ת.ז. ____________________ תאריך: 15/04/2012

קורס 67800 תשע"ב

בחינה מועד ב'

הנחיות:

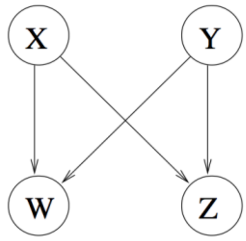
- בבחינה חמש שאלות. עליכם לענות על שלוש מתוכן (משקל כל שאלה 33 נקודות).
- לא ייבדקו יותר משלוש שאלות כך שאם פתרתם שאלות נוספות עליכם לבחור אילו שאלות ייבדקו ולסמן זאת בצורה ברורה.
- עליכם לנמק היטב כל תשובה ולהוכיח באופן מדויק כל טענה שנדרשתם להוכיח. תשובה נכונה ללא נימוק ו/או דרך לא תזכה אתכם בנקודות.
- עליכם לענות על כל השאלות הבאות באופן עצמאי. תלמיד שיתפס מעתיק, או שיהיה קיים חשש כי העתיק יועבר לטיפול רשויות האוניברסיטה.
- אין להשתמש בחומר עזר מכל סוג שהוא.
- (1 נקודות) ודאו כי מספר הזהות שלכם רשום על מחברת הבחינה.

משך הבחינה: שלוש שעות

בהצלחה !

1. בשאלה זו נבחן את השפעת אלגוריתם החיפוש עצמו על לימוד המבנה. נשתמש באלגוריתם חמדני המתחיל מרשת ריקה ובכל שלב מבצע שינוי לקשת אחת (למשל הוספה או הורדה של קשת). השינוי שנבחר הוא הטוב ביותר מבין אלו המותרים, כל עוד הוא מביא לשיפור ממש בפונקצית הציון BIC. אם לכמה שינויים יש את אותו הציון, נבחר ביניהם באופן שרירותי.

א. נתון מבנה הרשת הבייסיאנית הבא:



תנו דוגמא לפרמטרים של הרשת כך שאם נקבל מספר אינסופי של דוגמאות אימון (שנדגמו מההתפלגות המיוצגת על ידי הרשת), אלגוריתם המשתמש רק בפעולת 'הוספה' של קשת ילמד רשת פחות טובה (עם ציון פחות טוב) מאלגוריתם המשתמש בפעולת 'הוספה' וגם בפעולת 'הורדה'. הוכיחו את נכונות הדוגמה.

ב. תנו דוגמא להתפלגות המיוצגת על ידי רשת בייסיאנית (מבנה + פרמטרים) כך שגם אם נקבל מספר אינסופי של דוגמאות אימון שנדגמו מהרשת, אלגוריתם המשתמש בפעולת 'הוספה' וגם בפעולת 'הורדה' בהכרח נתקע במודל שאינו המודל הטוב ביותר. הוכיחו את נכונות הדוגמה.

ג. תנו דוגמא המקיימת את תנאי (ב) אך כך שאם נאפשר גם פעולת 'הפיכה' של קשת אזי האלגוריתם כן יגיע למבנה הטוב ביותר. הוכיחו את נכונות הדוגמה (אפשר להשתמש שוב באותה דוגמא של (ב) אם זה משרת אתכם).

2. נתונה רשת בה X ו-Y שניהם הורים של Z. נניח שכל המשתנים בינאריים, ל X ול Y התפלגות אחידה, ו-Z הוא ה XOR של X ו Y. בכיתה ראינו שאלגוריתם גיבס במקרה כזה לא מתכנס להתפלגות הפוסטריורית הנכונה.

א. כיצד יתנהג אלגוריתם גיבס אם ההתפלגות של Z מעט מורעשת כך שעבור $P(Z|X=x,Y=y)$ ערך ה 0 מוחלף ב-ε וערך ה 1 מוחלף ב $1-\varepsilon$? תנו תשובה איכותית ובנוסף חשבו באופן מפורש את תוחלת מספר האיטרציות באלגוריתם עד לשינוי השמה מההשמה ההתחלתית.

ב. כעת נבחן שינוי באלגוריתם כך שבכל שלב נדגמים <u>מספר</u> משתנים בהינתן האחרים. הראו שבמקרה זה אלגוריתם גיבס מחזיר את התשובה הנכונה גם במקרה ה XOR הדטרמיניסטי אם דוגמים שני משתנים בהינתן השלישי.

ג. נניח כעת כי נרצה להשתמש באלגוריתם המשופר לביצוע חישובים בעזרת דגימה על זוגות משתנים. לצורך כך, הראו כיצד ניתן לחשב את $P(X_i,X_j|\mathbf{V}-\{X_i,X_j\})$ כאשר **V** הם כל המשתנים ברשת, עבור רשת מרקובית כללית.

3. בשאלה זו נבחן ביצוע חישובים ברשת לא מכוונת עם מבנה שרשרת: $X_1 - X_2 - ... - X_n$.
נסמן ב L את מספר הערכים האפשרי לכל משתנה.

א. תארו את עץ הקליקות התואם לסדר אלימינציה המתחיל מאחד הקצוות ומתקדם לשני. הוכיחו כי עבור כל סדר אלימינציה אחר (שלא מתחיל בקצוות) מתקבל עץ קליקות עם רוחב עץ גדול ממש מזה שקיבלתם.

ב. נניח כעת כי ברצוננו לחשב הסתברויות שוליות על זוגות X_i ו X_j אשר אינם בהכרח שכנים ברשת המקורית. הציעו אלגוריתם לחישוב $P(X_i,X_j)$ שהסיבוכיות שלו לינארית ב n. כיצד תלויה הסיבוכיות שלו ב L?

ג. נניח כי אנחנו רוצים לחשב את ההסתברויות השוליות לכל הזוגות מהצורה (X_1, X_i) ברשת. הניחו כי ביכולתכם לשמור תוצאות ביניים ללא הגבלה והציעו אלגוריתם שמחשב את כל ההסתברויות השוליות בזמן לינארי ב n. כיצד תלויה הסיבוכיות שלו ב L? רמז: מיצאו אילו חישובים משותפים עבור משתנים סמוכים.

4. בהנתן פילוג $p(x_1,...,x_n)$ נאמר שגרף לא מכוון G הוא I-map מינימאלי ל p אם מתקיים ש $I(p) \supseteq I_{sep}(G)$ ועבור כל קשת שנוריד מ G דרישה זו כבר לא תתקיים. כמו כן נגדיר לכל משתנה X_i את קבוצת המשתנים $M(i)$ להיות הסט המינימלי של משתנים U כך ש: $\left(X_i \perp V - \{X_i, U\} \mid U\right) \in I(p)$, כאשר V היא קבוצת כל המשתנים. כעת נרכיב גרף G באופן הבא: לכל משתנה X_i ניצור קשת בינו לבין כל המשתנים ב $M(i)$. אם קיימת יותר מ $M(i)$ אחת, בוחרים אחת באופן שרירותי.

א. הניחו ש p חיובי (כלומר p>0 לכל ההשמות). הראו שהגרף המתקבל מקיים: $I(p) \supseteq I_{sep}(G)$.

ב. הניחו ש p חיובי (כלומר p>0 לכל ההשמות). הראו שהגרף המתקבל הוא I-map מינימלי.

ג. נתון פילוג על ארבעה משתנים בינאריים כך ש: $p(x_1=0,x_2=0,x_3=0,x_4=0) = p(x_1=1,x_2=1,x_3=1,x_4=1) = 0.5$. הראו שהגרף המורכב משתי קשתות $X_1 - X_2, X_3 - X_4$ הוא תוצאה אפשרית של האלגוריתם המתואר.

ד. הראו שהגרף מהסעיף הקודם אינו I-map מינימלי עבור הפילוג הנתון. כיצד זה מתיישב עם סעיפים א וב?

5. נתון גרף לא מכוון ללא מעגלים G. אנחנו מעוניינים באמידת הפרמטרים של מודל זה מתוך מדגם אמפירי. נסמן ב $p_D(x)$ את הפילוג האמפירי וב d_i את הדרגה של הקודקוד ה i (כלומר $d_i = |N(i)|$).

א. נניח מודל פרמטרי מהצורה: $p(x_1,...,x_n;\theta) \propto e^{\sum_{ij\in E}\theta_{ij}(x_i,x_j)+\sum_i \theta_i(x_i)}$. הראו שהפרמטרים הבאים מביאים למקסימום את הנראות:

$$\theta_{ij}(x_i,x_j) = \log p_D(x_i,x_j), \theta_i(x_i) = (1-d_i)\log p_D(x_i)$$

ב. נניח כעת מודל מהצורה: $p(x_1,...,x_n;\theta) \propto e^{\sum_{ij\in E}\theta_{ij}(x_i,x_j)}$. מצאו ביטוי סגור לפרמטרים $\theta_{ij}(x_i,x_j)$ שמביאים למקסימום את הנראות.

ג. עבור המודל **<u>מסעיף א</u>**, נניח כעת שהמשתנה X_1 הוא חבוי. תוכלו להניח שיש ל X_1 שכן יחיד X_2. תנו צעד עדכון מפורש לפרמטרים באיטרציה של אלגוריתם EM. אילו פרמטרים יהיו קבועים עבור כל האיטרציות?