Automi e Linguaggi Formali – A.A. 2016/17
II Compitino – Versione 1

Esercizio 1. Descrivete in italiano il linguaggio accettato dalla macchina di Turing definita dalla seguente tabella di transizione:

	0	1	X	Y	B
q_0	(q_2, X, R)	(q_1, X, R)		(q_0, Y, R)	(q_4, B, R)
q_1	(q_3, Y, L)	$(q_1, 1, \mathrm{R})$		(q_1, Y, R)	
q_2	$(q_2, 0, \mathrm{R})$	(q_3, Y, L)		(q_2, Y, R)	
q_3	$(q_3, 0, \mathrm{L})$	$(q_3, 1, \mathrm{L})$	(q_0, X, R)	(q_3, Y, L)	
$*q_4$					

Soluzione: la TM accetta il linguaggio costituito dalle stringhe con ugual numero di 0 e 1 (indipendentemente dal loro ordine). Ad es. 10, 0110,111000, etc.

Esercizio 2. (a) Definite una macchina di Turing M per il linguaggio $\{0^n1^n0^n \mid n \geq 1\}$. Definite la specifica formale della TM M, riportando δ come tabella di transizione.
(b) Descrivete in italiano il linguaggio $L(M)$ accettato dalla TM M.

Soluzione (una tra le possibili):

	0	1	X	B
q_0	(q_1, B, R)		(q_5, X, R)	
q_1	$(q_1, 0, \mathrm{R})$	(q_2, X, R)	(q_1, X, R)	
q_2	$(q_2, 0, \mathrm{R})$	$(q_2, 1, \mathrm{R})$		(q_3, B, L)
q_3	(q_4, B, L)			
q_4	$(q_4, 0, \mathrm{L})$	$(q_4, 1, \mathrm{L})$	(q_4, X, L)	(q_0, B, R)
q_5			(q_5, X, R)	(q_6, B, R)
$*q_6$				

(b) Il linguaggio L(M) accettato dalla TM M è costituito dall'insieme delle stringhe formate da n zeri, seguiti da n simboli 1, seguiti da n zeri, con n maggiore o uguale a 1 (ad es. 010, 001100, etc.).

Esercizio 3. Scrivete le descrizioni istantanee della TM M definita nell'Esercizio 2 quando il nastro di input contiene: (a) 0011 (b) 001100 (c) 0110.

Soluzione:
(a) $q_0 0011 \vdash Bq_1 011 \vdash B0q_1 11 \vdash B0Xq_2 1 \vdash B0X1q_2 B \vdash B0Xq_3 1$ *si arresta senza accettare*
(b) $q_0 001100 \vdash Bq_1 01100 \vdash B0q_1 1100 \vdash B0Xq_2 100 \vdash B0X1q_2 00 \vdash B0X10q_2 0 \vdash$

$B0X100q_2B$ |- $B0X100q_2B$ |- $B0X10q_30B$ |- $B0X1q_40B$ |- $B0Xq_410B$ |- $B0q_4X10B$ |- Bq_40X10B |- $Bq_4B0X10B$ |- Bq_00X10B |- Bq_1X10B |- Xq_110B |- XXq_20B |- $XX0q_2B$ |- XXq_30B |- Xq_4XB |- q_4XXB |- q_4BXXB |- q_0XXB |- Xq_5XB |- XXq_5B |- $XXBq_6B$ accetta
(c) q_00110 |- q_1110 |- Xq_210 |- $X1q_20$ |- $X10q_2B$ |- $X1q_30$ |- Xq_41 |- q_4X1 |- q_4BX1 |- q_0X1 |- Xq_51 si arresta senza accettare

Esercizio 4. Indicate quali fra le seguenti istanze di PCP hanno soluzione. Ognuna è presentata sotto forma delle due liste *A* e *B*; le *i*-esime stringhe delle due liste sono corrispondenti per *i*=1, 2, etc.
(a) *A* = (11, 1010, 01); *B* = (101, 10, 10)
(b) *A* = (1, 10111, 10); *B* = (111, 10, 0)
(c) *A* = (ab, aab, ba); *B* = (abb, ba, aa)

Soluzione: (a) Non ha soluzione. Si crea un loop: 2,1,3,3,... (b) Ha soluzione: 2,1,1,3. (c) Non ha soluzione. La prima scelta obbligata è i=1; dopo possiamo scegliere solo 3; successivamente, ognuno dei tre indici non rende possibile proseguire.

Esercizio 5. Qual è la relazione tra linguaggi ricorsivi, RE e non RE? Fate un esempio di linguaggio RE e di linguaggio non RE (tra quelli visti a lezione) e datene una breve descrizione.

Soluzione: L'insieme dei linguaggi ricorsivi è un sottoinsieme dei linguaggi RE. I linguaggi ricorsivi (o decidibili) sono quelli per cui ho garanzia che la TM si arresterà, sia che accetti o che non accetti l'input. I linguaggi RE non ricorsivi sono quelli per cui ho garanzia che la TM si arresterà se l'input è nel linguaggio, ma nel caso in cui non lo sia potrebbe proseguire in eterno. Per i linguaggi non RE non abbiamo alcuna garanzia sull'arresto.
Un esempio di linguaggio RE è il linguaggio universale L_u, definito come L_u=L(U), dove U è la TM universale. Un esempio di linguaggio non RE è il linguaggio di diagonalizzazione L_d, definito come l'insieme delle stringhe w_i tali che w_i è in $L(M_i)$.

Esercizio 6. (a) Date la definizione della classe dei problemi NP-completi. (b) Quando invece possiamo definire un problema come NP-arduo? (c) A quale classe di problemi appartiene il problema della soddisfacibilità (SAT)?

Soluzione: (a) Un problema è NP-completo se è definito da un linguaggio L tale che: (1) L è NP, e (2) per ogni altro L' in NP esiste una riduzione polinomiale di L' a L. (b) Un problema si dice NP-arduo se è possibile dimostrare la sola condizione (2) della definizione di NP-completezza, mentre non si può dimostrare la (1). Di solito ci riferiamo a questa classe di problemi come "intrattabili". (c) CSAT è il problema di soddisfacibilità di espressioni booleana in forma CNF ed è NP-completo (come SAT).

Esercizio 7. Dite quali tra le seguenti affermazioni è corretta:

(a) Un’espressione booleana si dice in forma CNF se è formata da una o più clausole tra di loro in OR logico.
(b) L’espressione $(x \vee \neg y) \wedge (\neg x \vee z) \wedge (x \vee \neg z)$ è in 2-CNF.
(c) Ogni linguaggio accettato da una TM multinastro è ricorsivamente numerabile.
(d) Se un linguaggio *L* e il suo complemento sono RE, allora anche il complemento di *L* è ricorsivo.
(e) La trattazione dell’intrattabilità si basa sull’ipotesi (non dimostrata) che P=NP.
Soluzione: (a) non corretta (devono essere in AND logico); (b) corretta; (c) corretta; (d) corretta; (e) non corretta (si basa sull’ipotesi, non dimostrata, che P≠ NP).

Automi e Linguaggi Formali – A.A. 2016/17
II Compitino – Versione 2

Esercizio 1. Descrivete in italiano il funzionamento della TM definita dalla seguente tabella di transizione:

	0	1	*B*
q_0	(q_1, B, R)	(q_5, B, R)	
q_1	$(q_1, 0, \mathrm{R})$	$(q_2, 1, \mathrm{R})$	
q_2	$(q_3, 1, \mathrm{L})$	$(q_2, 1, \mathrm{R})$	$(q_4, \mathrm{B}, \mathrm{L})$
q_3	$(q_3, 0, \mathrm{L})$	$(q_3, 1, \mathrm{L})$	$(q_0, \mathrm{B}, \mathrm{R})$
q_4	$(q_4, 0, \mathrm{L})$	$(q_4, \mathrm{B}, \mathrm{L})$	$(q_6, 0, \mathrm{R})$
q_5	(q_5, B, R)	$(q_5, \mathrm{B}, \mathrm{R})$	$(q_6, \mathrm{B}, \mathrm{R})$
$*q_6$			

Soluzione: La TM calcola l'operazione di sottrazione propria tra i due interi m e n, cioè $m \dot{-} n = max(m-n, 0)$. La macchina parte con un nastro in cui è presente la stringa di input 0^m10^n, dove m e n sono rappresentati dalla loro codifica unaria (usando il simbolo zero) e divisi dal simbolo 1. La macchina si arresta con $0^{m \dot{-} n}$.

Esercizio 2. Definite una macchina di Turing *M* che accetta il linguaggio costituito dalle stringhe con diverso numero di 0 e 1. Definite la specifica formale della TM *M*, riportando δ come tabella di transizione.

Soluzione (una tra le possibili):

	0	1	*X*	*Y*	*B*
q_0	(q_1, X, R)	(q_2, X, R)		(q_0, Y, R)	
q_1	$(q_1, 0, \mathrm{R})$	(q_3, Y, L)		(q_1, Y, R)	(q_4, B, R)
q_2	(q_3, Y, L)	$(q_2, 1, \mathrm{R})$		(q_2, Y, R)	(q_4, B, R)
q_3	$(q_3, 0, \mathrm{L})$	$(q_3, 1, \mathrm{L})$	(q_0, X, R)	(q_3, Y, L)	
$*q_4$					

Esercizio 3. Scrivete le descrizioni istantanee della TM *M* definita nell'Esercizio 2 quando il nastro di input contiene: (a) 1100 (b) 01100 (c) 10110.

Soluzione:
(a) $q_01100 \vdash Xq_2100 \vdash X1q_200 \vdash Xq_31Y0 \vdash q_3X1Y0 \vdash Xq_01Y0 \vdash XXq_2Y0 \vdash XXYq_20 \vdash XXq_3YY \vdash Xq_3XYY \vdash XXq_0YY \vdash XXYq_0Y \vdash XXYYq_0B$ *si arresta senza accettare*
(b) $q_001100 \vdash Xq_11100 \vdash q_3XY100 \vdash Xq_0Y100 \vdash XYq_0100 \vdash XYXq_200 \vdash XYq_3XY0 \vdash XYXq_0Y0 \vdash XYXYq_00 \vdash XYXYXq_1B \vdash XYXYXBq_4B$ *accetta*

(c) $q_0 10110 \vdash X q_2 0110 \vdash q_3 XY110 \vdash X q_0 Y110 \vdash XY q_0 110 \vdash XYX q_2 10 \vdash XYX1 q_2 0 \vdash XYX q_3 1Y \vdash XY q_3 X1Y \vdash XYX q_0 1Y \vdash XYXX q_2 Y \vdash XYXXY q_2 B \vdash XYXXYB q_4 B$ accetta

Esercizio 4. Indicate quali fra le seguenti istanze di PCP hanno soluzione. Ognuna è presentata sotto forma delle due liste A e B; le i-esime stringhe delle due liste sono corrispondenti per i=1, 2, etc.
(a) A = (01, 100, 010); B = (010, 00, 100)
(b) A = (xy, x, yz, z); B = (yz, xy, zx, x)
(c) A = (10, 110, 01); B = (100, 01, 11)

Soluzione: (a) Si crea un loop. Ad esempio 1,3,3,... oppure 3,3,... (b) Si crea un loop. La sequenza di indici 2,3 permette di andare avanti con la ricerca di corrispondenza. Dopo possiamo proseguire con 1,4 (cioè la sequenza diverrà 2,3,1,4), ma questo genera due stringhe di lunghezza diversa (la seconda ha un simbolo 'x' in più). Nulla indica che la stringa non può essere estesa a soluzione, ma proseguiremo all'infinito con una sequenza del tipo 2,3,1,4,1,4,... (c) Non ha soluzione. La prima scelta obbligata è i=1; dopo possiamo scegliere solo 3; successivamente, ognuno dei tre indici non rende possibile proseguire.

Esercizio 5. (a) Qual è la relazione tra linguaggi ricorsivi e RE? (b) Definite i due linguaggi L_e e L_{ne} e dite a quale classe di linguaggi appartengono.

Soluzione: (a) L'insieme dei linguaggi ricorsivi è un sottoinsieme dei linguaggi RE. I linguaggi ricorsivi (o decidibili) sono quelli per cui ho garanzia che la TM si arresterà, sia che accetti o che non accetti l'input. I linguaggi RE non ricorsivi sono quelli per cui ho garanzia che la TM si arresterà se l'input è nel linguaggio, ma nel caso in cui non lo sia potrebbe proseguire in eterno.
(b) $L_e=\{M \mid L(M)=0\}$, non è RE. $L_{ne}=\{M \mid L(M) \neq 0\}$, è RE ma non è ricorsivo.

Esercizio 6. (a) Date la definizione delle classi di problemi P, NP e NP-completi. (b) Date la definizione del problema CSAT ed indicate a quale classe appartiene.

Soluzione: (a) Un problema è nella classe P se è risolvibile da una TM deterministica in tempo polinomiale. Un problema è nella classe NP se è risolvibile da una TM non-deterministica in tempo polinomiale. Un linguaggio L si dice NP-completo se è in NP, e se per ogni altro L' in NP esiste una riduzione polinomiale di L' a L.
(b) CSAT è il problema di soddisfacibilità così definito: data una espressione booleana in forma CNF, dire se è soddisfacibile. CSAT (come SAT) è NP-completo.

Esercizio 7. Dite quali tra le seguenti affermazioni è corretta:
(a) Un'espressione booleana si dice in forma CNF se è formata da una o più clausole tra di loro in AND logico.
(b) L'espressione $(x \wedge \neg y \wedge \neg z) \vee (\neg x \wedge \neg y \wedge z)$ è in 3-CNF.

(c) Il linguaggio di diagonalizzazione L_d è definito come l'insieme delle stringhe w_i tali che w_i è in $L(M_i)$ (seguendo la codifica definita a lezione).
(d) Se un linguaggio L e il suo complemento sono RE, allora anche L è ricorsivo.
(e) La trattazione dell'intrattabilità si basa sull'ipotesi (non dimostrata) che P$\neq$ NP.

Soluzione: (a) corretta; (b) non corretta (AND e OR sono invertiti); (c) non corretta (L_d è definito come l'insieme delle stringhe tali che w_i NON è in $L(M_i)$); (d) corretta; (e) corretta.