

שאלות חזרה – מבוא לבינה מלאכותית

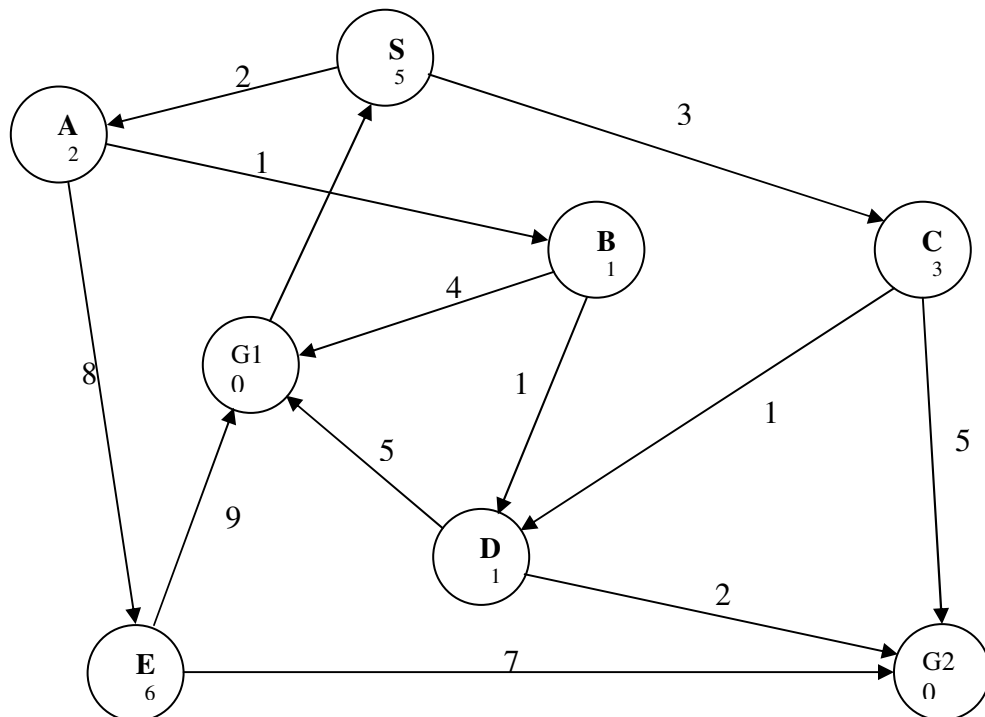
(תשס"ה, מועד ב)

נתון מרחב חיפוש כמצויר להלן. מצב ההתחלה מסומן באות S. מצבי הסיום הם G1, G2, G3. המספרים על הקשתות מגדירים את משקלי הקשתות והעלות המשווערת למטרה מכל קודקוד מוגדרת בתוך הקודקודים. לכל אחת מאלגוריתמי החיפוש בטבלה, כתבו לאיזה מקודקודי המטרה יגיע האלגוריתם (במידה והאלגוריתם אינו מוצא אף מטרה, סמנו F). כמו כן כתבו על פי הסדר את הקודקודים היוצאים מה-OPEN-LIST במהלך ריצת האלגוריתם.

הערות:

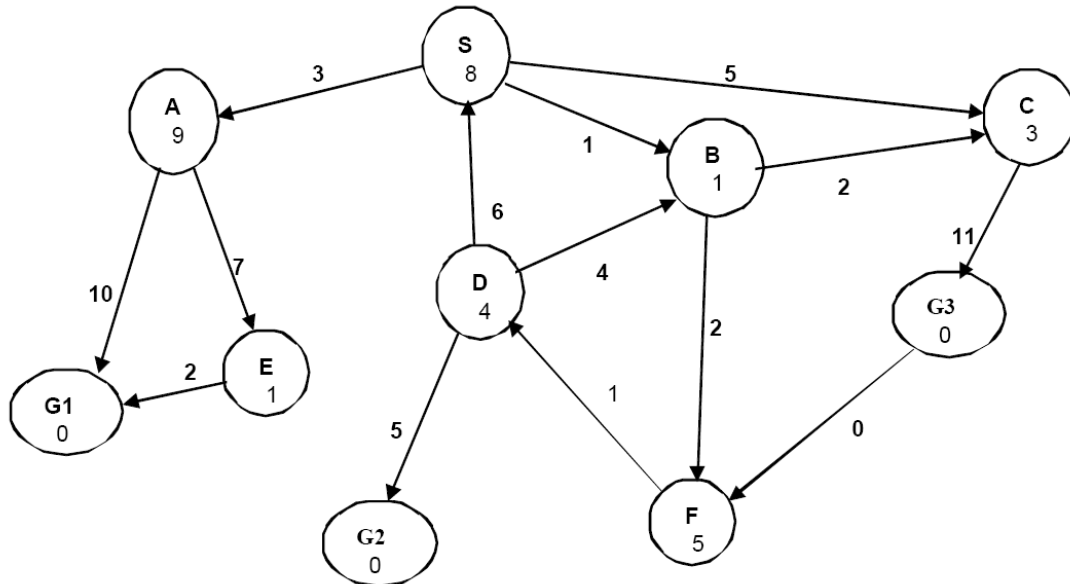
- א. הנח כי האלגוריתמים משתמשים ב-Closed-List.
- ב. במידה ולשני קודקודים או יותר יש עדיפות שווה, יש לבחור בסדר אלפביתי (A עדיף על B).
- ג. האלגוריתם עוצר כאשר הוא מגיע למטרה הראשונה.

האלגוריתם	מטרה	קודקודים שהוצאו מה-Open-List
BFS		S, A, C, B, D, E, G2
Iterative Deepening		S, A, C, S, A, B, E, C, D, G2
Greedy [Using h(n)]		S, A, B, G1
A*		S, A, B, D, C, G2
Uniform Cost Search		S, A, B, C, D, G2



#iter	Open-List	Out of Open list
BFS	S	
	A,C	S
	B,D,E,G2	A,C
	Goal found	B,D,E,G2
Iterative Deepening	S	
1	A,C	S
2	-	A,C
R2	S	
	A,C	S
	B,E,C	A
	C	B,E
	D,G2	C
	Goal found	D,G2
Greedy	S(5)	
	A(2),C(3)	S
	B(1),C(3),E(6)	A
	G1(0),D(1),C(3),E(6)	B
	Goal found	G1
A*	S(5)	
	A(4),C(6)	S
	B(4),C(6),E(16)	A
	D(5),C(6),G1(7),E(16)	B
	G2(6),G1(7),C(6),E(16)	D
	G2(6),G1(7),E(16)	C
	Goal found	G2
Uniform Cost Search	S(0)	
	A(2),C(3)	S
	B(3),C(3),E(10)	A
	C(3),D(4),G1(7),E(10)	B
	D(4),G1(7),G2(8),E(10)	C
	G2(6),G1(7),E(10)	D
	Goal found	G2

IDA* example
נתון מרחב החיפוש הבא:



קודקוד ההתחלה מסומן באות S. G1, G2, G3 הם קודקודים המקיימים את תנאי המטרה. המספרים על הקשתות מגדירים את משקלי הקשתות והעלות המשוערת למטרה מכל קודקוד מוגדרת בתוך הקודקודים. לכל אחת מאלגוריתמי החיפוש הבאים, כתבו לאיזה מקודקודי המטרה יגיע האלגוריתם (במידה והאלגוריתם אינו מוצא אף מטרה, סמנו F). כמו כן כתבו על פי הסדר את הקודקודים היוצאים מה- open-list במהלך ריצת האלגוריתם.

Iterative Deepening A* (IDA*) (i)

הערות:

- IDA* לא מכניס ל- open-list קודקודים שה- f שלהם גבוה יותר מערך ה- threshold הנוכחי.
- במידה ולשני קודקודים או יותר יש עדיפות שווה לפי האלגוריתם, יש לבחור בסדר אלפביתי (A) עדיף על (B).
- יש להשתמש ב- Duplicate pruning.
- האלגוריתם עוצר כאשר הוא מגיע למטרה הראשונה.

תשובה:

האלגוריתם	מטרה	קודקודים שהוצאו מה-Open-List
IDA*	G2	S, B, F, D, C, S, B, F, D, G2

הסבר:

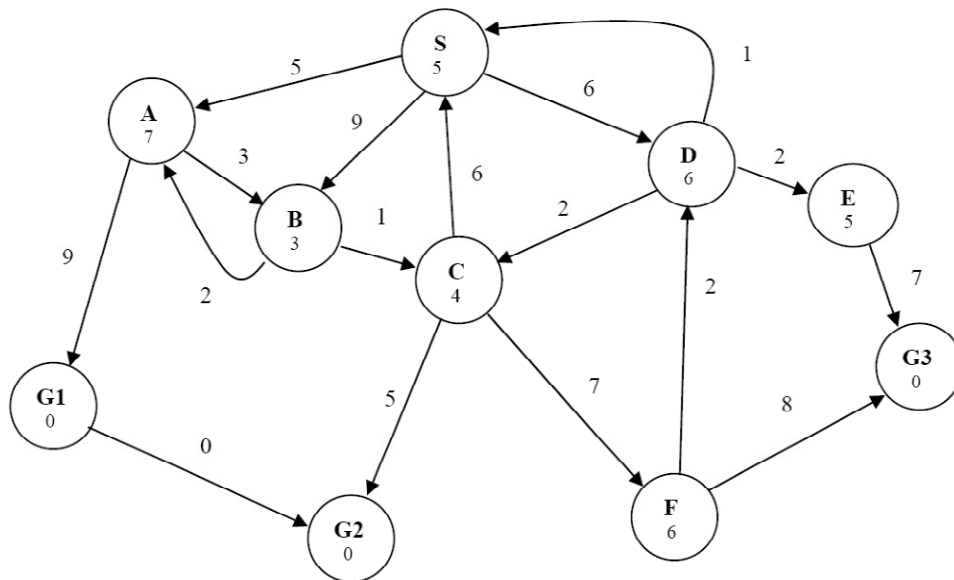
#iter	Open-List	Out of Open list
IDA*	S(8) , threshold set to 8	
	B(2), C(8)	S
	F(8), C(8)	B
	D(8), C(8)	F
	C(8)	D
	Empty - end of first iteration	C

2 nd iteration	S(8), threshold set to 9	
	B(2),C(8)	S
	F(8),C(8)	B
	D(8),C(8)	F
	G2(9),C(8)	D
	Goal found	G2

(תשס"ח מועד א')

שאלה 1:

א) נתון מרחב החיפוש בתמונה, כש-S הוא קודקוד ההתחלה ו-G1, G2, G3 הם קודקודי מטרה. הצלעות מסומנות במחיר המעבר עליהם והערכת המחיר למטרה נתונה בתוך הקודקודים.



לכל אחת משיטות החיפוש הבאות כתוב מה המטרה אליה יגיע (אם בכלל) וציין את כל הקודקודים שמוצאים מה- OPEN List. יש להשתמש ב- CLOSED List, למעט עבור אלגוריתם (ii).

כאשר כל הפרמטרים זהים, ההכנסה ל- OPEN List נעשית לפי סדר אלפא-בתי. כמו כן בעת מיון, קודקודים בעלי ערך זהה יסודרו בסדר אלפא בתי.

Breadth First (ii)

Iterative Deepening (iii)

A* (iv)

Greedy search (v)

לכל אלגוריתמים כתוב:

קודקוד המטרה שאליו יגיע:

הקודקודים שהוצעו מ-Open:

קודקודים שהוצעו מ-Open-List	מטרה	האלגוריתם
S,A,B,D,B,G1	G1	BFS
S S,A,B,D S,A,B,G1	G1	Iterative Deepening
S,A,B,B,D,C,S,C,E,G2	G2	A*
S,B,C,G2	G2	Greedy [Using h(n)]

הערה: אין להוריד נקודות למי שכתב את סדר הקודקודים עבור A* כ: S,A,B,B,D,C,C,E,G2

הסבר:

#iter	Open-List	Out of Open list
BFS	S	
	A,B,D	S
	B,G1,A,C,C,E,S	A,B,D
	Goal found	B,G1
A*	S(5)	
	A(12),B(12),D(12)	S
	B(11),B(12),D(12),G1(14)	A
	B(12),D(12),C(13),G1(14),A(17)	B
	C(12),S(12),C(13),E(13),G1(14),A(17)	B,D
	S(12),C(13),E(13),G2(13),G1(14),F(15),A(17),S(19)	C
	G2(13),G1(14),F(15),G3(15),A(17),S(19)	S,C,E
	Goal found	G2
Greedy	S(5)	
	B(3),D(6),A(7)	S
	C(4),D(6),A(7),A(7)	B
	G2(0),S(5),D(6),F(6),A(7),A(7)	C
	Goal found	G2

השתמש באותו מרחב חיפוש כמו ב-Simulated Annealing.

הנח כי הטמפרטורה היא 10 ו- $c=1$.

(i) אם אתה בקודקוד S ו-Simulated Annealing בחר בצורה רנדומלית את קודקוד B, מה

ההסתברות שהקודקוד הבא בחיפוש יהיה B? השתמש בפונקציה h הנתונה בקודקוד.

תשובה: 100% מכיון ש- $h(B) < h(S)$

(ii) אם אתה בקודקוד S ו-Simulated Annealing בחר בצורה רנדומלית בקודקוד A, מה

ההסתברות שהקודקוד הבא בחיפוש יהיה A?

כמו ב-(i) השתמש בפונקציה h שנתונה בקודקוד.

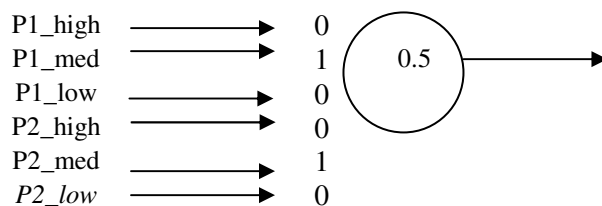
תשובה: $e^{-\frac{2}{10}}$

למידה

(תשס"ה, מועד א')

חוקר בודק את הקשר בין מינונים שונים של שני חומרים בתרופה חדשה על קבוצות מחקרים. נתונים החומרים p1 ו-p2. החוקר ניסה לשלב מיון גבוה, נמוך ובינוני של כל אחד מהחומרים על קבוצות שונות של נבדקים וסיווג את התוצאות לטובות ורעות. ציירו פרספטרון העושה שימוש בפונקציית מדרגה (step) המייצג את המידע בטבלה.

תוצאה	מינון p1	מינון p2
רעה	גבוה	גבוה
טובה	גבוה	בינוני
רעה	גבוה	נמוך
טובה	בינוני	נמוך
טובה	בינוני	בינוני



- ענו נכון / לא נכון, ונמקו בקצרה:
- ניתן לייצג כל פונקציה בוליאנית עם k משתנים (attributes) ע"י עצי החלטה עם k קודקודים פנימיים (כולל השורש).
 - אם לא פצלנו את עץ ההחלטה עפ"י התכונה המפרידה ביותר בכל שלב, נקבל פתרון שלא יתאים לכל הדוגמאות של הקלט.
 - במידה והייתה שגיאה בדוגמאות, גם עץ ההחלטה שיתקבל יהיה שגוי.

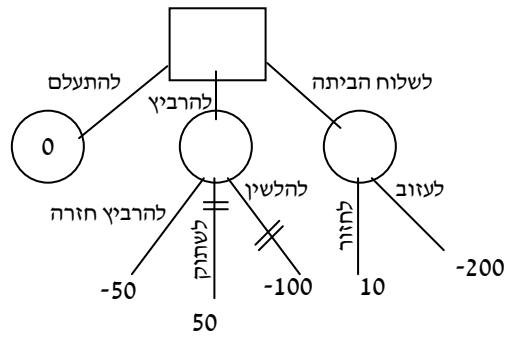
- הטענה אינה נכונה – יש צורך בהרבה יותר קודקודים פנימיים עבור האפשרויות של המשתנים.
- הטענה אינה נכונה – העץ שיתקבל עלול להיות יותר ארוך אך לא מנוגד לדוגמאות בהכרח.
- הטענה אינה נכונה – השגיאות בדוגמאות עשויות לא לבוא לידי ביטוי בעץ (ייתכן שהם בתכונה לא רלוונטית שלא מופיעה בכלל בעץ).

משחקים

(תשס"ג, מועד א)

- הרשלה התחצף אל המורה שלו. המורה נמצא בדילמה: הוא יכול להתעלם מהרשלה, להרביץ לו, לשלוח אותו הביתה לשבוע או להעיף אותו מבית ספר.
- אם המורה ירביץ להרשלה, הרשלה יכול להרביץ בחזרה, לשתוק או להלשין עליו למנהל.
- אם המורה ישלח את הרשלה הביתה לשבוע, הרשלה יכול לחזור אחרי שבוע או לעזוב את ביה"ס.
- אם המורה מתעלם מהרשלה אזי הרווח של שניהם הוא 0.
- אם המורה ירביץ להרשלה והרשלה ירביץ בחזרה, אזי המורה יפסיד 50 והרשלה ירוויח 50.
- אם המורה ירביץ להרשלה והוא ישתוק, אזי המורה ירוויח 50 והרשלה יפסיד 50.
- אם המורה ירביץ להרשלה והוא ילשין למנהל, אזי המורה יפסיד 100 והרשלה ירוויח 100.
- אם המורה ישלח את הרשלה הביתה לשבוע והוא חוזר אחרי שבוע אזי המורה ירוויח 10 והרשלה יפסיד 10.
- אם הרשלה עוזב את בית הספר אזי הוא מרוויח 200 והמורה מפסיד 200 (כשל חינוכי).

- שרטטו את עץ המשחק המתאים לבעיה שתוארה, כאשר הרשלה הוא minimizer והמורה הוא maximizer player.
- הפעילו את אלגוריתם α - β על העץ (משמאל לימין) ומצאו מהי התגובה האופטימלית של המורה להתנהגותו של הרשלה.



התגובה האופטימאלית של המורה היא להתעלם.

מצא את ה- pure and mixed Nash Equilibrium

	DL	CC
DL	50,50	80,20
CC	90,10	20,80

תשובה :

אין Pure. Mixed : נסמן ב- P את ההסתברות ששחקן השורות משחק DL. נקבל :

$$U(DL) = U(CC)$$

$$50p + 10(1-p) = 20p + 80(1-p)$$

$$P = 0.7$$

נסמן ב- Q את ההסתברות ששחקן העמודות משחק DL. נקבל :

$$U(DL) = U(CC)$$

$$50q + 80(1-q) = 90q + 20(1-q)$$

$$q = 0.6$$

שאלות חזרה - תכנון

1. נתונה הפעולה הבאה של STRIPS:

Operation: copy(X,Y)

Preconditions: in(X,V1), in(Y,V2)

Delete list: in(Y,V2)

Add list: in(Y,V1)

נתון כי מצב העולם הוא in(A,0), in(B,1), in(C,-1)

מה תהיה התוכנית ש-STRIPS ייצר כאשר ה-goal הוא $in(A,1) \wedge in(B,0)$ באמצעות הפעולה .copy

פתרון:

לא ניתן לבנות תוכנית STRIPS לבעיה זו. נדגים זאת להלן (המחסנית גדלה כלפי מטה):

המחסנית	הסבר
$in(A,1) \wedge in(B,0)$	
$in(A,1)$	פירוק ה-goal המורכב
$\neg in(B,0)$	
$copy(X,B)$	החלפת ה-goal הפשוט באופרטור
$in(X,0), in(B,V2)$	ודחיפת ה-preconditions
$in(B,V2)$	פירוק ה-goal המורכב
$in(X,0)$	

בשלב זה X מקבל את הערך A ואנחנו מעתיקים ל-B את הערך ב-A וכך דורסים אותו, ולכן $in(A,1)$ לא יתקיים. תופעה דומה קורה אם מחליפים את סדר תתי המטרה.

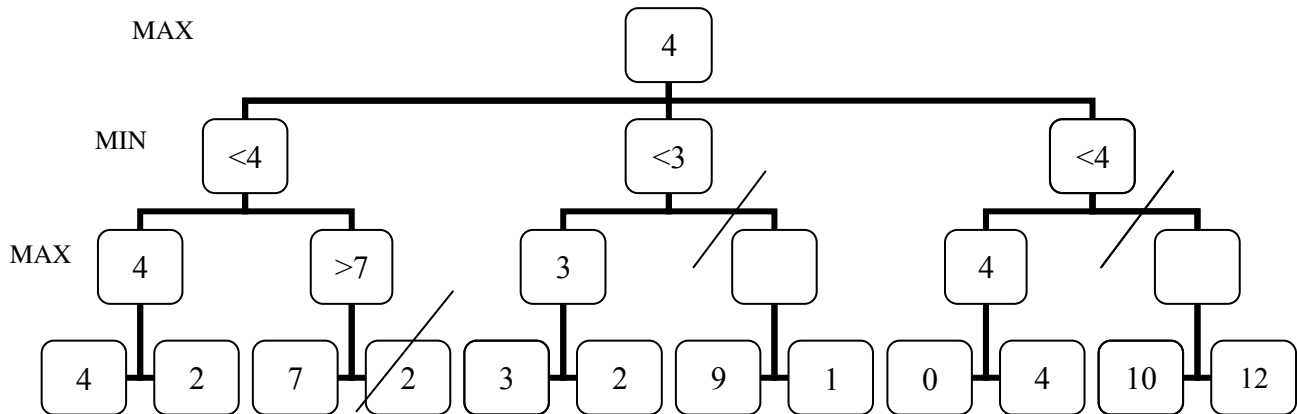
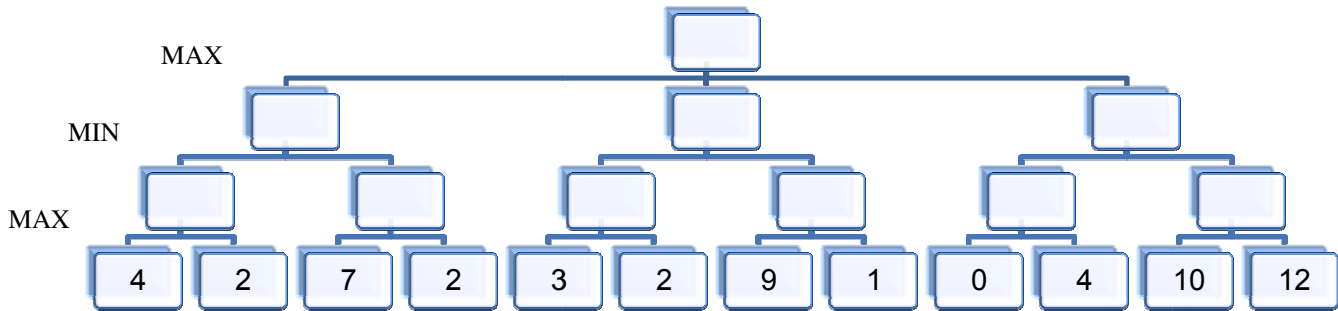
2. ענו נכון/לא נכון ונמקו:

- a. ניתן לפתור בעזרת STRIPS כל בעיית תכנון
- b. לפניכם מתג שאם לוחצים עליו כשהאור דולק, הוא נכבה, ואם לוחצים עליו כשהאור כבוי, הוא נדלק. פעולת הלחיצה נקראת Toggle. לא ניתן לתת תיאור STRIPS לפעולה Toggle בשימוש בפרדיקטים $On(x)$ ו- $Off(x)$ בלבד.

- a. הטענה אינה נכונה – היות והאלגוריתם אינו שלם אפילו אם יש פתרון לבעיית תכנון מסוימת האלגוריתם עלול לא למצוא אותו.
- b. הטענה נכונה – אכן לא ניתן לתת תיאור STRIPS לבעיה עם הפרדיקטים הנתונים בלבד, היות וגם ה-Preconditions וגם ה-effect מותנים.

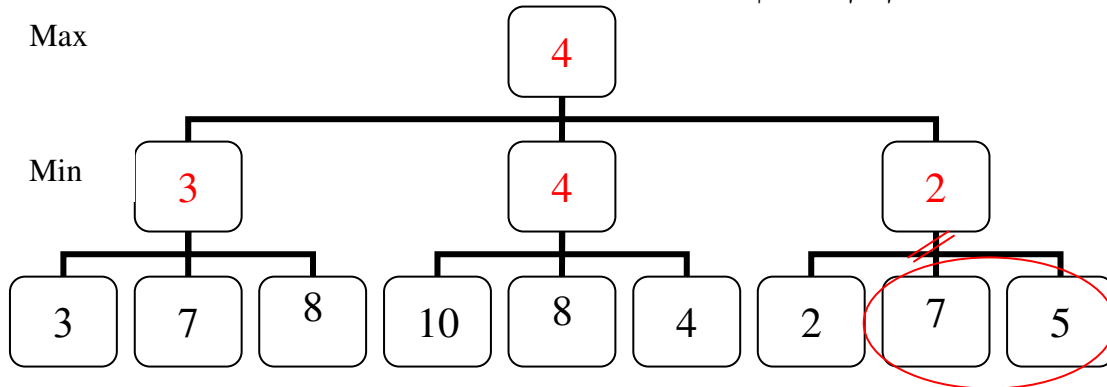
שאלות חזרה – עצי משחק

1. (תשס"ג א') נתון עץ המשחק הבא (רמה ראשונה היא MAX). מה תהיה התזוזה הראשונה של השחקן לפי $\alpha - \beta$? אילו קודקודים לא יפותחו?
פתרון:

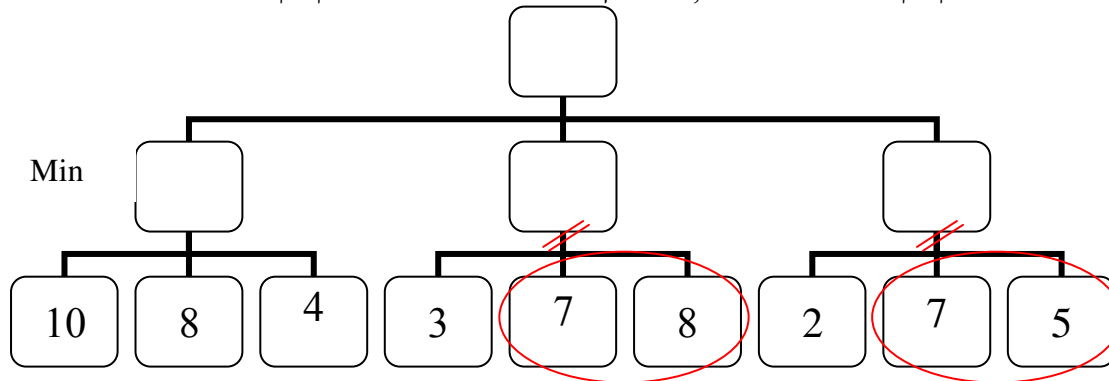


השחקן יזוז לענף הראשון.

1. מועד א' 2009: נתון עץ המשחק הבא:



- (א) השלם את 4 הערכים של הקודקודים הריקים לפי אלגוריתם ה-Minimax
 (ב) סמן את הקודקודים שנגזמו (=לא יפותחו) ע"י אלגוריתם Alpha-beta. הנה שסדר פיתוח הקודקודים הוא משמאל לימין.
 (ג) סדר מחדש את הקודקודים כך ש-Alpha-beta יגזום כמות מקסימאלית של קודקודים. ניתן לשנות את הסדר של הקודקודים ברמות 1 ו-2, אבל אין לשנות את האבא של כל קודקוד.



מועד ב' 2007: שאלה 4

נתון עץ המשחק להלן. הנה כי כעת תור המחשב (maximizer).

א. לאיזה מצב יבחר המחשב להגיע בתור זה? הענף השמאלי

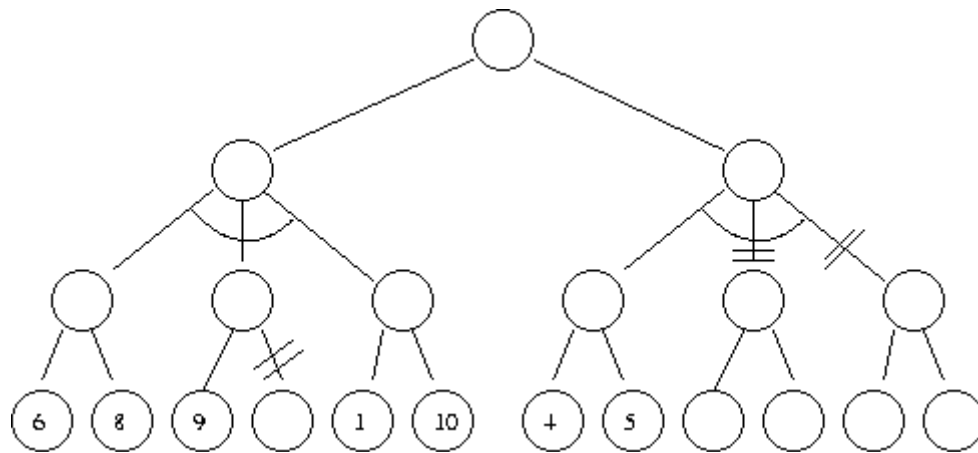
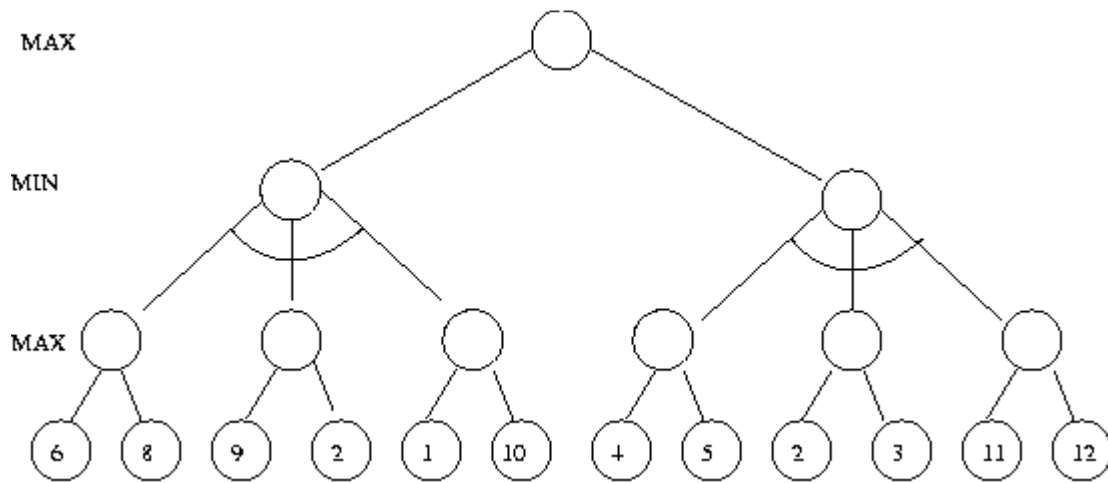
ב. מה יהיו הערכים בקודקודים בסוף ריצת אלגוריתם min-max?

עומק 2: 8,9,10, 5,3,12

עומק 1: 8,3

עומק 0: 8

ג. הראו את השפעת הגיזום של אלגוריתם alpha-beta בהנחה שהעץ נסרק משמאל לימין.



שאלות חזרה - עצי החלטה

2. (תשס"ג ב') נתונה סוכנות הכרויות עם database גדול של אנשים. על מנת לספק שירות יעיל אתם מעוניינים לבנות מודל לכל לקוח חדש בהתבסס על קבוצה מייצגת קטנה, אותה הלקוח כבר קטלג (ה-training set). ברגע שהמודל נלמד ניתן לסנן עבורו את ה-database ולשלוח רק את המועמדים הרלוונטיים.
- נתונה קבוצת הדוגמאות הבאות ומטרתכם ללמוד את הפונקציה Candidate המחזירה yes / no, המציין את פוטנציאל המועמד ללקוח.

<i>i</i>	<i>Sex</i>	<i>Status</i>	<i>Age</i>	<i>Education</i>	<i>Location</i>	<i>Looks</i>	<i>Candidate</i>
1	M	Single	<30	HS	North	Good	No
2	F	Single	30-40	BA	Center	Good	Yes
3	F	Single	>40	MA	South	Great	No
4	M	Divorced	30-40	MA	South	Great	No
5	F	Single	<30	MA	North	Great	No
6	F	Single	30-40	HS	Center	Good	No
7	F	Single	30-40	BA	South	Good	Yes
8	F	Single	>40	HS	Center	Great	No
9	F	Divorced	30-40	MA	South	Great	Yes
10	F	Divorced	<30	HS	South	Good	No
11	F	Single	>40	BA	Center	Good	No
12	F	Divorced	<30	MA	Center	Great	No

ציירו עץ החלטה לקבוצת הדוגמאות. אם קיימות מספר תכונות עם אותו gain יש לבחור לפי סדר הא"ב. יש לציין בקודקודי העץ את ערכי ה-gain הרלוונטיים.

פתרון:

$$I(3/12,9/12) = -(1/4)\log(1/4) - (3/4)\log(3/4) = 0.5 + 0.311 = 0.811$$

$$\begin{aligned} \text{Remainder}(\text{Age}) &= (4/12) * I(0/4,4/4) + (5/12) * I(3/5,2/5) + (3/12) * I(0/3,3/3) \\ &= 0 + (5/12) * [-0.6\log(0.6) - 0.4\log(0.4)] + 0 = 0.405 \end{aligned}$$

$$IG(\text{Age}) = I(3/12,9/12) - \text{Remainder}(\text{Age}) = 0.811 - 0.405 = 0.41$$

בצורה דומה נבדוק את שאר המאפיינים. המאפיין AGE נבחר כמאפיין בעל ה-IG הגדול ביותר.

נחלק את טבלת הדוגמאות על פי ערכי AGE האפשריים. אם גיל המתמודד קטן מ-30 או גדול מ-40 לכל הדוגמאות יש את אותו הערך ולכן נגדיר אותם כעלים. עתה נותרנו עם כל הדוגמאות בהן גיל המתמודד הוא בין 30 ל-40:

<i>I</i>	<i>Sex</i>	<i>Status</i>	<i>Age</i>	<i>Education</i>	<i>Location</i>	<i>Looks</i>	<i>Candidate</i>
2	F	Single	30-40	BA	Center	Good	Yes
4	M	Divorced	30-40	MA	South	Great	No
6	F	Single	30-40	HS	Center	Good	No
7	F	Single	30-40	BA	South	Good	Yes
9	F	Divorced	30-40	MA	South	Great	Yes

על מנת לבחור את המאפיין לתת העץ נחשב עבור כל מאפיין את IG:

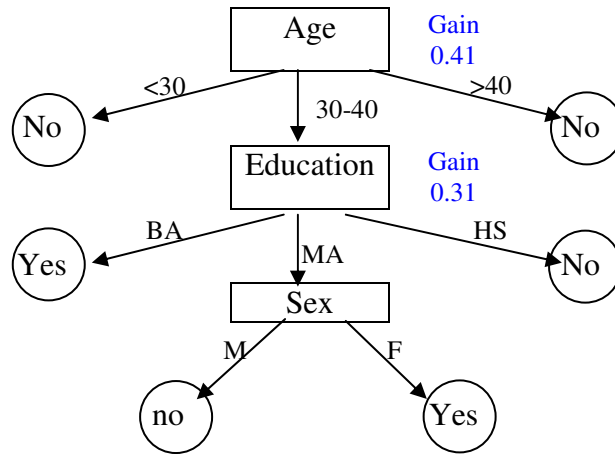
$$IG(\text{Education}) = I(3/5,2/5) - \text{Remainder}(\text{Education})$$

$$I(3/5,2/5) = -0.6\log(0.6) - 0.4\log(0.4) = 0.442 + 0.529 = 0.971$$

$$\text{Remainder}(\text{Education}) = (1/5)I(0/1, 1/1) + (2/5)I(2/2, 0/2) + (2/5)I(1/2, 1/2) = 0 + 0 + 0.4 = 0.4$$

$$\text{IG}(\text{Education}) = 0.971 - 0.4 = 0.571$$

העץ המתקבל הוא :



3. (תשס"א א') נתונות הדוגמאות הבאות לעץ החלטה בנושא של האם להצביע בבחירות או לא. הפרמטרים הם מזג אוויר, מידת תמיכה, וסקרים.

החלטה	סקרים	מידת תמיכה	מזג אוויר
לא	תיקו	מועטה	סוער
כן	שלי מוביל	גדולה	בהיר
כן	שלי מוביל	גדולה	מעונן
לא	יריב מוביל	גדולה	סוער
כן	תיקו	מועטה	מעונן
לא	תיקו	גדולה	סוער
לא	שלי מוביל	מועטה	בהיר
כן	יריב מוביל	מועטה	בהיר
לא	שלי מוביל	מועטה	מעונן
כן	יריב מוביל	מועטה	מעונן

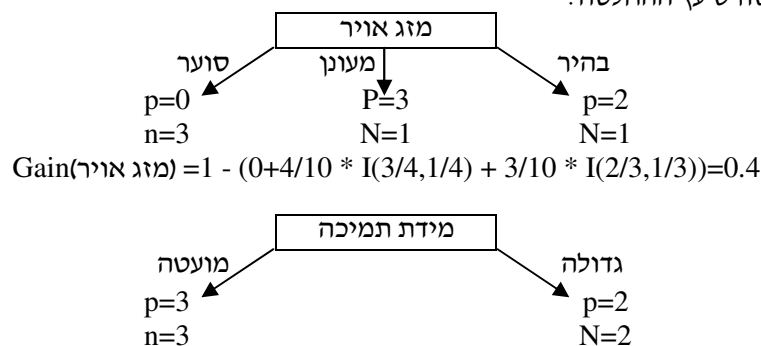
- בנו עץ החלטה מתאים, במקרה שה-gain שווה בחרו עפ"י סדר הא"ב.
- ציינו בקודקודי העץ הרלוונטיים את ה-gain של התכונה שבחרתם.
- תארו נוסחה לוגית שיכולה לתאר את העץ שקבלתם.

פתרון :

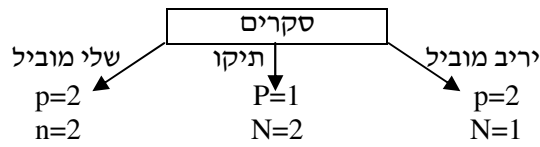
א.ב. בניית עץ החלטה לשאלה האם להצביע בבחירות :
נחשב את האנטרופיה ההתחלתית –

$$p=5, n=5 \rightarrow I(1/2, 1/2) = 1$$

נבחר תכונה לשורש עץ ההחלטה :

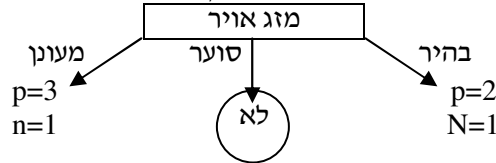


$$\text{Gain(מידת תמיכה)}=0$$



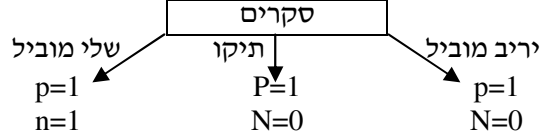
$$\text{Gain(סקרים)} = 1 - (4/10 * I(1/2, 1/2) + 3/10 * I(1/3, 2/3) + 3/10 * I(2/3, 1/3)) = 0.05$$

ה-Gain הכי גדול הוא של התכונה "מזג אוויר" ולכן היא תהיה שורש העץ.

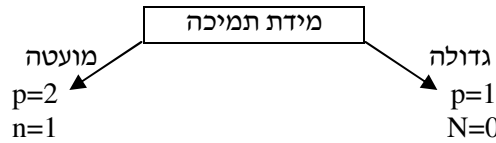


נבחר תכונה עבור תת העץ השמאלי בו מזג האוויר הוא מעונן :

$$I(3/4, 1/4) = 0.81 \text{ היא עץ זה}$$

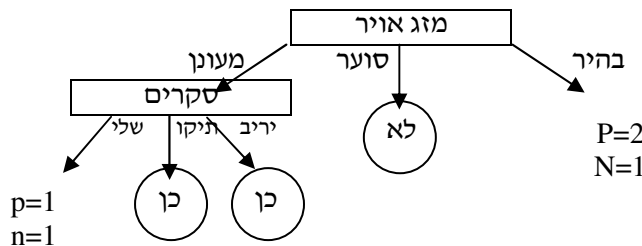


$$\text{Gain(סקרים)} = 0.81 - (2/4 * I(1/2, 1/2) + 0) = 0.31$$

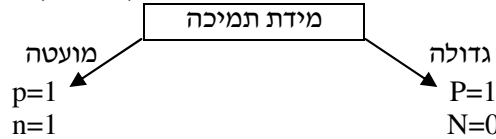


$$\text{Gain(מידת תמיכה)} = 0.81 - (3/4 * I(2/3, 1/3) + 0) = 0.81 - 0.69 = 0.12$$

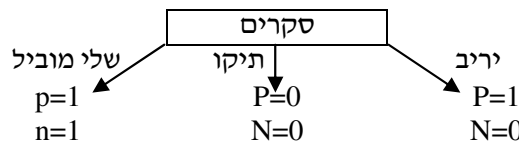
שורש תת העץ יהיה אם כן "סקרים"



נבחר תכונה לתת העץ הימני. האנטרופיה ההתחלתית היא $I(1/3, 2/3) = 0.918$

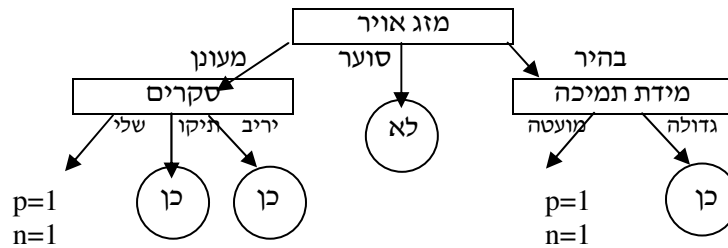


$$\text{Gain(מידת תמיכה)} = 0.918 - (2/3 + 0) = 0.251$$

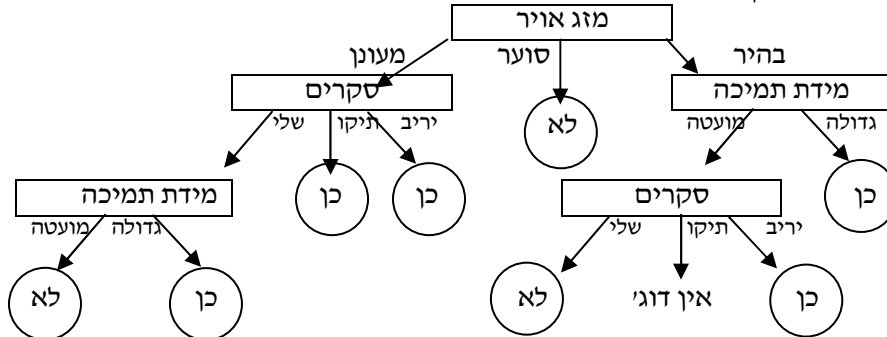


$$\text{Gain(סקרים)} = 0.918 - (2/3 + 0) = 0.251$$

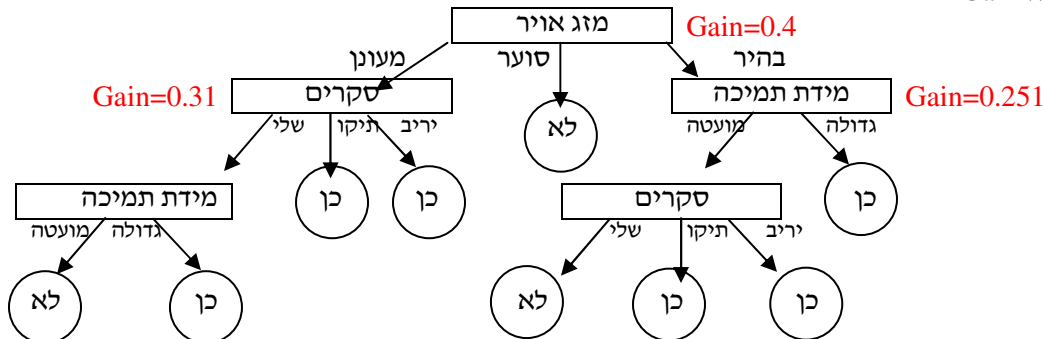
שורש תת העץ יהיה אם כן "מידת תמיכה", שכן במקרה שה-Gain שווה יש לבחור לפי סדר הא"ב. עד כה העץ הוא:



נציב את התכונות הנותרות בתתי העץ:



היות ואין דוגמאות עבור סקרים=תיקו ולא ניתן לנו ערך ברירת מחדל, נבחר עפ"י הרוב "כן", והעץ המתקבל עם ערכי ה-Gain:



ג. נסמן ב- W את מזג האוויר, T את מידת התמיכה ו- S את הסקרים. הנוסחא שמתארת את העץ היא:

$$((W = \text{בהיר}) \wedge (T = \text{גדולה})) \vee ((W = \text{בהיר}) \wedge (T = \text{מועטה}) \wedge \sim(S = \text{שלי})) \vee ((W = \text{מערונן}) \wedge \sim(S = \text{שלי})) \vee ((W = \text{מערונן}) \wedge (T = \text{גדולה}))$$

4. ענו נכון / לא נכון, ונמקו בקצרה:

א. ניתן לייצג כל פונקציה בוליאנית עם k משתנים ($attributes$) ע"י עצי החלטה עם k קודקודים פנימיים (כולל השורש).
 הטענה אינה נכונה – יש צורך בהרבה יותר קודקודים פנימיים עבור האפשרויות של המשתנים.

ב. אם לא פצלנו את עץ ההחלטה עפ"י התכונה המפרידה ביותר בכל שלב, נקבל פתרון שלא יתאים לכל הדוגמאות של הקלט.
 הטענה אינה נכונה – העץ שיתקבל עלול להיות יותר ארוך אך לא מנוגד לדוגמאות בהכרח.

ג. במידה והייתה שגיאה בדוגמאות, גם עץ ההחלטה שיתקבל יהיה שגוי.

הטענה אינה נכונה - השגיאות בדוגמאות עשויות לא לבוא לידי ביטוי בעץ (ייתכן שהם בתכונה לא רלוונטית שלא מופיעה בכלל בעץ).

ד. (תשס"ב ב') נתונות קבוצת דוגמאות עם התכונות A, B, C ו-D. בנו מהקבוצה Decision Tree כאשר קודקוד השורש הוא לפי A ולו 3 תתי עצים. יצרו עץ חדש באופן הבא: שורש העץ החדש זהה לקודם ואף לו 3 תתי עצים, אך כל תת עץ נבנה ע"י שליחת כל הדוגמאות בקבוצה, תוך התעלמות מ-A. האם נוצר אותו עץ כמו במקור? אם כן, נמקו, אם לא הביאו דוגמא נגדית.
פתרון: לא נוצר בהכרח אותו עץ. בעץ החדש אי אפשר להשתמש במאפיין A ולכן יוצרו שלושה תתי עצים זהים. עץ זה יוצר אם למאפיין A אין משמעות כלומר $IG(A)=0$ (ויש מספיק דוגמאות).

5. מועד א 2005: שאלה 2

נתונה טבלה ובה נתוני אימון. מוגדרים שלושה מאפיינים F1, F2 ו-F3. לכל מאפיין ישנם שלושה ערכים אפשריים: a, b או c.

א. עליכם לבנות עץ החלטה המייצג את מאפיין הפלט output על סמך האלגוריתם שנלמד בשיעור. אם קיימות מספר תכונות עם אותו f1 gain עדיף על f2 ו-f2 עדיף על f3. ערך ה-default כאשר אין דוגמא הוא false.

ב. ציינו בקודקודי העץ הרלוונטיים את ה-gain של התכונה שבחרתם.

ג. כתבו את הנוסחה הלוגית המתארת את העץ שקבלתם.

טבלת הנתונים:

#	f1	f2	f3	output
1	b	b	a	t
2	a	a	a	f
3	b	a	a	f
4	b	c	a	t
5	c	b	a	f
6	a	c	b	f
7	b	a	c	t
8	a	a	c	f
9	c	b	c	f

$$I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) = -\frac{p}{p+n} \log_2 \frac{p}{p+n} - \frac{n}{p+n} \log_2 \frac{n}{p+n}$$

$$\text{remainder}(A) = \sum_{i=1}^v \frac{p_i + n_i}{p+n} I\left(\frac{p_i}{p_i + n_i}, \frac{n_i}{p_i + n_i}\right)$$

$$IG(A) = I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) - \text{remainder}(A)$$

$$I(3/9,6/9)=0.92$$

$$\text{Reminder}(f1)=3/9*0+4/9*I(3/4,1/4)+2/9*0=4/9*0.81=\underline{\underline{0.36}}$$

$$\text{Reminder}(f2)=4/9*I(1/4,3/4)+3/9*I(1/3,2/3)+2/9*I(1/2,1/2)=4/9*0.81+3/9*0.92+2/9*1=\underline{\underline{0.889}}$$

$$\text{Reminder}(f3)=5/9*I(2/5,3/5)+1/9*I(0,1)+3/9*I(1/3,2/3)=5/9*0.97+0+3/9*0.92=\underline{\underline{0.845}}$$

$$\mathbf{IG(f1)}=0.92-0.36=\underline{\underline{0.56}}$$

#	f1	f2	f3	output
1	B	b	a	t
3	B	a	a	f
4	B	c	a	t
7	B	a	c	t

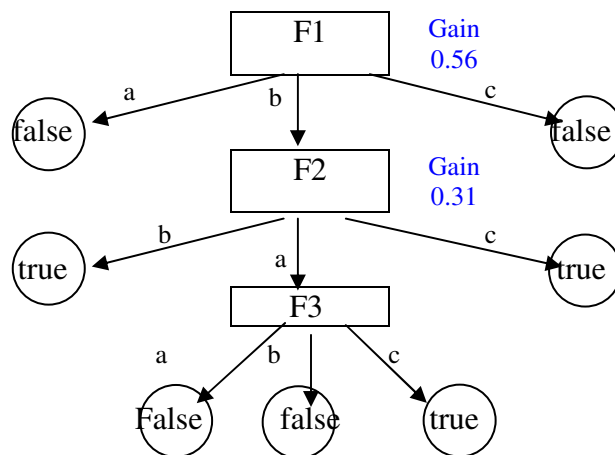
$$I(3/4,1/4)=0.81$$

$$\text{Remainder}(f2)=2/4*I(1/2,1/2)+1/4*I(1,0)+1/4*I(1,0)=0.5+0+0=\underline{\underline{0.5}}$$

$$\text{Remainder}(f3)=3/4*I(2/3,1/3)+1/4*I(1,0)=3/4*0.92=\underline{\underline{0.69}}$$

$$\mathbf{IG(f2)}=0.81-0.5=\underline{\underline{0.31}}$$

#	f1	f2	f3	output
3	B	a	a	f
7	B	a	c	t



Boolean Function:

$$\mathbf{f1=b^{\wedge}\{(f2=b)\vee(f2=c)\vee[(f2=a)^{\wedge}(f3=c)]\}}$$

שאלות חזרה - רשתות נוירונים

1. ציירו perceptron שבו 5 קלטים שערכם 0 או 1, המחשב את פונקצית הרוב (majority).

פתרון:

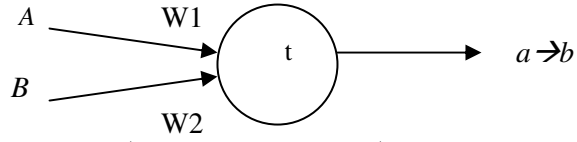
הרשת אמורה להוציא "1" כאשר לפחות 3 מהקלטים הם "1" ולכן משקולות יחידות הקלט הן 1 וביחידת הפלט נשתמש בפונקצית step עם threshold 2.5 (זו תשובה אפשרית מיני רבות).

2. ציירו רשת נוירונים המייצגת את $a \rightarrow b$.

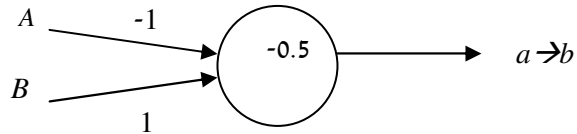
פתרון:

כדוגמאות ניקח את טבלת האמת של הביטוי:

a	B	$a \rightarrow b$
0	0	1
0	1	1
1	0	0
1	1	1



יש למצוא את המשקולות וערך הסף כך שהפלט יקיים את הדרוש. ניתן לתת ערכים רנדומליים למשקולות ולהפעיל את אלגוריתם הלמידה, אך ניתן גם למצוא את הערכים המתאימים בשיטת הניסוי והטעייה:



ניתן לראות את הבעיה כמערכת משוואות בת 3 נעלמים:

$$w_0 > 0$$

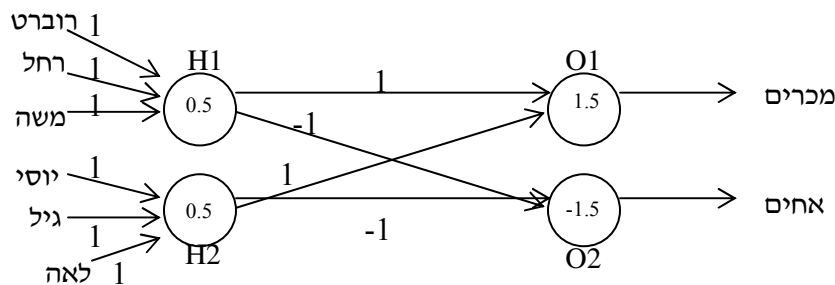
$$w_2 + w_0 > 0$$

$$w_1 + w_0 < 0$$

$$w_1 + w_2 + w_0 > 0$$

נבחר $w_0 = 0.5$. על מנת שמשוואה 3 תהיה נכונה, נבחר $w_1 = -1$. על מנת שמשוואה 4 תהיה נכונה נבחר $w_2 = 1$. שימו לב כי w_0 מוגדר להיות המספר $\text{thresh} * (-1)$.

3. (תשס"א ב') נתונה רשת הנוירונים הבאה:



בקודקודים מחושבת פונקצית $\text{Step}(t)$ כאשר t הוא המספר בקודקוד. בהינתן זוג של שמות (1-ים במקום המתאים בקלט), $O1$ מחזירה 0 ו- $O2$ מחזירה 1 אם הם אחים, ולהפך אם הם מכרים.

- א. האם רחל ומשה אחים?
- ב. האם גיל ולאה אחים?
- ג. האם גיל ורחל מכרים?
- ד. מהי הכלל על פיו פועלת הרשת?

פתרון:

א. נחשב את ערכי הפלט ביחידות:

$$H_1 = \text{Step}(\text{Robert} \cdot 1 + \text{Rachel} \cdot 1 + \text{Moshe} \cdot 1) = \text{Step}(0 \cdot 1 + 1 \cdot 1 + 1 \cdot 1) = \text{Step}(2) = 1 \quad (2 > 0.5)$$

$$H_2 = \text{Step}(\text{Yosi} \cdot 1 + \text{Gil} \cdot 1 + \text{Lea} \cdot 1) = \text{Step}(0 \cdot 1 + 0 \cdot 1 + 0 \cdot 1) = \text{Step}(0) = 0 \quad (0 \not> 0.5)$$

$$O_1 = \text{Step}(H_1 \cdot 1 + H_2 \cdot 1) = \text{Step}(1 \cdot 1 + 0 \cdot 1) = \text{Step}(1) = 0 \quad (1 \not> 1.5)$$

$$O_2 = \text{Step}(H_1 \cdot (-1) + H_2 \cdot (-1)) = \text{Step}(-1 \cdot 1 - 0 \cdot 1) = \text{Step}(-1) = 1 \quad (-1 > -1.5)$$

← רחל ומשה אחים.

ב. באופן זהה נקבל שגיל ולאה אחים.

ג. נחשב את ערכי הפלט ביחידות:

$$H_1 = \text{Step}(\text{Robert} \cdot 1 + \text{Rachel} \cdot 1 + \text{Moshe} \cdot 1) = \text{Step}(0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1) = \text{Step}(1) = 1 \quad (1 > 0.5)$$

$$H_2 = \text{Step}(\text{Yosi} \cdot 1 + \text{Gil} \cdot 1 + \text{Lea} \cdot 1) = \text{Step}(0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1) = \text{Step}(1) = 1 \quad (1 > 0.5)$$

$$O_1 = \text{Step}(H_1 \cdot 1 + H_2 \cdot 1) = \text{Step}(1 \cdot 1 + 1 \cdot 1) = \text{Step}(2) = 1 \quad (2 > 1.5)$$

$$O_2 = \text{Step}(H_1 \cdot (-1) + H_2 \cdot (-1)) = \text{Step}(-1 \cdot 1 - 1 \cdot 1) = \text{Step}(-2) = 0 \quad (-2 \not> -1.5)$$

← רחל וגיל מכרים.

ד. אם זוג השמות הם מאותה קבוצה – הם אחים. אם השמות הם מ-2 קבוצות שונות הם מכרים.

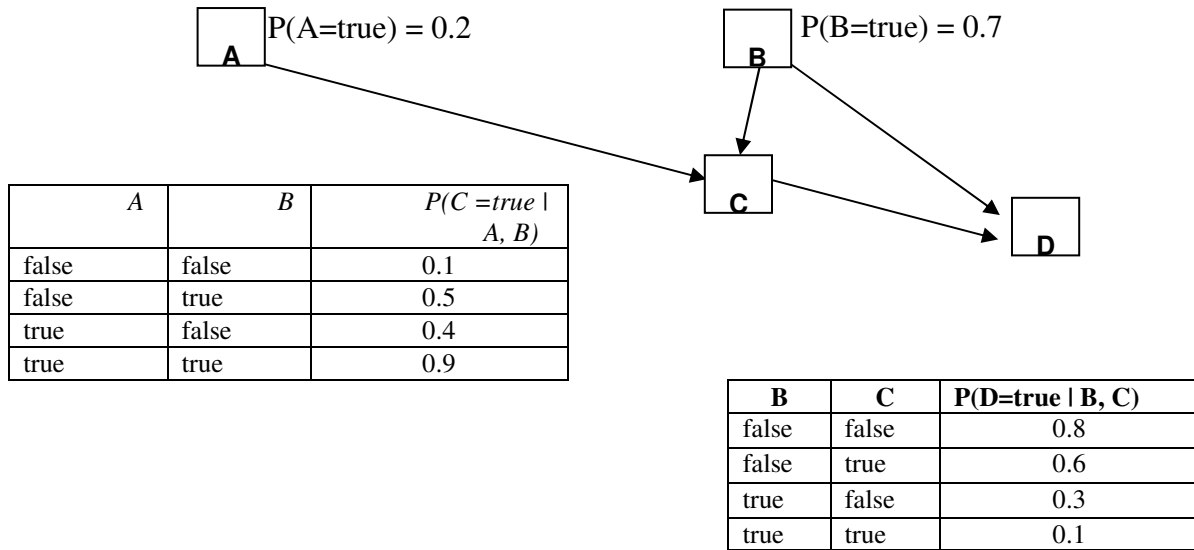
שאלות חזרה – Uncertainty, Naïve Bayes & Bayesian Networks

1. נתונה רשת בייסיאנית להלן. המשתנים A,B,C ו-D הם משתנים בוליאניים.

א. מהי ההסתברות שערך כל המשתנים הוא false?

ב. מהי ההסתברות ש: $B=C=true, D=false$?

ג. מהי ההסתברות $P(C=true|D=true)$?



פיתרון

א.

$$\begin{aligned}
 &P(A=f \wedge B=f \wedge C=f \wedge D=f) \\
 &= P(A=f) \cdot P(B=f) \cdot P(C=f | A=f \wedge B=f) \cdot P(D=f | B=f \wedge C=f) \\
 &= 0.8 \cdot 0.3 \cdot 0.9 \cdot 0.2 = 0.0432
 \end{aligned}$$

ב.

$$\begin{aligned}
 &P(B=t \wedge C=t \wedge D=f) \\
 &= P(B=t) \cdot P(C=t | B=t) \cdot P(D=f | B=t \wedge C=t) \\
 &= P(B=t) \cdot [P(A=t) \cdot P(C=t | A=t \wedge B=t) + P(A=f) \cdot P(C=t | A=f \wedge B=t)] \cdot P(D=f | B=t \wedge C=t) \\
 &= 0.7 \cdot (0.2 \cdot 0.9 + 0.8 \cdot 0.5) \cdot 0.9 = 0.3654
 \end{aligned}$$

ג.

$$\begin{aligned}
 P(C|D) &= \frac{P(D|C) \cdot P(C)}{P(D)} \\
 &= [P(D|C \wedge B) \cdot P(B) + P(D|C \wedge \bar{B}) \cdot P(\bar{B})] \cdot \\
 &\quad [P(C|A \wedge B) \cdot P(A \wedge B) + P(C|A \wedge \bar{B}) \cdot P(A \wedge \bar{B}) + P(C|\bar{A} \wedge B) \cdot P(\bar{A} \wedge B) + P(C|\bar{A} \wedge \bar{B}) \cdot P(\bar{A} \wedge \bar{B})] / \\
 &\quad [P(D|C \wedge B) \cdot P(C \wedge B) + P(D|C \wedge \bar{B}) \cdot P(C \wedge \bar{B}) + P(D|\bar{C} \wedge B) \cdot P(\bar{C} \wedge B) + P(D|\bar{C} \wedge \bar{B}) \cdot P(\bar{C} \wedge \bar{B})]
 \end{aligned}$$

נשתמש במבנה הרשת כדי לקבל-

$$P(A \wedge B) = P(A) \cdot P(B)$$

$$P(C \wedge B) = P(C|B) \cdot P(B) = [P(C|B \wedge A) \cdot P(A) + P(C|B \wedge \bar{A}) \cdot P(\bar{A})] \cdot P(B)$$

2. מחלה מסויימת (מסומנת כ d) נפוצה באוכלוסייה ביחס של אחד למליון. קיימת בדיקה להמצאות המחלה (מסומנת כ test). הבדיקה מדוייקת ב-98% כלומר בשני אחוזים מהבדיקות אדם בריא יקבל תשובה חיובית (תשובה חיובית לבדיקה פירושה חולי) ובשני אחוזים מהבדיקות אדם חולה יקבל תשובה שלילית. אדם נבדק והבדיקה נתנה תוצאות חיוביות. השתמש בהיסק ב- Bayesian reasoning על מנת לחשב את הסיכוי שאדם זה אכן חולה.

Bayes Rule: $P(A|B) = P(B|A) \cdot P(A) / P(B)$

A= "d=true", B="test=positive"

P (d=true | test=positive)

$$= (P(\text{test=positive} | d=\text{true}) \cdot p(d=\text{true})) / p(\text{test=positive})$$

$$= (0.98 \cdot 0.000001) / (0.000001 \cdot 0.98 + 0.999999 \cdot 0.02)$$

$$= 0.00000098 / (0.00000098 + 0.01999998) = 0.00000098 / 0.02000096 \sim 4.9 \cdot 10^{-5}$$

3. גברת כהן היא סטודנטית למדעי המחשב שמאזינה למוזיקה בכל מקום. יש לה לעיתים תכופות שיעורי בית ורבים מהם דורשים תוכנות. יש לנו מספר דוגמאות על בחירת סוג המוזיקה שלה בפעמים שונות. נניח כי ראיתם את גברת כהן בבוקר (Morning). היו לה שיעורי בית (HomeworkDue=Yes) שלא דורשים תוכנות (Programming=No).
א. מה סוג המוזיקה ש- Naive Bayes ינבא שהיא שומעת? הראה את החישובים שלך והסבר.

TimeOfDay	HomeworkDue?	Programming?	MusicType
Morning	Yes	No	Classical
Morning	No	No	Pop
Morning	No	Yes	Classical
Morning	Yes	No	Classical
Afternoon	Yes	Yes	Pop
Afternoon	No	No	Pop
Evening	No	Yes	Pop
Evening	Yes	Yes	Classical

תשובה:

נחשב קודם את ה-prior: $P(\text{Pop}) = 4/8$, $P(\text{Classical}) = 4/8$

נשתמש בהנחת Naïve Bayes:

$$P(X_1, X_2, X_3 | C) = P(X_1 | C) \cdot P(X_2 | C) \cdot P(X_3 | C)$$

$$P(\text{Morning} | \text{Classical}) = 3/4$$

$$P(\text{HomeworkDue=Yes} | \text{Classical}) = 3/4$$

$$P(\text{Programming=No} | \text{Classical}) = 2/4$$

$$P(\text{Morning} | \text{Pop}) = 1/4$$

$$P(\text{HomeworkDue=Yes} | \text{Pop}) = 1/4$$

$$P(\text{Programming=No} | \text{Pop}) = 1/2$$

נקבל עבור Classical:

$$P(\text{Morning, Yes, No} | \text{Classical}) = 3/4 \cdot 3/4 \cdot 2/4 \cdot 4/8 = 9/64$$

ועבור Pop:

$$P(\text{Morning, Yes, No} | \text{Pop}) = 1/4 \cdot 1/4 \cdot 1/2 \cdot 4/8 = 1/64$$

לכן Naïve Bayes ינבא Classical.

ב. מה סוג המוזיקה שינבא k-nearest neighbor עם $k=3$?
 תשובה:
 נחשב עבור כל שורה, מה המרחק שלה מ- $\langle \text{Morning, Yes, No} \rangle$:

TimeOfDay	HomeworkDue?	Programming?	MusicType	Distance
Morning	Yes	No	Classical	0
Morning	No	No	Pop	1
Morning	No	Yes	Classical	2
Morning	Yes	No	Classical	0
Afternoon	Yes	Yes	Pop	2
Afternoon	No	No	Pop	2
Evening	No	Yes	Pop	3
Evening	Yes	Yes	Classical	2

3 הדוגמאות הקרובות ביותר נותנות לנו את הסיווג: Classical, Pop, Classical.
 Classical.

4. מועד א 2009: שאלה 3:

במפעל לייצור מנועים הוחלט ליישם מערכת בקרת איכות מבוססת Naïve Bayes. כל מנוע נבדק ביציאה מפס הייצור באמצעות 3 חיישנים: חיישן תנועה שבודק רעידות (M), גלאי קול שבודק את הרעש שנוצר (N), ומדחום שמודד את טמפרטורת המנוע (T). כל אחד נותן תוצאה בוליאנית- תקלה זוהתה (true) או לא (false). בסופו של דבר צריך לקבל החלטה האם המנוע תקין או לא. בטבלה הבאה מצוינים נתונים שנאספו מכמה מנועים שיצאו מפס הייצור:

Result	M	N	T
Ok	false	false	true
Ok	false	true	false
Ok	false	true	false
Ok	true	false	false
Ok	false	false	false
Bad	true	false	false
Bad	true	true	false
Bad	false	true	true

הגיע מנוע חדש. חיישן התנועה דיווח שהכול תקין, גלאי הקול דיווח על תקלה והמדחום דיווח שהכול תקין. האם יש תקלה במנוע או לא?

ת:

נחשב קודם את ה- prior:

$$P(\text{Result}=\text{Ok}) = 5/8, P(\text{Result}=\text{Bad}) = 3/8$$

$$P(M=\text{false}|\text{Result}=\text{Bad}) = 1/3$$

$$P(N=\text{true}|\text{Result}=\text{Bad}) = 2/3$$

$$P(T=\text{false}|\text{Result}=\text{Bad}) = 2/3$$

$$P(M=\text{false}|\text{Result}=\text{Ok}) = 4/5$$

$$P(N=\text{true}|\text{Result}=\text{Ok}) = 2/5$$

$$P(T=\text{false}|\text{Result}=\text{Ok}) = 4/5$$

נקבל עבור Bad:

$$P(M=\text{false}, N=\text{true}, T=\text{false}|\text{Result}=\text{Bad})P(\text{Result}=\text{Bad}) = 1/3 * 2/3 * 2/3 * 3/8 = 1/18$$

$$P(M=\text{false}, N=\text{true}, T=\text{false}|\text{Result}=\text{Ok})P(\text{Result}=\text{Ok}) = 4/5 * 2/5 * 4/5 * 5/8 = 4/25$$

לכן Naïve Bayes ינבא שאין תקלה במנוע.

5. שאלה 6 מועד א 2007

א. נתון:

A – חולי בשפעת

B – שתיית תה צמחים לחיזוק המערכת החיסונית

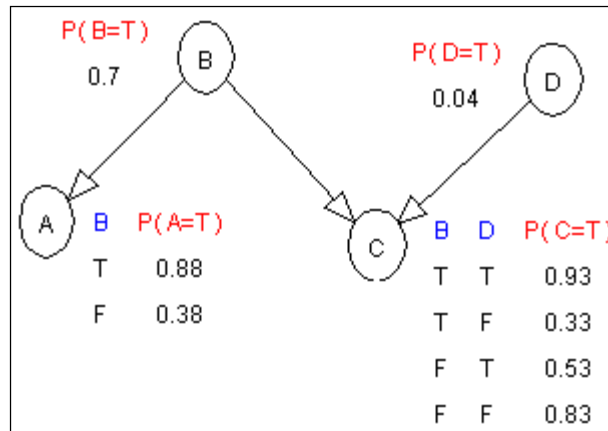
$$P(A)=0.75, P(B)=0.5, P(A|B)=0.1$$

נתון אדם החולה בשפעת. מהי ההסתברות שהוא שתה תה צמחים?

ב. נתונה הרשת הבייסיאנית להלן.

1. מהי ההסתברות ש: $A=B=C=D=false$?

2. מהי ההסתברות ש: $C=false$?



תשובה:

א. $P(B|A)=P(A|B)P(B)/P(A)=0.1*0.5/0.75=0.0667$

ב. $P(-A,-B,-C,-D)=P(-B)*P(-D)*P(-A|-B)*P(-C|-D)=.3*.96*.62*.17=0.03$

$$P(-C)=P(B)P(D)P(-C|B,D) + P(B)P(-D)P(-C|B,-D) + P(-B)P(D)P(-C|-B,D) + P(-B)P(-D)P(-C|-B,-D)$$

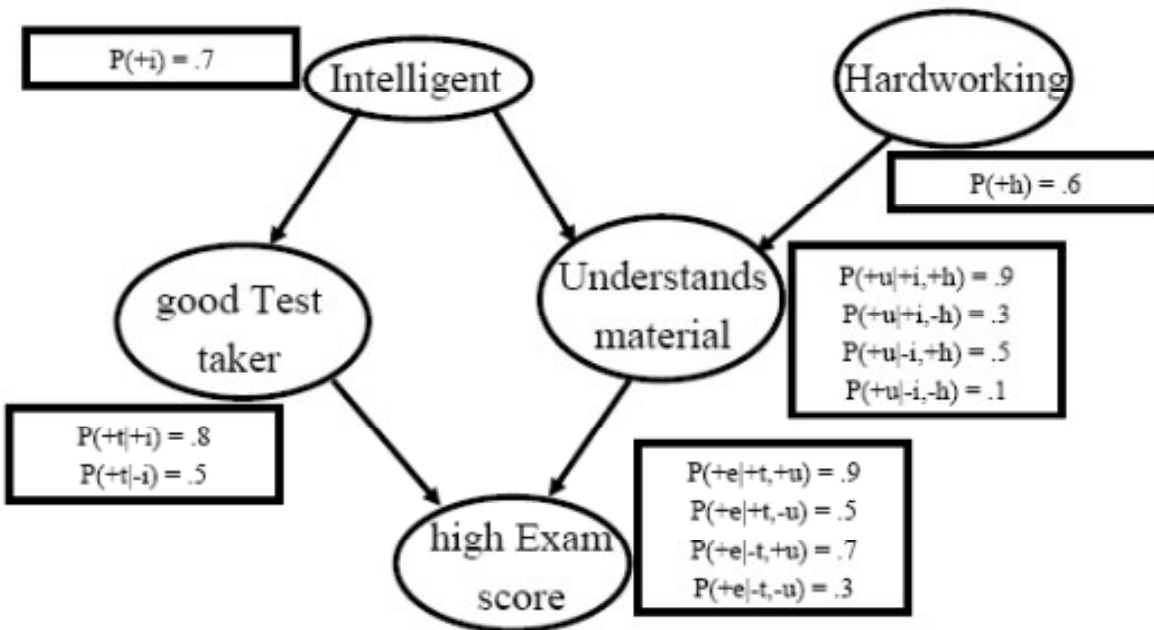
$$=.7*.04*.07$$

$$+.7*.96*.67$$

$$+.3*.04*.47$$

$$+.3*.96*.17=0.00196+0.45024+0.0564+0.04896=0.55756$$

בתור מרצה, אני מעוניינת לדעת האם סטודנט הבין את החומר כאשר ההערכה מתבצעת באמצעות מבחן. התרשים הבא מתאר את הרשת הבייסיאנית המתאימה. כל המשתנים בוליאניים, ונסמן $+x/-x$ כדי לבטא ערך true/false עבור משתנה.



ניתן לראות שציון גבוה במבחן (e) מושפע מהבנת החומר (u) ומהיכולת להתרכז במבחן (t). שני יכולות אלו מושפעות מהאינטליגנציה של הסטודנט (i). הבנת החומר מושפעת גם מחריצות הסטודנט (h). (א) חשב את ההסתברות שסטודנט שהצליח במבחן הבין את החומר, ז"א את $P(+u|+e)$.

ת:

לפי חוק בייס

$$P(u|e) = P(e|u)P(u)/P(e)$$

לפי ההסתברות השלמה

$$P(e|u) = P(e|u, t)P(t) + P(e|u, -t)P(-t)$$

$$P(t) = P(t|i)P(i) + P(t|-i)P(-i)$$

$$P(-t) = 1 - P(t)$$

$$P(u) = P(u|h, i)P(h, i) + P(u|-h, i)P(-h, i) + P(u|h, -i)P(h, -i) + P(u|-h, -i)P(-h, -i)$$

לכן

$$P(e|u) = 0.9*(0.8*0.7+0.5*0.3)+0.7*(1-[0.8*0.7+0.5*0.3]) = 0.842$$

$$P(u) = 0.9*0.6*0.7 + 0.3*0.4*0.7 + 0.5*0.6*0.3 + 0.1*0.4*0.3 = 0.564$$

במקום לחשב את $P(e)$ נחשב את $P(-u|e)$ ונגרמל

$$P(-u|e) = P(e|-u)P(-u)/P(e)$$

$$P(e|-u) = P(e|-u, t)P(t) + P(e|-u, -t)P(-t)$$

$$P(-u) = P(-u|h, i)P(h, i) + P(-u|-h, i)P(-h, i) + P(-u|h, -i)P(h, -i) + P(-u|-h, -i)P(-h, -i)$$

לכן

$$P(e|-u) = 0.5*(0.8*0.7+0.5*0.3)+0.3*(1-[0.8*0.7+0.5*0.3]) = 0.442$$

$$P(-u) = 0.1*0.6*0.7 + 0.7*0.4*0.7 + 0.5*0.6*0.3 + 0.9*0.4*0.3 = 0.436$$

נקבל ש-

$$a^* < 0.842 * 0.564, 0.442 * 0.436 > = 1 \Rightarrow < 0.711336129, 0.288663871 >$$

P(u|e) = 0.711336129 ז"א

(ב) אלו מהטענות הבאות נובע ממבנה הרשת? נמק בקצרה כל תשובה. אין להסתמך על הערכים המספריים אלא רק על מבנה הרשת.

- a. t ו- u בלתי תלויים.
- b. t ו- u בלתי תלויים בהינתן i, e ו- h.
- c. t ו- u בלתי תלויים בהינתן i ו- h.
- d. e ו- h בלתי תלויים בהינתן u.

ת:

- a. לא נובע ממבנה הרשת. להיפך, מכיוון שלשניהם יש אב משותף הם תלויים אחד בשני.
- b. לא נובע ממבנה הרשת. להיפך, מכיוון שלשניהם יש בן משותף, e, שנתון בשאלה.
- c. נובע ממבנה הרשת. מכיוון ש- i ו- h הם אבות של t, t בלתי תלוי ב- u (שאינו צאצא שלו) בהינתן i ו- h.
- d. לא נובע ממבנה הרשת. בהינתן u, i ו- h לא בלתי תלויים, i ו- i הוא אב קדמון (ב- causal chain) של e.

מועד א: 2010: שאלה 2: Bayes Nets

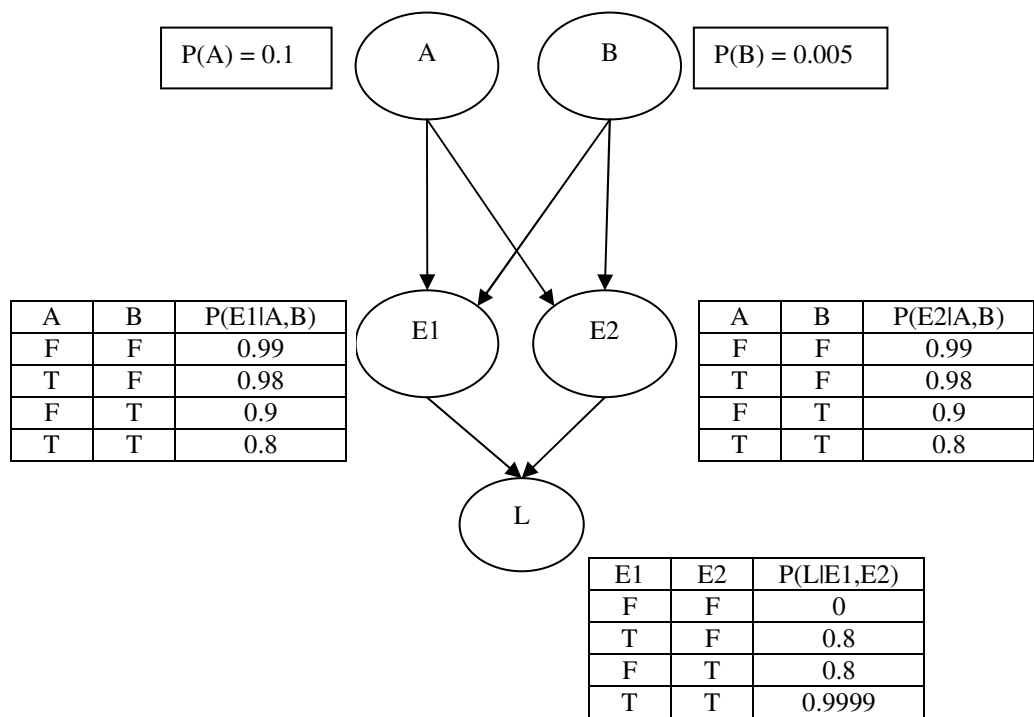
למטוס בואינג 747 ישנם 2 מנועים. ההסתברות לנחיתה מוצלחת כשרק מנוע אחד פועל היא 80%, 0% אם שניהם אינם פועלים ו- 99.99% אם שניהם פועלים כרגיל. בתנאים רגילים, לכל מנוע יש סיכוי של 1 ל- 100 שיפסיק לפעול במהלך הטיסה. ההסתברות גדולה פי 10 אם המטוס נפגע מלהקת ציפורים. דבר נוסף שעלול להשפיע הוא גיל המטוס: במטוס ישן הסיכוי לתקלה כפול ממטוס חדש. כמובן שההשפעה של פגיעת ציפורים במנוע של מטוס ישן גם היא גדולה יותר- יש סיכוי של 20% שהמנוע יפסיק לפעול. בסקר בטיחות שנעשה בשדות תעופה בעולם, התגלה שמתוך סך הטיסות, רק 0.5% נפגעו מציפורים. בנוסף, חברות התעופה מקפידות שאחוז המטוסים הישנים, מתוך כלל המטוסים, יהיה 10% בלבד. נשתמש במשתנים הבאים:

- L - נחיתה מוצלחת.
- E1 - מנוע 1 עובד.
- E2 - מנוע 2 עובד.
- B - פגיעה בלהקת ציפורים.
- A - המטוס ישן.

- א. צייר את הרשת הבייסאנית היעילה ביותר לייצוג בעיה זו, כולל ה- CPT.
- ב. הוכח או הפרך: בכל סדר הכנסה שונה של קודקודים בבניית רשת בייסיאנית מקבלים מבנה אחר של הרשת.

תשובה:

א.



ב. לא נכון. לדוגמה ברשת הקודמת, אם היינו מכניסים בהתחלת הבנייה קודם את A ואח"כ את B, או קודם את B ואח"כ את A, היינו עדיין מקבלים את אותה הרשת.