Quello che ci proponiamo di fare adesso è di mostrare una serie di problemi che sono algoritmicamente insolubili, cioè che non ammettono un algoritmo generale di soluzione.

Ci limiteremo ad enunciare i risultati senza darne le dimostrazioni relative in questo momento.

<u>Definizione</u>

Il <u>problema della parola</u> per un processo di (semi) Thue Π è il problema di determinare per ogni coppia data di parole u, v sull' alfabeto di Π se $u \overset{*}{\underset{\Pi}{\Rightarrow}} v$.

Si dimostra che:

Esiste un processo di (semi) Thue su un alfabeto di due simboli il cui problema della parola è insolubile. Inoltre tutte le regole di riscrittura $g \to h$ di tale processo sono tali che $g, h \neq 0$.

SISTEMI DI CORRISPONDENZA DI POST

Un sistema di corrispondenza di Post è dato da un alfabeto finito A ed un insieme finito di coppie di parole $(w_i, \overline{w}_i)$ su A $(1 \leq i \leq m)$.

SI DICE CHE IL SISTEMA AMMETTE UNA SOLUZIONE SE ESISTE UNA SUCCESSIONE FINITA DI INDICI DI COPPIE

$i_1, i_2, \ldots, i_k$, $1 \leq i_j \leq m$ tale che

$$w_{i_1} w_{i_2} \ldots w_{i_k} \equiv \overline{w}_{i_1} \overline{w}_{i_2} \ldots \overline{w}_{i_k} \qquad (*)$$

Una parola come quella che compare in $(*)$ si dice PAROLA SOLUZIONE DEL SISTEMA DI CORRISPONDENZA DI POST.

Una visualizzazione efficace si può ottenere con una rappresentazione che faccia uso di tessere di domino sulle quali sono stampate le due parole della coppia.

Quindi $(w_1, \overline{w}_1)$ viene rappresentata come $\begin{array}{|c|}\hline w_1 \\ \hline \overline{w}_1 \\ \hline\end{array}$

Consideriamo un esempio.

Sia il seguente sistema di Post

(s_2, s_2) ; $(s_1 s_2, s_2 s_1 s_1)$; $(s_1 s_2, s_2 s_1 s_2)$; $(s_1 s_2 s_1, s_1)$

Una rappresentazione tipo-domino è la seguente:

s_2	$s_1 s_2$	$s_1 s_2$	$s_1 s_2 s_1$
s_2	$s_2 s_1 s_1$	$s_2 s_1 s_2$	s_1
1	2	3	4

Il problema della corrispondenza di Post consiste nel mettere una accanto all'altra le tessere (eventualmente con ripetizioni) in modo tale che la parola che si legge nella traccia superiore è identica a quella che si legge nella traccia inferiore.

LA SEGUENTE SUCCESSIONE DI TESSERE, FORNISCE UNA SOLUZIONE: 4-2-1-3.

Infatti:

$s_1 s_2 s_1$	$s_1 s_2$	s_2	$s_1 s_2$
s_1	$s_2 s_1 s_1$	s_2	$s_2 s_1 s_2$

In realtà una soluzione si ha già utilizzando solo tre tessere:

$s_1 s_2 s_1$	$s_1 s_2$	$s_1 s_2$
s_1	$s_2 s_1 s_1$	$s_2 s_1 s_2$

il che mostra che le tessere $\boxed{\frac{a}{a}}$ sono superflue

Consideriamo un altro esempio:

Sia il seguente sistema di corrispondenza di Post:

$$(s_1, s_1 s_1) \, ; \, (s_2 s_2, s_2) \, ; \, (s_1, s_2 s_2)$$

rappresentato da:

s_1	$s_2 s_2$	s_1
$s_1 s_1$	s_2	$s_2 s_2$
1	2	3

È facile convincersi che non esiste alcuna soluzione che fa uso di due sole tessere e neanche di tutte e tre le tessere usate una sola volta (la parola di sopra sarebbe di lunghezza 4, quella di sotto di lunghezza 5).

Dobbiamo quindi necessariamente usare almeno una tessera più di una volta.

Una soluzione al problema è fornita da:

s_1	s_1	$s_2 s_2$	$s_2 s_2$
$s_1 s_1$	$s_2 s_2$	s_2	s_2
1	3	2	2

La ricerca di una soluzione del sistema di Post si chiama:

PROBLEMA DELLA CORRISPONDENZA DI POST

Post stesso nel 1946 ha dimostrato che IL PROBLEMA DELLA CORRISPONDENZA È INDECIDIBILE, CIOÈ NON ESISTE NESSUN <u>ALGORITMO</u> IN GRADO DI DECIDERE SE UN ARBITRARIO SISTEMA DI CORRISPONDENZA DI POST AMMETTE O NO UNA SOLUZIONE.

GRAMMATICHE

— Non esiste nessun algoritmo per determinare, data una grammatica T', se

- $L(T') = \phi$
- $L(T')$ è infinito
- data una parola v, se $v \in L(T')$

— Non esiste nessun algoritmo per determinare, date due grammatiche T'_1 e T'_2, se

- $L(T'_1) \subseteq L(T'_2)$
- $L(T'_1) = L(T'_2)$

LA GERARCHIA DI CHOMSKY

Esaminiamo adesso il problema di fornire modelli di calcolo, via via più specifici, sempre nel contesto dei processi di Semi-Thue.

Questo può essere fatto imponendo delle restrizioni via via più forti ad una <u>grammatica a struttura di frase</u> che verrà indicata in questo contesto come <u>grammatica di tipo 0</u>

UNA GRAMMATICA Γ È DETTA DI <u>TIPO 1</u> O <u>GRAMMATICA CONTESTUALE</u> SE, PER CIASCUNA PRODUZIONE $g \to \bar{g}$ DI Γ SI HA CHE $|g| \leq |\bar{g}|$.

UNA GRAMMATICA Γ È DETTA DI <u>TIPO 2</u> O <u>LIBERA DA CONTESTO</u> SE CIASCUNA PRODUZIONE È DEL TIPO $X \to h$, DOVE X È UNA VARIABILE ED h UNA QUALSIASI PAROLA SULL'ALFABETO TOTALE.
(In generale h può essere la parola vuota; se nessuna produzione di Γ è tale che $h = 0$ allora Γ si dice <u>GRAMMATICA LIBERA DA CONTESTO POSITIVA</u>)

UNA GRAMMATICA Γ È DETTA DI <u>TIPO 3</u> O <u>REGOLARE</u> SE CIASCUNA DELLE SUE PRODUZIONI È DEL TIPO $U \to aV$ OPPURE $U \to a$, CON U e V VARIABILI ED a SIMBOLO TERMINALE

A ciascun tipo di grammatica corrisponde una classe di linguaggi generati.

Si ha che:

Un linguaggio L è regolare (di tipo 3)	SE	esiste una grammatica regolare T' tale che $L = L(T')$ o $L = L(T') \cup \{o\}$
Un linguaggio L è libero da contesto (tipo 2)	SE	esiste una grammatica libera da contesto tale che $L = L(T')$ o $L = L(T') \cup \{o\}$
Un linguaggio L è contestuale (tipo 1)	SE	esiste una grammatica contestuale tale che $L = L(T')$ o $L = L(T') \cup \{o\}$.
Un linguaggio L è r.e. (o di tipo 0)	SE	esiste una grammatica a struttura di frase T' che lo genera.

Semplicemente guardando le definizioni date si vede che:

- Ogni linguaggio regolare è libero da contesto
- Ogni linguaggio libero da contesto è contestuale
- Ogni linguaggio contestuale è r.e.

Si ha allora l'inclusione

$$L_3 \subset L_2 \subset L_1 \subset L_0$$

avendo indicato con L_i ciascuna famiglia di linguaggi.

Le inclusioni sono tutte in senso stretto come è, correttamente, indicato, in quanto è possibile portare esempi di linguaggi che appartengono alla classe più vasta ma non a quella immediatamente inferiore nella gerarchia.

Passiamo, molto rapidamente, in rassegna, alcuni risultati riguardanti la decidibilità o l'indecidibilità:

- ESISTE UN ALGORITMO IN GRADO DI DETERMINARE, DATA UNA GRAMMATICA LIBERA DA CONTESTO T' ED UNA PAROLA u, SE u APPARTIENE O NO AD L(T').

- ESISTE UN ALGORITMO PER DECIDERE, DATA UNA GRAMMATICA LIBERA DA CONTESTO T', SE L(T') È O NON È VUOTO.

- ESISTE UN ALGORITMO PER DECIDERE, DATA UNA GRAMMATICA LIBERA DA CONTESTO T', SE L(T') È FINITO O INFINITO

La situazione, però, non è del tutto felice neanche qui.

Si ha infatti che:

— NON ESISTE NESSUN ALGORITMO PER DETERMINARE, DATE DUE GRAMMATICHE LIBERE DA CONTESTO T'_1 e T'_2, SE

$$L(T'_1) \cap L(T'_2) = \phi$$

— NON ESISTE NESSUN ALGORITMO PER DECIDERE SE UNA DATA GRAMMATICA LIBERA DA CONTESTO È <u>AMBIGUA</u>.

(Una grammatica libera da contesto T' si dice <u>ambigua</u> se esiste almeno una parola appartenente al linguaggio generato da T' che può essere ottenuta mediante due derivazioni <u>di un certo tipo standard</u>)

Per arricchire il quadro dei problemi che è possibile porre nell'ambito di una <u>pura</u> teoria della calcolabilità, presentiamo ora (per il momento senza dimostrazione) altri due teoremi, uno di tipo positivo ed uno di tipo negativo.

I. ESISTE UN NUMERO e TALE CHE LA FUNZIONE DI UNA SOLA VARIABILE CALCOLATA DAL PROGRAMMA NUMERO e SI COMPORTA COME SEGUE:

$$\Phi_e(x) = e$$

CIOÈ, QUALUNQUE SIA LA VARIABILE D'INGRESSO DÀ COME USCITA e.

Tale funzione sembrerebbe essere soltanto una funzione costante. La sua novità risiede nel fatto che, poiché nella nostra teoria i numeri sono anche codici di programmi, possiamo interpretare tale codice e sottolineare il fatto che TALE FUNZIONE "<u>REAGISCE A QUALSIASI STIMOLO ESTERNO</u>" RIPROPONENDO IL SUO STESSO CODICE.

Questo aspetto "autoriproduttivo" è potenzialmente interessante sia in direzione tecnologica (avere una teoria degli automi in grado di produrre copie di sé stessi) sia in direzione biologica collegandosi, in particolare ad impostazioni teoriche che assegnano la specificità dei sistemi biologici alla loro tendenza ad automantenere sé stessi.

II.

Sia adesso T' un insieme di funzioni parzialmente calcolabili di una sola variabile.
Chiamiamo <u>insieme di indici</u> di T' l'insieme:

$$R_{T'} = \{ t \in N \mid \Phi_t \in T' \}$$

Teorema (Rice)

$R_{T'}$ È RICORSIVO SOLO NEL CASO IN CUI T' È L'INSIEME VUOTO OPPURE È L'INSIEME DI TUTTE LE FUNZIONI PARZIALMENTE CALCOLABILI.

ovvero (in modo equivalente)

SIA T' UN INSIEME DI FUNZIONI PARZIALMENTE CALCOLABILI DI UNA VARIABILE. SE ESISTONO DUE FUNZIONI PARZ. CALCOLABILI f E g TALI CHE $f \in T'$ e $g \notin T'$ ALLORA $R_{T'}$ NON È RICORSIVO.

L'idea allettante ~~alla ricerca di un risultato~~ ~~ancora dietro tale risultato~~ è quella di fissare una qualsiasi proprietà delle funzioni e poi chiedere una verifica "<u>meccanica</u>" di tale proprietà per una specifica funzione.
Il teorema di Rice ci dice che tale idea, per quanto allettante, non è realizzabile.