ת.ז. ____________________ תאריך: 09/02/2011

קורס 67800 תשע"א
בחינה מועד א'

הנחיות:

- בבחינה חמש שאלות. עליכם לענות על שלוש מתוכם (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 המחברת.

משך הבחינה: שעתיים וחצי

בהצלחה !

1. בהנתן פילוג $p(x_1, x_2, \dots, x_n)$ 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. בהינתן פילוג p חיובי (כלומר p>0 לכל ההשמות) נבנה גרף באופן הבא: אם $X_i \perp X_j | X_{V\setminus\{i,j\}}$ איננה ב $I(p)$ הוסיפו את הקשת i-j ל G.

א. (11 נקודות) הראו שעבור הגרף המתקבל מתקיים $I(p) \supseteq I_{sep}(G)$.

ב. (11 נקודות) הראו שהגרף הוא I-map מינימאלי ל p.

ג. (11 נקודות) נתונה רשת הבייסיאנית $X_1 \rightarrow X_2 \leftarrow X_3$. נגדיר את F להיות משפחת כל ההתפלגויות p שמתפרקות על פי הרשת. מצאו את p ב-F כך שגרף הלא מכוון G שהוא I-map מינימאלי של p, הוא בעל מספר קשתות מקסימאלי. הוכיחו את נכונות תשובתכם.

2.

א. (15 נקודות) נתון המודל הפרמטרי $p(x_1, x_2, x_3; \theta) \propto e^{0.5\sum_{ij \in E} \theta(1 - x_i x_j)}$ כאשר E קשתות של הגרף המלא על שלושה קודקודים, ו θ פרמטר סקלרי יחיד. כמו כן $x_i \in \{-1,1\}$. בהינתן מדגם מצאו ביטוי סגור לאומד ML עבור θ כפונציה של המדגם.

ב. (7 נקודות) וודאו שעבור המדגם האחיד מתקבלת תוצאה הגיונית בסעיף א. הסבירו.

ג. (11 נקודות) נניח שיש לכם מדגם בגודל אחד כאשר הערך של x_1 חבוי ועבור המשתנים האחרים: $x_2 = 1, x_3 = 1$. מהו הערך של θ אחרי איטרצייה אחת של אלגוריתם EM עם ערך התחלתי $\theta = 0$.

3. בידינו רשת בייסיאנית סבוכה המייצגת את ההתפלגות P(**X**). ברצוננו לחשב P(**H** |**E**=**e**) עבור **H**∪**E**=**X** כאשר אין ביכולתנו לחשב או לדגום P(**H** |**E**=**e**) באופן ישיר. נניח כי בידינו Q(**X**) הקרוב ל-P(**X**) ושמיוצג על ידי רשת בייסיאנית עם מבנה של עץ.

   א. (13 נקודות) הסבירו במפורט כיצד לבנות עץ קליקות ל-Q וכיצד תשמשו בעץ לחשב את Q(**H** |**E**=**e**).

   ב. (13 נקודות) הסבירו במפורט כיצד תשמשו בחישוב של (א) על מנת לייצר דגימה מ Q(**H** |**E**=**e**).

   ג. (7 נקודות) תארו במדויק כיצד נוכל להשתמש בדגימות של (ב) על מנת לחשב את P(**H** =**h**|**E**=**e**) עבור **h** נתון.

4. ביישום אלגוריתם אלימינציה נתון (variable elimination), פקטורים מוכפלים מדי פעם בקבוע שרירותי על מנת להבטיח יציבות נומרית. מספר ההכפלות בחישוב מסוים אינו ידוע, בכל הכפלה ייתכן שימוש בקבוע אחר, ואף אחד מהקבועים לא נשמר. אנו מתעניינים בשימוש באלגוריתם זה לחישוב התפלגויות שוליות $P(X_i)$, התפלגויות שוליות מותנות $P(X_i|\mathbf{E}=\mathbf{e})$ והסתברות התצפיות $P(\mathbf{E}=\mathbf{e})$.

א. (10 נקודות) האם ניתן לחשב את הסתברות התצפיות בעזרת האלגוריתם? נמקו במדויק מדוע לא ניתן לבצע את החישוב או הסבירו בקצרה כיצד ניתן לבצע אותו והוכיחו את נכונות החישוב.

ב. (8 נקודות) האם ניתן לחשב התפלגויות שוליות מותנות בעזרת האלגוריתם? נמקו במדויק מדוע לא ניתן לבצע את החישוב או הסבירו בקצרה כיצד ניתן לבצע אותו, והוכיחו את נכונות החישוב.

ג. (5 נקודות) האם ניתן לחשב התפלגויות שוליות (לא מותנות) בעזרת האלגוריתם? נמקו במדויק מדוע לא ניתן לבצע את החישוב או הסבירו בקצרה כיצד ניתן לבצע אותו והוכיחו את נכונות החישוב.

ד. (10 נקודות) נניח כעת כי ידועים הקבועים בהם הוכפלו הפקטורים במהלך החישובים, אך לא ידוע איזה פקטור הוכפל באיזה קבוע. האם במקרה זה ניתן לחשב את הסתברות התצפיות? נמקו במדויק מדוע לא ניתן לבצע את החישוב או הסבירו בקצרה כיצד ניתן לבצע אותו והוכיחו את נכונות החישוב.

5. נניח כי ברצוננו ללמוד את המבנה של רשת בייסיאנית עבור שני משתנים מקריים בלבד X ו-Y. לשם פשטות נניח לאורך השאלה ש X ו Y בינאריים.

א. (12 נקודות) הוכיחו שעבור מדגם כלשהוא, ציון הנראות (likelihood score) הוא זהה עבור המבנה בו X הורה של Y ועבור המבנה בו Y הורה של X. שימו לב שאינכם יכולים להניח כי ידועה הטענה הכללית יותר שהזכרנו בכיתה לפיה ציון הנראות (או ה-BIC) זהה לכל אוסף רשתות המקודדות בדיוק את אותן אי-תלויות.

ב. (11 נקודות) ניתן לתאר מדגם על X ו-Y עד ידי השכיחות יחסית לכל השמה (כלומר טבלה $p_D(x,y)$ שאבריה אי-שליליים ומסתכמים לאחד) ומספר הדוגמאות הכולל M (לצורך פשטות, נניח בסעיף זה כי ניתן לקיים כל שכיחות יחסית בכל גודל מדגם). תארו איכותית מדגם שעבורו קיים M כך הרשת האופטימאלית לפי ציון הנראות (likelihood score) שונה מהרשת האופטימאלית לפי ה- Bayesian Information Criterion (BIC). כמו כן דאגו שהשכיחות של כל משתנה תהיה אחידה (כלומר $p_D(x)=p_D(y)=0.5$ לכל ערכי x ו y). שימו לב שעליכם לתאר איכותית מדוע קיים M כזה אבל אין צורך להוכיח מתמטית את קיומו.

ג. (10 נקודות) הוכיחו פורמאלית כי קיים M כך שעבור מדגם עם אותה שכיחות יחסית שתיארתם ב-(ב), יעדיפו שני הציונים את אותה הרשת.