

Institutt for Data og Informasjonsvitenskap

## Eksamensoppgave i TDT4171 Metoder i AI

**Faglig kontakt under eksamen: Gleb Sizov**

**Tlf.: : 735 91802**

**Eksamensdato: 5. Juni 2013**

**Eksamenstid (fra-til): 0900 - 1300**

**Hjelpemiddelkode/Tillatte hjelpemidler: D**

*Ingen trykte eller håndskrevne hjelpemidler tillatt.*

*Bestemt, enkel kalkulator tillatt*

### **Annen informasjon:**

Sensurfrist: 26. Juni, 2013.05.27.

Alle spørsmål og alle delspørsmål skal besvares. Hvert spørsmål vektes som angitt i oppgaveteksten.

**Målform/språk: Bokmål**

**Antall sider: 4 (inkludert denne forsiden)**

**Antall sider vedlegg: 0**

**Kontrollert av:**

---

Dato

Sign

## Oppgave 1 – 35%

- a) Beskriv syntaks og semantikk for Bayesianske Nettverk.
- b) Lag en modell for det følgende domenet med et Bayesiansk nettverk. Lag modellen så enkel og lettforståelig som mulig:  
*Melk fra en ku er infisert med sannsynlighet 0.01. Det finnes en test for å detektere om melken er infisert eller frisk. Testen gir enten et positivt resultat (som indikerer infisert melk) eller et negativt resultat (som indikerer frisk melk). Testen er ikke perfekt, da den gir et positivt resultat på frisk melk med sannsynlighet 0.05 og et negativt resultat på infisert melk med sannsynlighet 0.02.*

Du skal oppgi **både** den grafiske strukturen **og** de betingede sannsynlighetsfordelingene når du besvarer denne del-oppgaven.

- c) En test blir gjort for å analysere melken, og kommer tilbake med et positivt resultat, dvs. testen indikerer at melken er infisert. Regn ut sannsynligheten for at melken faktisk er infisert, gitt at testen ga et positivt resultat. Bruk tallene fra modellen du definerte i Spørsmål (b).
- d) Utvid modellen fra Spørsmål (b) på følgende måte:  
*Etter å ha fått testresultatet kan bonden bestemme seg for enten å lagre melken eller tømme den ut. Hvis han tar vare på melken, og den ikke er infisert, tjener han \$100 på at han kan selge den senere. Hvis han lagrer melken og den faktisk er infisert, vil den ødelegge all melken bonden har lagret unna allerede, og dermed taper bonden \$1000. Hvis han tømmer ut melken vil han tjene \$0, uansett om melken er frisk eller infisert.*

Du skal oppgi **både** den kvalitative delen (den grafiske strukturen) **og** den kvantitative delen (tallene i modellen) når du besvarer denne del-oppgaven.

- e) Hva kaller vi den typen modeller som du laget i Spørsmål (d), og hvordan kan bonden bruke den til å ta den for ham beste beslutningen? Gjør beregningene, og foreslå den beste beslutningen for bonden. Husk at testen ga positivt resultat. Hvilket prinsipp har du brukt for å anbefale hvilket valg bonden bør ta?

**Merk:** Hvis du ikke klarte å beregne  $P(\text{melken er infisert} \mid \text{testresultatet er positivt})$  i Spørsmål (c) kan du bare anta en verdi for denne sannsynligheten nå, og gjøre beregningene fra det startpunktet. Skriv tydelig i besvarelsen din hvilke tall du har benyttet.

- f) Utvid modellen fra Spørsmål (b) med følgende informasjon:  
*Det er kjent at mer melk infiseres når det er høy luftfuktighet. Videre er testen mindre pålitelig når lufttemperaturen er høy.*

Du skal kun oppgi graf-strukturen, og trenger **ikke** angi betingede sannsynlighetstabeller i dette delspørsmålet.

- g) Synes du Bayesianske nettverk er et *naturlig* modellerings-rammeverk for dette problemområdet? Hvordan kan man karakterisere problemområder der Bayesianske nettverk med suksess kan brukes? Kan du gi et eksempel på et problemområde der Bayesianske nettverk passer *dårlig*?

### **Spørsmål 2 - 15%**

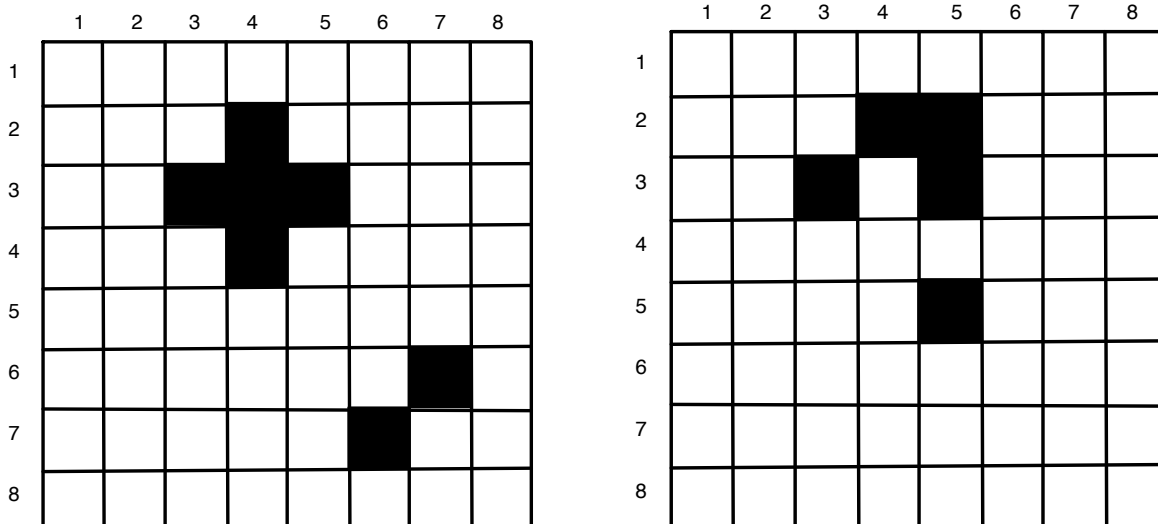
- a) *Beskriv* hovedideene med case-basert resonering (CBR). Du kan for eksempel gjøre dette ved å sammenligne CBR med en lærealgoritme for beslutningstrær for å belyse poengene dine.
- b) Gi en kortfattet beskrivelse av hva som skjer i hvert av de fire skrittene av CBR-syklusen, og diskuter hvordan generell domenekunnskap kan benyttes i hvert av disse fire skrittene.

### **Spørsmål 3 - 20%**

- a) En av antagelsene som ligger under *Skjulte Markov-kjeder (Hidden Markov Models)* er den såkalte "Markov-antagelsen." Forklar hvorfor denne antagelsen blir gjort. Vis den grafiske strukturen av en Skjult Markov-kjede, og angi hvordan denne strukturen ville sett ut dersom Markov-antagelsen ikke ble gjort
- b) I hvilke situasjoner kan man benytte en Markov Beslutningsprosess (Markov Decision Process) for å løse et beslutningsproblem? Gi en kort generell beskrivelse, samt ett konkret eksempel. Hva er hovedideen bak "*value iteration*" algoritmen, som kan brukes til å løse Markov Beslutningsprosessene.
- c) Diskuterer påstanden "*Dersom Kunstig Intelligens blir en realitet vil det bety slutten for menneskeheten*". Hvilke grunner kan man ha for å komme med en slik påstand, og hva kan gjøres for å unngå at dette blir en realitet? Tror du på denne påstanden? Hvorfor (ikke)?

#### Oppgave 4 – 30%

Du har blitt spurt om å analysere en stor mengde med binære bilder. Hvert bilde består av 8 ganger 8 piksler i et grid, og du kan lese av verdien for pikselen på hver posisjon i gridet. Hvite piksler har verdi 1, mens svarte piksler har verdi 0. To eksempler er angitt i figuren under.



For bildet på vestre hånd vil man, for eksempel, kunne få vite at verdien i celle (1,1) = 1 fordi bildet har en hvit piksel øverst til venstre. . Videre er verdien av (2,4)=0, fordi elementet på posisjon (2,4) er svart. Oppgaven din er å lage et system som automatisk angir om bildet inneholder et svart kors. Bildet over til venstre er eksempel på et slikt, da pikslene i posisjon (2,4), (3,3), (3,4), (3,5) og (4,4) alle er svarte. Bildet på høyre hånd er et eksempel på et bilde som *ikke* inneholder et kors.

- I stedet for manuelt å bygge en modell som kan løse dette problemet bør du foretrekke å lære denne modellen fra dataene. Hva karakteriserer situasjoner der det er nyttig å lære modeller fra data (i stedet for å generere dem manuelt)?
- Vil du foretrekke å bruke et nevralt nettverk (artificial neural network) eller et beslutnings-tre (decision tree) til denne oppgaven? Begrunn svaret ditt.
- Beskriv hvordan du ville gjennomført lærefasen:
  - Hvordan ville du ha strukturert input-dataene?
  - Hvilken lærealgoritme vil du bruke, og hva er det denne lærealgoritmen gjør?Gi en *kortfattet* beskrivelse.
- Du blir så bedt om å tilpasse løsningen din, slik at systemet ikke bare detekterer om det er (minst ett) kors i et bilde, men også kan finne ut *hvor mange kors* det er i hvert bilde. Hvordan vil du videreutvikle løsningen din for å få til dette?