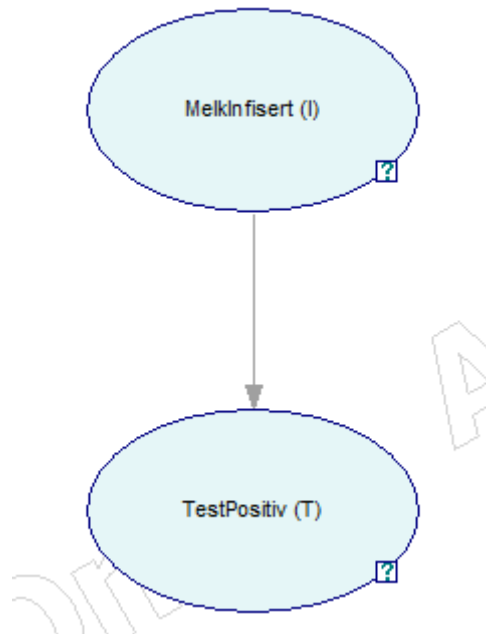


Eksamen TDT4171

Oppgave 1.I

Strukturen på dette bayesianske nettverket er svært enkel – den består av to noder, MelkInfisert (I) og TestPositiv (T), og en kobling mellom disse, $I \rightarrow T$



De betingende sannsynlighetene er som følger:

| P (I = sann) | P (I = usann) |
|---------------------|----------------------|
| 0.01 | 0.99 |

| I | P (T = sann I) | P (T = usann I) |
|--------------|-------------------------|--------------------------|
| <i>sann</i> | 0.98 | 0.02 |
| <i>usann</i> | 0.05 | 0.95 |

Disse verdiene er hentet rett fra oppgaveteksten – sannsynligheten for infisert melk er 0.01, sannsynligheten for positiv test ved frisk melk er 0.05 og sannsynligheten for negativ test på infisert melk er 0.02

Oppgave 1.II

En test kommer tilbake med positivt resultat. Regn ut sannsynligheten for at melken faktisk er infisert, gitt positiv test.

Vi skal finne $P(i | t)$. Ved å bruke Bayes' regel kan vi skrive dette om til

$$P(i | t) = \frac{P(t | i) * P(i)}{P(t)}$$

$P(i)$ og $P(t | i)$ kan vi lese direkte av fra sannsynlighetstabellene. $P(t)$ kan vi finne ved å bruke $P(A) = P(A \wedge B) + P(A \wedge \neg B) = P(A | B) * P(B) + P(A | \neg B) * P(\neg B)$:

$$P(t) = P(t | i) * P(i) + P(t | \neg i) * P(\neg i)$$

$$P(t) = 0.98 * 0.01 + 0.05 * 0.99$$

$$P(t) = 0.0593$$

Vi kan sette inn denne verdien, samt de andre fra sannsynlighetstabellene:

$$P(i | t) = \frac{P(t | i) * P(i)}{P(t)}$$

$$P(i | t) = \frac{0.98 * 0.01}{0.0593}$$

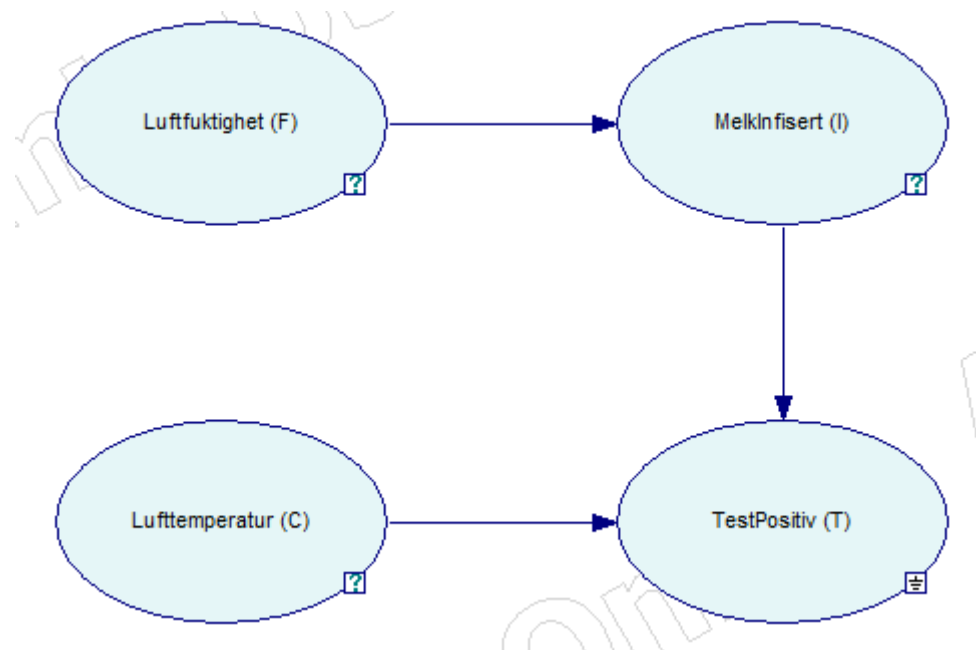
$$P(i | t) = \mathbf{0.165}$$

Sannsynligheten for at melken er infisert hvis testen er positiv er $0.165 = 16.5\%$

Oppgave 1.III

Det er kjent at mer melk infiseres når det er høy luftfuktighet. Videre er testen mindre pålitelig når lufttemperaturen er høy. Oppgå grafstruktur og foreslå relevante verdier for de betingede sannsynlighetstabellene.

Strukturen er fortsatt forholdsvis enkel – det er to nye noder, luftfuktighet og lufttemperatur, som påvirker hver sin originale node i henhold til oppgaveteksten.



Jeg har gjort følgende antakelser:

- Både Luftfuktighet (F) og Lufttemperatur (C) har to tilstander – høy og lav. Dette indikerer når verdiene er høye nok til å påvirke hhv. infeksjonsraten og testene.
- Jeg har antatt at de ikke påvirker hverandre. Jeg er usikker på hvor relevant dette er, da jeg ikke har særlig kunnskap om meteorologi. Jeg føler likevel det er en relativt trygg antakelse.
- Jeg antar at gården ligger relativt nært kysten, og at luftfuktigheten derfor er høy nesten halve året.
- Jeg antar at testen fungerer optimalt store deler av året, men at den er mindre pålitelig grunnet temperatur et par måneder i løpet av sommeren.
- Jeg antar at testen fungerer dårligere «i begge retninger» når temperaturen er høy. Med dette mener jeg at sannsynligheten for både falske negative og falske positive resultater øker når luftfuktigheten er høy.
- Jeg antar verdiene vi fikk over er ved lav luftfuktighet og temperatur, altså er de utgangspunktet ved optimale forhold.

Basert på disse antakelsene vil jeg foreslå følgende sannsynlighetstabeller (se neste side):

| P (F = høy) | P (F = lav) |
|--------------------|--------------------|
| 0.4 | 0.6 |

| P (C = høy) | P (C = lav) |
|--------------------|--------------------|
| 0.2 | 0.8 |

| F | P (I = sann F) | P (I = usann F) |
|------------|-------------------------|--------------------------|
| <i>høy</i> | 0.05 | 0.95 |
| <i>lav</i> | 0.01 | 0.99 |

| I | C | P (T = sann I, C) | P (T = usann I, C) |
|--------------|------------|----------------------------|-----------------------------|
| <i>sann</i> | <i>Høy</i> | 0.9 | 0.1 |
| | <i>Lav</i> | 0.98 | 0.02 |
| <i>usann</i> | <i>Høy</i> | 0.1 | 0.9 |
| | <i>Lav</i> | 0.05 | 0.95 |

Det viktigste å ta med seg fra dette er at sannsynligheten for infisert melk øker når temperaturen er høy, og testen treffer dårligere når luftfuktigheten er høy.