

# Informatica Teorica



Anno Accademico 2022/2023

Fabio Zanasi

<https://www.unibo.it/sitoweb/fabio.zanasi>

Ottava lezione

# Nelle puntate precedenti

Abbiamo visto varie tecniche (in particolare: la **mapping-reducibility**), per dimostrare che un problema non è calcolabile (indecidibile o non-riconoscibile).

# In questa lezione

Ci concentriamo su risultati generali che chiariscano la nozione di calcolabilità ad ampio raggio.

Dimostriamo che tutti i problemi decidibili, in un certo senso, triviali (Teorema di Rice).

Dimostriamo che ci sono molti più problemi non riconoscibili che riconoscibili  
(usando l'argomento diagonale di Cantor).

# Teorema di Rice

# Quali risultati positivi?

Abbiamo visto che non possiamo decidere se una TM

- ferma su un dato input
- ferma su input vuoto (la stringa vuota)
- è equivalente a un'altra TM.

Allora, quali problemi riguardanti le TM sono decidibili?

# Quali risultati positivi?

Per esempio, possiamo decidere se una TM data:

- ferma sempre in un certo numero di passi
- ha meno di un certo numero di stati

Cos'altro?

# Ripasso: i linguaggi

I **linguaggi** sono sottoinsiemi di  $\Sigma^*$ . Senza perdita di generalità (modulo la corretta codifica di  $\Sigma^*$  in  $\{0,1\}^*$ ), possiamo assumere  $\Sigma = \{0,1\}$ .

Il **linguaggio della TM  $\mathcal{M}$**  è definito come:

$$L_{\mathcal{M}} = \{x \in \{0,1\}^* \mid \mathcal{M} \text{ accetta } x\}.$$

# Proprietà di TM-linguaggi

Una **proprietà di linguaggio**  $P$  è una funzione da un insieme di TM a  $\{0,1\}$  (falso/vero), tale che  $L_{\mathcal{M}} = L_{\mathcal{M}'}$  implica  $P(\mathcal{M}) = P(\mathcal{M}')$ .

Questo assicura che  $P$  dipenda solo dal linguaggio descritto dalla macchina. Per esempio: ``ferma in 42 step'' non è proprietà del linguaggio.

Questa proprietà è **non-triviale** se esiste una TM  $\mathcal{M}$  tale che  $P(\mathcal{M}) = 1$  e una TM  $\mathcal{M}'$  tale che  $P(\mathcal{M}') = 0$ .

Formalmente, identifieremo le TM che soddisfano la proprietà  $P$  con l'insieme  $\{y \in \Sigma^* \mid y = \text{code}(\mathcal{M}) \text{ e } P(\mathcal{M}) = 1\}$ .

# Esempi

$\{y \mid y = \text{code}(\mathcal{M}) \text{ e } L_{\mathcal{M}} \text{ è finito}\}$

$\{y \mid y = \text{code}(\mathcal{M}) \text{ e } L_{\mathcal{M}} \text{ è vuoto}\}$

$\{y \mid y = \text{code}(\mathcal{M}) \text{ e } L_{\mathcal{M}} = L\}$

Queste sono *proprietà di linguaggio* non-triviali:  
alcune TM le hanno, altre no.

$\{y \mid y = \text{code}(\mathcal{M}) \text{ e } \mathcal{M} \text{ riconosce qualche linguaggio}\}$

Questa è una proprietà di linguaggio triviale: per definizione, ogni TM riconosce un linguaggio.

# Il Teorema di Rice

**Teorema.** Se  $P$  è *language property* non triviale, allora il problema “ $\mathcal{M}$  ha proprietà  $P$ ” è indecidibile.

**Strategia di dimostrazione.** Per contraddizione.

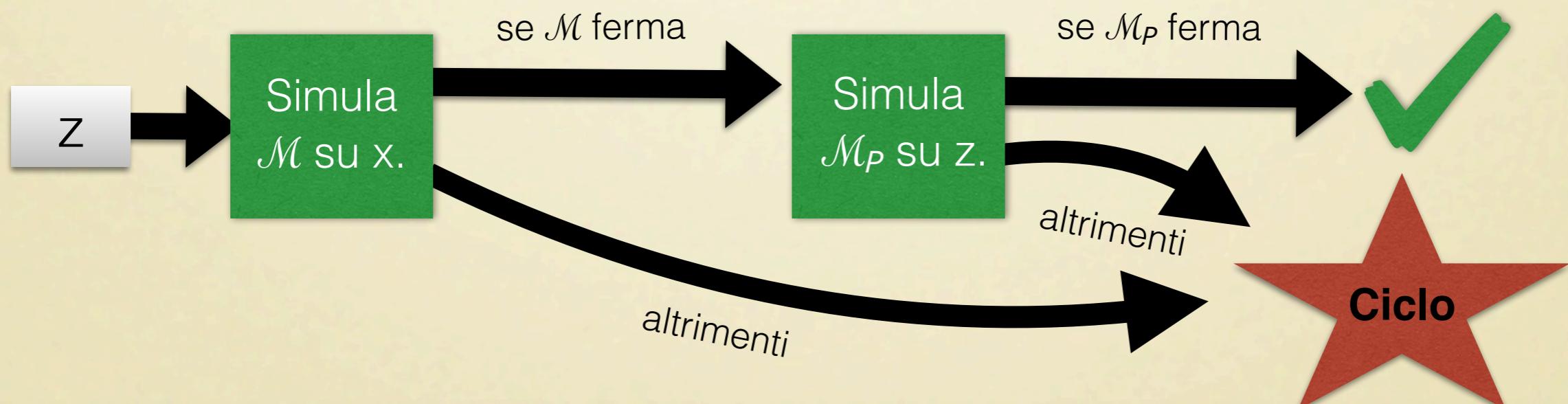
Dimostriamo che se “ $\mathcal{M}$  ha proprietà  $P$ ” fosse decidibile, allora il problema della fermata sarebbe decidibile.

# Teorema di Rice

$\mathcal{M}_\emptyset$  è una TM che riconosce il linguaggio vuoto.

## Dimostrazione.

- Considera una proprietà P. Assumiamo  $P(\mathcal{M}_\emptyset) = 0$ .
- Poiché P è non-triviale, possiamo considerare una TM  $\mathcal{M}_P$  tale che  $P(\mathcal{M}_P) = 1$ .
- Fissiamo  $\mathcal{M}$  e  $x$  come parametri e costruiamo la seguente TM  $\mathcal{M}_{\mathcal{M},x}$ :



# Teorema di Rice

## Dimostrazione

$\mathcal{M}$   
ferma  
su  $x$ .

$\Rightarrow \mathcal{M}_{\mathcal{M},x}$  ferma su  $z$   
ogni volta che  
 $\mathcal{M}_P$  ferma su  $z$ .

$$\Rightarrow L_{\mathcal{M}_P} = L_{\mathcal{M}_{\mathcal{M},x}}$$

$P(\mathcal{M}_{\mathcal{M},x}) = P(\mathcal{M}_P)$ ,  
allora  $\mathcal{M}_{\mathcal{M},x}$  ha  
proprietà  $P$ .

$\mathcal{M}$  non  
ferma su  $x$ .

$\Rightarrow \mathcal{M}_{\mathcal{M},x}$  non  
ferma mai  
su  $z$ .

$$\Rightarrow L_{\mathcal{M}_{\mathcal{M},x}} = \emptyset$$

$P(\mathcal{M}_{\mathcal{M},x}) = P(\mathcal{M}_\emptyset)$ , allora  
 $\mathcal{M}_{\mathcal{M},x}$  non ha proprietà  
 $P$ .

Conclusione: se potessimo decidere se  $\mathcal{M}_{\mathcal{M},x}$  ha la proprietà  $P$ , potremmo decidere il problema della fermata.

Allora,  $\{y \mid y = \text{code}(\mathcal{M}) \text{ e } P(\mathcal{M}) = 1\}$  è indecidibile.

# Teorema di Rice

## Dimostrazione.

Abbiamo assunto  $P(\mathcal{M}_\emptyset) = 0$ . Se  $P(\mathcal{M}_\emptyset) = 1$ ?

In questo caso, ripetiamo lo stesso argomento, ma per la proprietà  $\neg P$  (“ $\mathcal{M}$  non ha la proprietà  $P$ ”).

Osserva che questo funziona perché:

- dato che  $P$  è non-triviale, anche  $\neg P$  è non-triviale.
- dato che  $P(\mathcal{M}_\emptyset) = 1$ , allora  $\neg P(\mathcal{M}_\emptyset) = 0$ .

Concludiamo che  $\{y \mid y = \text{code}(\mathcal{M}) \text{ e } \neg P(\mathcal{M}) = 1\}$  è indecidibile. Questo implica che anche  $\{y \mid y = \text{code}(\mathcal{M}) \text{ e } P(\mathcal{M}) = 1\}$  sia indecidibile.

# Avvertenza: Proprietà Decidibili

Attenzione a possibili fraintendimenti: il Teorema di Rice riguarda proprietà di **linguaggio**, non proprietà **algoritmiche**; riguarda funzioni (specifiche), non programmi (implementazioni).

Per esempio, non possiamo usare il teorema di Rice per derivare l'indecidibilità del problema della fermata (e simili).

# Avvertenza: Proprietà Decidibili

In generale, ci sono tre tipi di proprietà riguardo le TM:

- **Proprietà di linguaggio.** Quelle non triviali sono indefinibili (teorema di Rice).
- **Proprietà strutturali.** Per esempio “ $\mathcal{M}$  ha 13 stati”. Queste sono tipicamente decidibili poiché si possono verificate staticamente sulla (codifica della) descrizione di TM.
- **Proprietà comportamentali (o algoritmiche).** Per esempio “ $\mathcal{M}$  non si muove a sinistra su input 0101”. Alcune sono decidibili, altre no, e la classificazione non è ovvia.

# La cardinalità dei problemi irrisolvibili

# Obiettivo

Vogliamo mostrare che la maggior parte dei linguaggi non sono riconoscibili (e, quindi, indecidibili).

Il nostro argomento consiste nel mostrare che ci sono ‘molti’ più linguaggi che TM.

Per rendere la dimostrazione precisa, introduciamo un metodo che si rivela utile anche in altri contesti:  
l’argomento diagonale di Cantor.

# Insiemi infiniti numerabili

Un insieme  $S$  è **infinito numerabile** se c'è una funzione totale biettiva  $f: \mathbb{N} \rightarrow S$ .

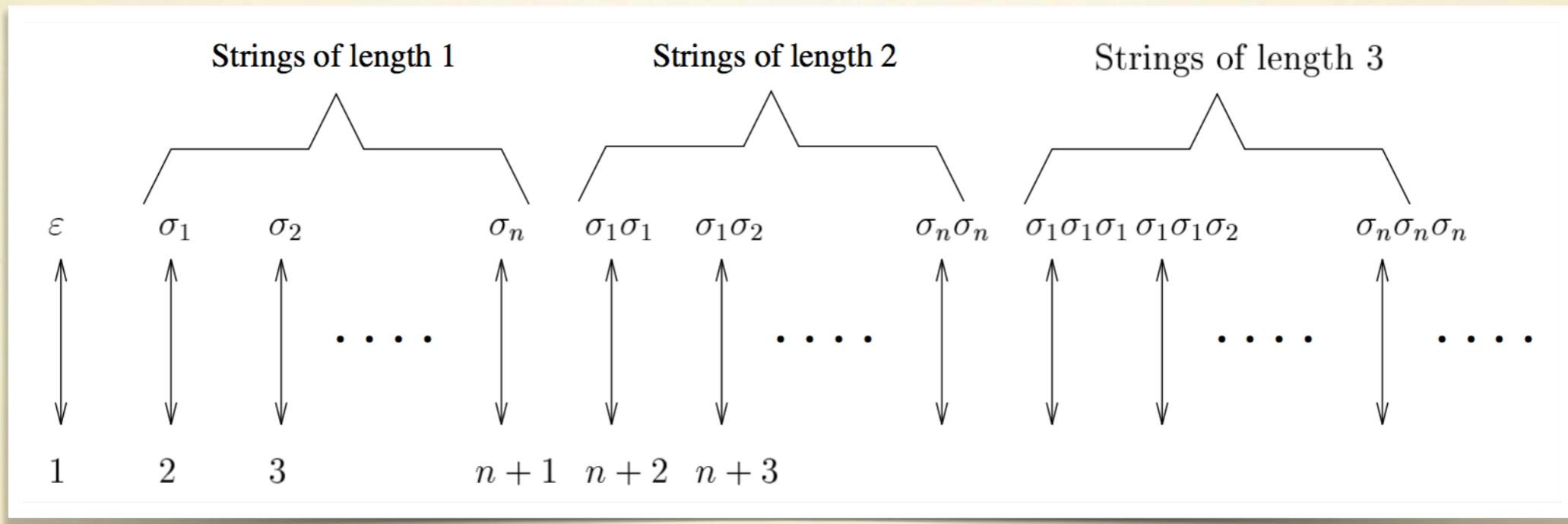
Idea: un insieme infinito numerabile ha tanti elementi quanti sono i numeri naturali.

Esempio: l'insieme dei numeri dispari  $D$  con la funzione  $f: \mathbb{N} \rightarrow D$  definita come  $f(n) = 2n - 1$ .

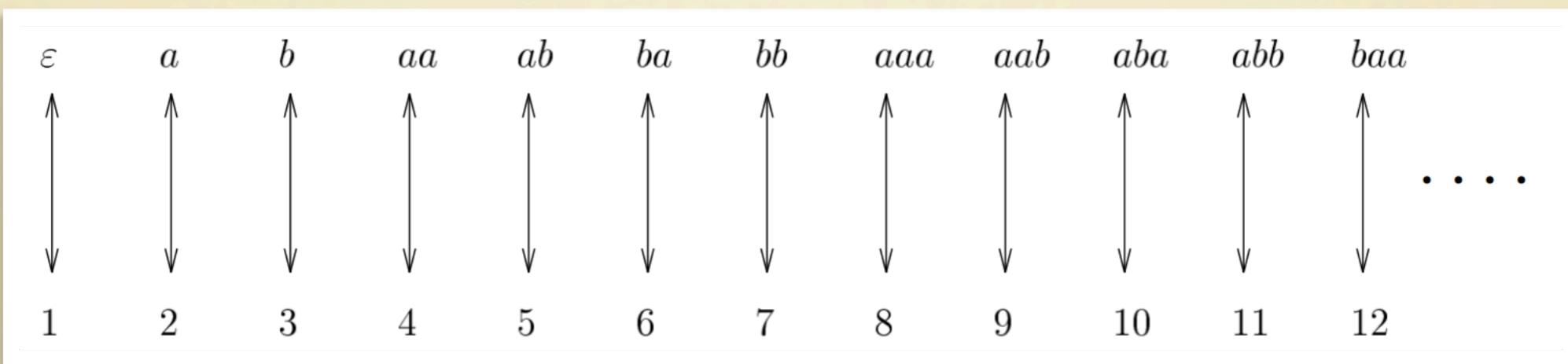
**Lemma.** Se  $S_1$  e  $S_2$  sono infiniti numerabili,  $S_1 \cup S_2$  è infinito numerabile.

# Insiemi infiniti numerabili

Esempio: l'insieme  $\Sigma^*$  di stringhe su alfabeto finito  $\Sigma$ .  
Assumi  $|\Sigma| = n$ , la biezione con  $\mathbb{N}$  è costruita come:



Per esempio, sia  $\Sigma = \{a,b\}$ :



# Counting TM

Abbiamo visto che ogni TM può essere codificata come una stringa per un alfabeto  $\Sigma$  con  $|\Sigma| = 2$  (e.g.,  $\Sigma = \{0,1\}$ ).

Allora, l'**insieme di tutte le TM** è infinito numerabile.

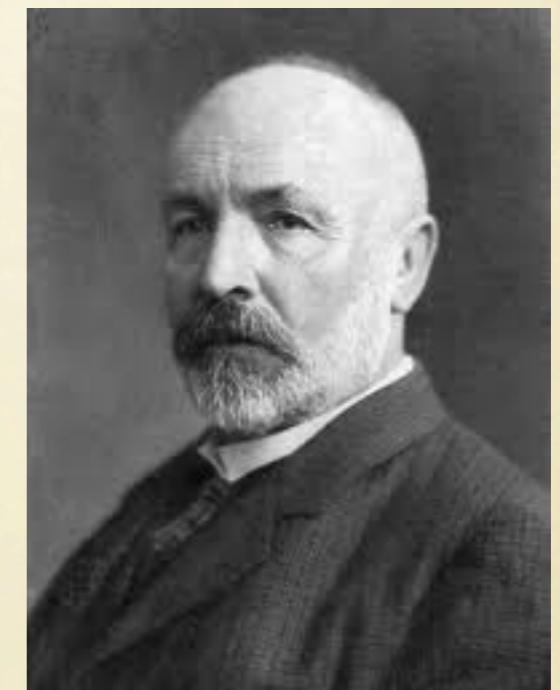
Anche l'**insieme di tutti i linguaggi riconoscibili** è infinito numerabile. Questo perché, per definizione, un linguaggio è riconoscibile se c'è una TM che lo riconosce.

Per la stessa ragione, l'**insieme delle funzioni  $\mathbb{N} \rightarrow \mathbb{N}$  computabili** da una TM è infinito numerabile.

# Insiemi non-numerabili

Come visto, insiemi infiniti numerabili (compreso l'insieme delle TM) sono gli insiemi con tanti elementi quanti sono i numeri naturali (possono essere “contati”).

Georg Cantor (1845-1918) insegna che ci sono infiniti insiemi **non-numerabili (*uncountable*)**, con “più” elementi di  $\mathbb{N}$ .



Cantor introduce una tecnica, detta **diagonalizzazione**, per mostrare che un insieme è non-numerabile e la usa per mostrare che i numeri reali sono non-numerabili.

# I linguaggi sono non-numerabili

Sia  $S_\Sigma$  l'insieme di tutti i linguaggi sull'alfabeto finito  $\Sigma$ .

**Teorema.** L'insieme  $S_\Sigma$  non è numerabile.

**Dimostrazione.** Ricorda che un linguaggio  $L$  è un sottoinsieme di  $\Sigma^*$ . Abbiamo già visto che  $\Sigma^*$  è infinito numerabile, quindi possiamo scriverlo come  $\Sigma^* = \{\sigma_1, \sigma_2, \sigma_3, \dots\}$ .

Allora un linguaggio, diciamo  $L_1 = \{\sigma_1, \sigma_4\}$ , può essere rappresentato come una riga in una tabella:

	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_4$	$\sigma_5$	$\dots$
$L_1$	1	0	0	1	0	$\dots$

# I linguaggi sono non-numerabili

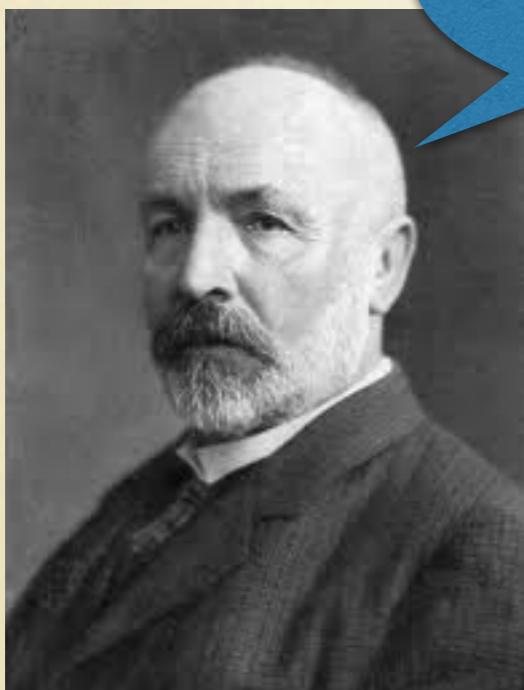
*Ciascun* linguaggio su  $\Sigma$  può essere rappresentato in questo modo.

Per contraddizione, assumi che  $S_\Sigma$  sia un insieme numerabile. Allora possiamo assegnare un numero naturale ai suoi elementi, così che ogni  $L_i \in S_\Sigma$  appare come riga a destra.

	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_4$	$\sigma_5$	$\dots$
$L_1$	1	0	0	1	0	$\dots$
$L_2$	0	1	1	0	1	$\dots$
$L_3$	0	0	0	0	0	$\dots$
$L_4$	1	1	1	0	1	$\dots$
$L_5$	1	1	1	1	1	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

# I linguaggi sono non-numerabile

Davvero ogni linguaggio  $L$  di  $S_\Sigma$  appare in una riga?



No! Guarda la  
diagonale.

	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_4$	$\sigma_5$	• • •
$L_1$	1	0	0	1	0	• • •
$L_2$	0	1	1	0	1	• • •
$L_3$	0	0	0	0	0	• • •
$L_4$	1	1	1	0	1	• • •
$L_5$	1	1	1	1	1	• • •
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•
•	•	•	•	•	•	•

# I linguaggi sono non-numerabili

Davvero ogni elemento  $L$  di  $S_{\Sigma}$  compare su una riga?

Definisci  $L$  come  $00110\dots$ ,  
allora  $\sigma_i \in L$  sse  $\sigma_i \notin L_i$ .

Quindi  $L$  è diverso da ogni  
linguaggio  $L_i$  sulla riga.

Dunque  $L$  non può essere in  
una riga! **Contraddizione.**

Quindi,  $S_{\Sigma}$  non è  
numerabile.

	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_4$	$\sigma_5$	$\dots$
$L_1$	1	0	0	1	0	$\dots$
$L_2$	0	1	1	0	1	$\dots$
$L_3$	0	0	0	0	0	$\dots$
$L_4$	1	1	1	0	1	$\dots$
$L_5$	1	1	1	1	1	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

# Riassumendo...

Dato un alfabeto finito  $\Sigma$ , abbiamo visto che:

- l'insieme di linguaggi riconoscibili da una TM è infinito numerabile.
- l'insieme di tutti i linguaggi è non-numerabile.

Ne abbiamo visti alcuni:  
 $HALT^-$ ,  $EQ$ ,  $EQ^-$

Quindi, esistono linguaggi che non sono riconoscibili da alcuna TM.

**Ulteriore domanda: quanti linguaggi non riconoscibili ci sono?**

# Quanti insiemi non riconoscibili?

La nostra risposta deriva da un risultato generale.

**Teorema.** Se  $S$  è un insieme infinito, non-numerabile e  $S'$  è un sottoinsieme infinito numerabile di  $S$ , allora  $S \setminus S'$  non è infinito numerabile.

**Dimostrazione.** Assumi  $S \setminus S'$  sia infinito numerabile. Allora, poiché i linguaggi infiniti numerabili sono chiusi per unione (vedi Lemma precedente),  $(S \setminus S') \cup S' = S$  è numerabile. Contraddizione!

**Corollario.** L'insieme di linguaggi non riconoscibili non è infinito numerabile — allora ci sono più linguaggi non-riconoscibili che riconoscibili.