a series that converge to and we work with techniques to allow rounding the error, such as truncating the series or rounding the computation

Let’s give an interesting example:

* + More generally, it can be written, for example as
    - (where iff means “if and only if”)

The naïve idea of this last one is:

* compute all the digits of
* check if there are digits of in a row

This, however, is not an algorithm, because we can’t exclude entirely the generation on at some point. Since ’s decimal expansion is non-repeating and doesn't follow a simple pattern, we cannot guarantee that the algorithm won't eventually find the desired sequence of n 5's (given is an irrational number), so we may run it indefinitely and will eventually become infeasible, because we have no way of returning 0.

Is this function computable? In the case of this function, we don't have an effective procedure known to us to determine whether it's computable or not (hence, it’s *not an effective procedure*). The fact that we can't exclude the existence of an effective procedure *doesn't mean* the function is computable, but it also *doesn't definitively prove* that it is computable.

Let’s consider now a slightly different example, for a function :

* + We deduce that somehow, we will reach as constant substituting the values
  + More generally:

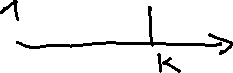
Consider

We then have two possibilities (with plot of the functions reported here, given its quite simple shape):

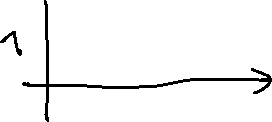










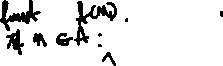


This implies the function is computable, because it behaves regularly (step function, so either 1 or 0, or just a constant function, so they can be computed by simple programs). Even though we won’t know the exact shape of the function, this way we proved it’s computable (the function shape is irrelevant in knowing wich program will compute the function).

Can we use the same argument for ?

Let

and take:



Problem is, is not computable in the slightest, because the set is possibly infinite and there is no such a thing as a finite representation for it (in the notes, it’s also present an example of a function which is , otherwise; since the condition does not depend on the variable, it can have either way or as value, so the function remains computable, but if posed inside the set would be equally incomputable).

This poses the question for the existence of non-computable functions, because it suggests is computable, because the set is possibly infinite and we can’t provide a finite representation.

A good algorithm should satisfy the following characteristics which can be ideally implemented in a theoretical machine we call *computational model*, this way being considered *effective*:

* it has a *finite length*
* there exists a *computing agent* able to execute the algorithm instructions
  + this agent has a *memory* to store the input, results and steps and it is *unbounded*
    - even if the algorithm will be finite, we assume it is unbounded for the sake of analyzing if it’s computable or not (large, but never using the full space)
    - this way, we will be able to define algorithms working on any possible input and there is no limit on the memory that can be used
  + the computation consists in *discrete steps*, not probabilistic or not-deterministic
  + finite limit to number of instructions and the power of their complexity
    - this way representing a finite machine
* the computation can
  + terminate in a finite yet unbounded number of steps output
  + diverge (never terminate) no output

Let’s recall the *math notation* needed to understand the subsequent inference of non-computable functions for evert “effective” computational model.

* set of *natural numbers* (so finite and always with a successor)
* as *Cartesian product* (combine two sets to create an ordered one)
  + We will write, having set,
* *binary relation* or *predicate* as







* , the *partial* function, special relation such that
  + We write
* In words, we essentially say it’s a mathematical relationship that associates elements from a set to elements in a set , but it may not be defined for all elements in (for example, not all pairs)
* When you apply the partial function to an element in its domain, you write to indicate that the function is defined and yields a result. Conversely, if you try to apply the function to an element outside its domain, you write to signify that the function is undefined for that input.

Given a set , we indicate with the cardinality (number of elements), then we define, for sets and :

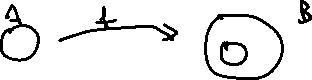


* (unique and complete mapping)

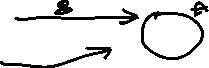


* (no two different inputs map to the same output)



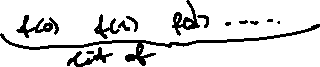


* (covering the entire codomain – all possible outputs – with

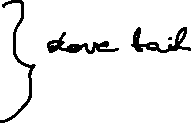


Observe also that if , having injectivity in between.

* (listing all the elements one after the other)



Idea (just to visualize the whole thing, place the elements in a matrix and enumerating them in diagonals):



This so called “dove tail enumeration” means systematically listing all functions from to :

* Begin by listing the element at position in the matrix, which is the function that maps to .
* Then, move along the diagonals of the matrix, listing the elements in order
* A countable union of countable sets is countable:

Let’s come back to the existence of non-computable functions: we focus on unary function over the natural numbers (function that takes a single argument or input variable and produces a single output):

We then fix a model of computation, which then induces a set of algorithms, for example a set of all algorithms inside of it. Given an algorithm we compute a function , which is said to be *computable* in our model if there exists an algorithm that computes it.

Hence, we define *functions computable in*  like:



Clearly we have . Is this inclusion strict? (so, , which means (is there a non-computable function?)



The answer is yes, because the algorithms are too few to compute all the functions, so they must be countable in some way, hence by logical closure computable.

By assumption, an algorithm is a finite sequence of instructions from an instruction set , which we assume finite. We can interpret all of this as a big union of finite algorithms.

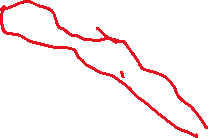
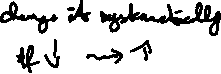
(countable union of countable sets 🡪 countable)

Given and since we have (which means it is surjective be definition), we have:

What we say in words is this: the set of all algorithms in our fixed computational model and , the set of computable functions are as many as the natural numbers.

On the other hand, the set of all functions is not countable. Why? Assume for the sake of contradiction that it is so:

We can list the elements of like we did before (taking, with diagonalization, the main diagonal, then systematically changing diagonal values):



then build a function on that, like this one:



is a function which is *total* (so, defined for every natural number) in so there is

* (meaning is not defined at , since and it means we are not enumerating the current function inside the natural numbers, which we assume we can always do since is countable; so there is the contradiction)
* (again, not defined in and we do not enumerate the function assuming we can, hence another contradiction)

Since is distinct from all the functions in the enumeration, it demonstrates that the set of all functions is uncountable because it cannot be put in one-to-one correspondence with the set of natural numbers .

Summing up in math notation (there are more function than natural numbers, even though finite algorithms are as many as natural numbers):

Note that we can’t count non-computable functions, so:

We conclude that:

* no computational model can compute all functions
* there are more non-computable than computable functions

# URM Computability

To give a good notion of computability, we must choose a good model of computation, inducing a class of algorithms and computable functions. There can be:

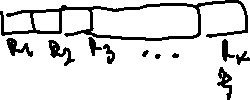
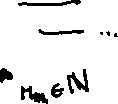
* Turing Machines (finite-state control, reading, writing, initial/final configuration)
* -calculus (a design of programming where one designs/applies functions based on primitives)
* Partial recursive functions (functions calculated with specific function that build partially)
* Canonical deduction systems (system used to create proofs logically via connectives and trees)
* URM (Unlimited Register Machines)

Whatever the model, we may concern if a specific theory may be valid for the specific model.

According to the Church-Turing thesis, a function is computable by an *effective procedure* if and only if it’s computable by a Turing machine (we resort to this to shorten the proof that a certain function is computable and it’s used informally as notion of effectiveness, then must be supported by evidence). This says that a function if computationally robust and we can choose whatever model one likes.

The notion of computable function will be formalized by using the URM-machine, abstraction based on the Von Neumann’s model. It has many characteristics:

* *memory is unbounded*, using an infinite number of *registers* storing each a natural number (where a sequence of registers is called *configuration*);



* it executes a program, based on a finite list of instructions (and a *computing agent* able to execute it);

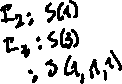
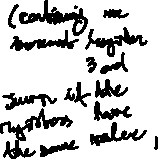


* it has arithmetic instructions, characterized by the fact that the instruction to be executed in the next step is the one following the current instruction in the program. They are:
  + *zero* , which sets the content of register to zero:
  + *successor* , which increments by 1 the content of register :
  + *transfer* , which transfers the content of register into , which staying untouched:
* another last instruction we have is:
  + *conditional jump*: , which compares the content of register and , so:
    - if then jumps to (jumps to -th instruction)
    - otherwise, it will continue with the next instruction

The computation:

* starts from an initial configuration of registers and executes
* terminates if
  + the instruction to be executed does not exist
  + it’s the last instruction
  + you jump out of the program yourself

An example might be the following one:



In LaTeX form, to not kill your eyes that much, coming from the notes:

Immagine che contiene testo, Carattere, numero, linea

Descrizione generata automaticamente

As we’re using the Church-Turing thesis, we’re defining a machine, so we must describe which *states* it has: there is a *register configuration* , taking the register content and index the next instruction via a program *counter* . Also, *operational semantics* can be defined via .

A computation can possibly diverge (not terminate); consider for instance this program:

Immagine che contiene testo, Carattere, linea, schermata

Descrizione generata automaticamente

Immagine che contiene numero, Carattere, linea, testo

Descrizione generata automaticamenteLet be an URM program. Given a sequence of natural numbers , indicates the computation of from :

* if the computation eventually terminates (*halts*)
* if the computation diverges (*never halts*)

We work on computations that start from an initial configuration where only a *finite number of registers contain a non-zero value*. So, given denotes for .

The notation then extends to the previous ones, stating that at the end of a program we will have a valid value 🡪 for and in final configuration .

For URM-computable functions, given a function (possibly partial), we say is URM-computable if there is a URM program such that, , and .

In words, for any input tuple of natural numbers, if you run the URM program on this input it will eventually have a result equal to the output of the function for input. This way, computes .

We then define, as the classes of computable functions. Therefore is the union of all of them.

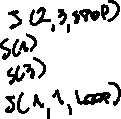
We next list some examples of URM-computable functions, providing the corresponding programs (courtesy of notes and my drawings from lesson notes):

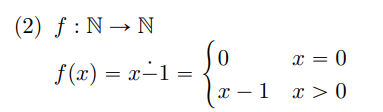
Immagine che contiene testo, Carattere, schermata, bianco

Descrizione generata automaticamente



Idea: Incrementing and until contain the same value, resulting in adding to the content of . Specifically:



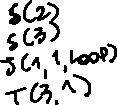




Now, let's analyze how this program works:

* If , it will jump to instruction , which presumably indicates the end of the program.
* If , it will go through instructions and , effectively setting to .
* If , it will go through instructions , and , effectively setting to .





The core concept behind this program is to continually subtract from the input value, considering the partial nature of the function. This means that the program might not always terminate, even if the function is computable, or it might terminate when the function is not computable.

In this specific example, the program checks if two values are equal; if they are, it jumps to a different instruction. If they are not equal, it subtracts one from the value. This subtraction continues until there is memory available for further operations.

Courtesy of notes (slightly different, but same example):

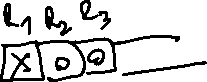
Immagine che contiene testo, schermata, Carattere, algebra

Descrizione generata automaticamente

Let’s consider a different function:

Immagine che contiene testo, Carattere, bianco, numero

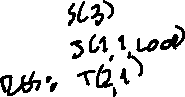
Descrizione generata automaticamente



The function behaves as follows:

* If the input number is even, the function returns half of (store an increasing even number in then storing its’ half in )
* If the input number is odd, the function does not terminate (indicated by the symbol ).

The program continues executing these instructions in a loop. If the initial input is even, it will eventually reach a point where equals the even number in , and it will jump to instruction , halving the input . If is initially odd, the program keeps increasing the even number in , and it never reaches the halting condition, resulting in a non-termination, as indicated by .



Given a program , for some fixed number of parameters , there exists a unique function computed by that we denote as follows as . More precisely:

Immagine che contiene testo, Carattere, linea, calligrafia

Descrizione generata automaticamente

In words: given a fixed number of parameters, the program halts if there is a final character of computation, otherwise the function will terminate when the program terminates.

Remember:

* a program terminates or not, a function is defined or not. A function is not computing, only the program does (they are correlated, of course).
* the same function can be computed via different algorithms, which means different problems

Question for us: given how many computing are there computing ?

* We can have infinitely many if and the only if the function is computable
* Answer: or infinitely many

Exercise

Consider , class of URM machines without transfer instructions (so, no ). We indicate the class of URM computable functions. How does compare to ? (in math notation, )

(Thoughts)

We can use as we can zero in and increment the register until it reaches .

The idea is:

(if the registers are equal, it exits the loop)

(back the program to the loop beginning)

In plain terms, this program aims to achieve a similar effect as the transfer instruction by repeatedly incrementing the value in register until it matches the value in register . Once they are equal, it exits the loop.

Proof

We show that . Let computable in i.e. there is in . Just observe that is also a URM program . As said in thoughts, ideally we replace the transfer instruction as the step with the previous subroutine:

Immagine che contiene testo, Carattere, bianco, design

Descrizione generata automaticamente

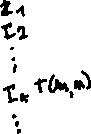
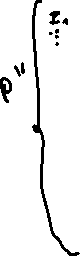
Let then . Hence there is a URM program such that .

We show there exists machine such that .

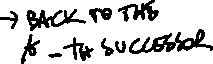
Remember: We’re assuming is well formed: if it terminates, it will at instruction

We proceed by induction on , which is the number of transfer instructions in . As notes say, we can assume, without loss of generality, that when a program halts it does so at the index of the last instruction plus one (induction logic at its core, in words).

* : trivial, with with no transfer instructions is already a program, hence
* : let be the URM program with transfer instructions. Hence, we replace with a jump to the subroutine:











We call this program which has transfer instructions and . By inductive hypothesis there program such that . Putting things together:



*Note*: for any URM-program there is a well-formed program computing the same function. In fact:

In words:

For any URM program , you can create an equivalent well-formed program that computes the same function. To do this, you replace any conditional jump instruction where is greater than the total number of instructions () with . This ensures that the program always jumps to a valid instruction and remains well-formed while achieving the same computational result as the original program.

Exercise

Variant machine where there are no traditional transfer instructions such that , but (swap instructions) like to exchange context of registers. How does compare to ? (is ? ) [To note: The teacher says the proof is simple and very similar to the one we did before]

Proof

Let us prove . We can replace the swap instructions with a few transfer instructions, formalizing how can be encoded in means of the routine. We can explain this in terms of how a swap instruction works in programming: we allocate a new register/new variable, we assign the variable to save to the new variable, then the new variable will get assigned to the second variable.

So, we create something like:

Formally, having a function we have a program as a program such that . Proceeding by induction:

* , the program is already a URM program and (such as before)
* , where the program, by injection, must have at least some transfer instructions to realize how a swap works. So, if only if this program uses both and instructions (the swap can’t be explained otherwise, and we need this statement to make this work correctly).
  + This way we can prove that with the swap, will have at least swap instructions, given the swap will be given via a jump instruction reaching the transfer instructions, hence creating the swap.
  + Inductively, there exists a URM program for and , concluding for having swaps recursively

Exercise

Consider without jump (where the apex indicates “minus minus”). How does compare with ? (is ?). [To note: it’s difficult, but one can start characterizing the shape of the functions in ]

Proof

Let us prove . As said from the hint, we can characterize the shape of functions inside of it. We first observe that is strictly contained in C, since there are total computable functions in C that cannot be computed by a machine due to the lack of jumps.

If we try to think logically, we have zero, transfer, successor as the available functions. This means this function is strictly linear and can only perform execution as a fixed sequence of inputs, potentially up until a constant number of operations. This says they always terminate, so we can have:

Or also (having as the constant we were discussing above, which will be inside ):

This can be proven by induction, but operating with something that makes the computation possible. In this case, it should be something with a number, just to prove can come out of it. So, recursively it must recreate the shape of . We will use a register describing the execution of a given number of steps, say , so is equal to .

* , we have , fine because with the base case is trivial, having already or which will turn it as alone
* , so in this case the only thing this can do is the other three functions:
  + , concluding trivially because the next step, having fr and we conclude we’re inside and this is hence respected. When , infact, the operation resets to and the function will keep its form
  + , so will allow us to get the sum of the instruction, given , again concluding by inductive hypothesis. This way the function will have its expanded form , because we continuously sum
  + ; this is uncertain because the function depends on two values this time around, ; when they are different from each other, the result way be unknown (one can be 1, the other we can’t know for sure, making the underlying function assume shapes unknown.)
    - When (or equal) we will know will do exactly the transfer of steps; otherwise, if , we won’t know what happen for sure, it can jump many instructions
    - The proof goes well if we assume we have exactly steps, so the function for or even can go exactly linearly assuming we will execute exactly *only that* number of steps. This happens because we will keep inside the function

# Decidable problems