Luento 5: Tietämyksen esittäminen moniagenttijärjestelmissä

Raul Hakli

Tietämyksen esittäminen

- Tietämyksen esittäminen (knowledge representation) on yksi tekoälyn tärkeimpiä ongelmia.
- Perinteinen tapa on käyttää logiikkaa, ts. jotakin formaalia kieltä, joka mahdollistaa myös päättelyn.
- Tekoäly-kurssilla käsitellyt propositio- ja predikaattilogiikka eivät tarjoa riittävän vahvoja välineitä moniagenttijärjestelmien tarpeisiin. Nyt tarkastellaan *modaalilogiikkaa*, joka mahdollistaa usean agentin tietämyksen esittämisen sekä myös muiden *propositionaalisten asenteiden* (kuten uskomukset ja halut) käsittelyn.

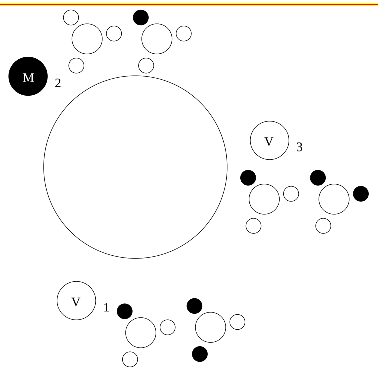
Ympäristön havaittavuus

- Peliteoriassa oletetaan yleensä, että pelitilanne, toimintavaihtoehdot ja lopputulemat ovat kaikkien agenttien tiedossa; itse asiassa oletetaan jopa että ne ovat yleistä tietoa (common knowledge).
- Yleensä tilanne ei kuitenkaan ole näin yksinkertainen, vaan eri agentit tietävät eri asioita ja joutuvat sekä kommunikoimaan että päättelemään mitä toiset agentit tietävät tai uskovat.

Värilliset hatut

- Oletetaan että n henkilöä istuu pöydän ääressä joko musta tai valkoinen hattu päässä. Kaikki voivat nähdä muiden hattujen värin paitsi omansa.
- Oletetaan että on yleistä tietoa että kaikki puhuvat totta, näkevät toistensa hatut, kuulevat hyvin ja ovat täydellisiä loogisia päättelijöitä.
- Ulkopuolinen tarkkailija astuu sisään ja ilmoittaa: "Jokaisella teistä on joko musta tai valkoinen hattu ja ainakin yhden hattu on valkoinen".
- Tarkkailija kysyy "Kuinka moni teistä tietää oman hattunsa värin?" Jos kukaan ei tiedä, hän kysyy uudestaan.
- Kun kysymys on esitetty k kertaa, missä k on valkoisten hattujen lukumäärä, jokainen jolla on valkoinen hattu vastaa myöntävästi.

Agenttien tiedolliset tilat



Modaalilogiikka

Modaalilogiikka laajentaa tavanomaista logiikkaa *modaalioperaattoreilla* $\Box$ ja $\Diamond$. Tavanomaisessa *aleettisessa* modaalilogiikassa niiden intuitiivinen tulkinta on seuraava:

$\Box p$ välttämättä p

$\Diamond p$ mahdollisesti p

Operaattorit voidaan määritellä toistensa avulla: $\Box p =^{\text{def}} \neg\Diamond\neg p$. Operaattorit ovat toistensa *duaalit*. Usein modaalilogiikassa määritellään myös *tiukka* implikaatio: $p \,\Box\!\!\rightarrow q =^{\text{def}} \Box p \rightarrow q$.

Mahdolliset maailmat

- Modaalilogiikan semantiikassa käytetään hyväksi *mahdollisen maailman* käsitettä. Mahdollisiksi maailmoiksi ajatellaan kaikki vaihtoehtoiset asiantilat. Jos nyt *aktuaalisessa maailmassa* esimerkiksi Helsingissä sataa vettä, niin mahdollisiksi maailmoiksi voidaan ajatella sellaiset tilanteet, joissa Helsingissä paistaa aurinko, sataa lunta, on pilvipoutaa jne.
- Aleettisessa logiikassa vaaditaan tavallisesti että maailmat ovat *normaaleja*, millä tarkoitetaan että niissä vallitsevat klassisen logiikan lait.
- Nyt ajatellaan, että lause $\Box p$ on tosi, joss lause p on tosi kaikissa loogisesti mahdollisissa maailmoissa. Vastaavasti lause $\Diamond p$ on tosi, joss lause p on tosi jossakin mahdollisessa maailmassa.

Episteeminen ja doksastinen logiikka

- *Episteemisellä* logiikalla tarkoitetaan modaalilogiikkaa, jossa modaalioperaattorille annetaan tulkinta "agentti tietää että". *Doksastisessa* logiikassa "agentti uskoo että".
- Episteemisen (ja doksastisen) logiikan idea on peräisin Georg Henrik von Wrightiltä ja sen tunnetuin kehittäjä on Jaakko Hintikka.
- Episteemisen logiikan syntaksissa operaattori $\Box$ korvataan operaattorilla K (doksastisessa B). Monen agentin tapauksessa käytetään vielä alaindeksejä $\rightarrow$ multimodaalinen logiikka.
- Mahdolliset maailmat tulkitaan nyt agentin episteemisiksi tai doksastisiksi vaihtoehdoiksi joilla tarkoitetaan niitä maailmoja, jotka ovat yhteensopivia agentin tiedon (tai uskomusten) kanssa.
- Nyt $K_a p$ on tosi, joss p on tosi kaikissa agentin a tietämysmaailmoissa. Vastaavasti B_a on tosi, joss p on tosi kaikissa agentin a uskomusmaailmoissa.

Esimerkki

- On myös mahdollista yhdistää episteemisen ja doksastisen logiikan operaattoreita yhdeksi logiikaksi. (Yleensäkin on mahdollista yhdistää toisiinsa useita erilaisia modaalilogiikoita ja näin usein tehdäänkin moniagenttijärjestelmien yhteydessä. Tämä on yksi tärkeimmistä modaalilogiikan suosion syistä.)
- Logiikassa voidaan nyt ilmaista monimutkaisia lauseita, jotka kuvaavat agenttien tieto- ja uskomustiloja.
- Esimerkiksi lause "Matti uskoo, että Satu tietää, että Raul ei tiedä, tietääkö Juho, että Lontoossa sataa" voidaan ilmaista lyhyesti muodossa

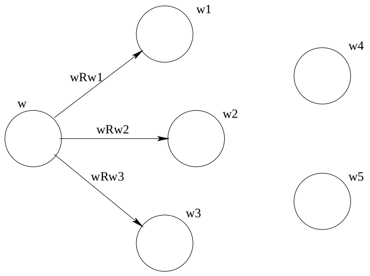
$$B_M K_S(\neg K_R K_J p \wedge \neg K_R \neg K_J p),$$

missä p vastaa atomilausetta "Lontoossa sataa".

Saavutettavuusrelaatio

- Ajatellaan vielä, että aina ei tarvitse tarkastella kaikkia mahdollisia maailmoja yhtäaikaa, vaan että se mitkä maailmat voivat olla vaihtoehtoisia maailmoja, riippuu nykyisestä maailmasta.
- Näin saadaan *saavutettavuusrelaatio* tai *vaihtoehtorelaatio* R, joka kertoo mitkä maailmat ovat toistensa mahdollisia vaihtoehtoja.
- Kun maailmojen joukko W ja saavutettavuusrelaatio R on kiinnitetty, saamme *Kripke-kehyksen* $\langle W, R\rangle$.

Kripke-kehys



$W = \{w, w_1, w_2, w_3, w_4, w_5\}$
$R = \{(w, w_1), (w, w_2), (w, w_3)\}$

Valuaatiot

- Kun Kripke-kehykseen liitetään *valuaatio* (tai totuusarvojakelu) π, joka kertoo mitkä lauseet ovat tosia missäkin maailmassa, saadaan *Kripke-malli* $M = \langle W, R, \pi\rangle$.
- Tässä siis π ilmoittaa kustakin logiikan lausemuuttujasta p ja maailmasta w onko p tosi maailmassa w. (Teknisesti π voidaan määritellä monella tavalla: funktioksi lausemuuttujien joukolta maailmojen potenssijoukolle, maailmojen joukolta lausemuuttujien potenssijoukolle tai maailmojen joukon ja lausemuuttujien joukon karteesiselta tulolta totuusarvojen joukolle.)
- Nyt meillä on tarvittavat ainekset modaalilogiikan *semantiikan* määrittelemiseksi.

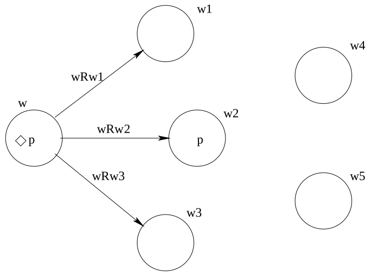
Semantiikka

Olkoon M Kripke-malli $\langle W, R, \pi\rangle$. Jos lause p on tosi mallin M maailmassa w, merkitään $\langle M, w\rangle \models p$, ja jos epätosi, niin merkitään $\langle M, w\rangle \not\models p$. Loogisten operaattoreiden totuusehdot voidaan nyt määritellä seuraavasti:

1. $\langle M, w\rangle \models p$ joss $p \in \pi(w)$ (tms.)
2. $\langle M, w\rangle \models \neg\phi$ joss $\langle M, w\rangle \not\models \phi$
3. $\langle M, w\rangle \models \phi \vee \psi$ joss $\langle M, w\rangle \models \phi$ tai $\langle M, w\rangle \models \psi$
4. $\langle M, w\rangle \models \Box\phi$ joss kaikille $w' \in W$ joille pätee wRw', $\langle M, w'\rangle \models \phi$.
5. $\langle M, w\rangle \models \Diamond\phi$ joss on olemassa $w' \in W$ jolle pätee wRw', $\langle M, w'\rangle \models \phi$.

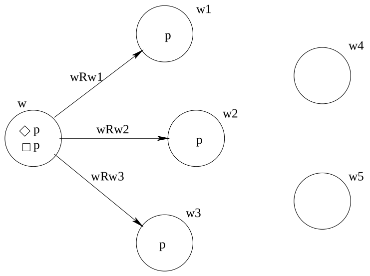
Kaava ϕ on *validi mallissa* $M = \langle W, R, \pi\rangle$, merkitään $M \models \phi$, joss $\langle M, w\rangle \models \phi$ kaikilla $w \in W$. Kaava ϕ on *validi kehyksessä* $\langle W, R\rangle$, merkitään $\langle W, R\rangle \models \phi$, joss ϕ on validi jokaisessa mallissa $\langle W, R, \pi\rangle$.

Kripke-malli (1)



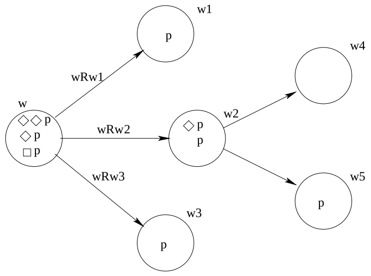
$W = \{w, w_1, w_2, w_3, w_4, w_5\}$
$R = \{(w, w_1), (w, w_2), (w, w_3)\}$
$\pi(w) = \{\}, \pi(w_1) = \{\}, \pi(w_2) = \{p\}, \pi(w_3) = \{\}, \pi(w_4) = \{\}, \pi(w_5) = \{\}$

Kripke-malli (2)



$W = \{w, w_1, w_2, w_3, w_4, w_5\}$
$R = \{(w, w_1), (w, w_2), (w, w_3)\}$
$\pi(w) = \{\}, \pi(w_1) = \{p\}, \pi(w_2) = \{p\}, \pi(w_3) = \{p\}, \pi(w_4) = \{\}, \pi(w_5) = \{\}$

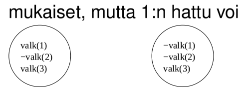
Kripke-malli (3)



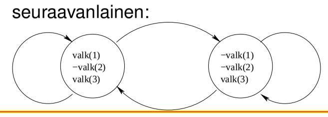
$W = \{w, w_1, w_2, w_3, w_4, w_5\}$
$R = \{(w, w_1), (w, w_2), (w, w_3), (w_2, w_4), (w_2, w_5)\}$
$\pi(w) = \{\}, \pi(w_1) = \{p\}, \pi(w_2) = \{p\}, \pi(w_3) = \{p\}, \pi(w_4) = \{\}, \pi(w_5) = \{p\}$

Kripke-mallit episteemisessä logiikassa

- Kun tarkastellaan episteemistä logiikkaa, mahdolliset maailmat ymmärretään agentin episteemisiksi vaihtoehdoiksi.
- Esimerkiksi tilanteessa, jossa agentti 1 tietää kahden muun agentin hattujen värin mutta ei omaansa, agentti 1 pitää mahdollisina niitä kahta maailmaa, joissa 2:n ja 3:n hatut ovat agentin 1 tiedon mukaiset, mutta 1:n hattu voi olla joko musta tai valkoinen:

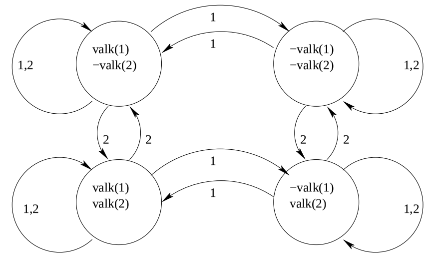


- Agentti pitää molempia maailmoja mahdollisina riippumatta siitä kumpi maailma on oikea, joten saavutettavuusrelaation on oltava seuraavanlainen:



Monen agentin tapaus

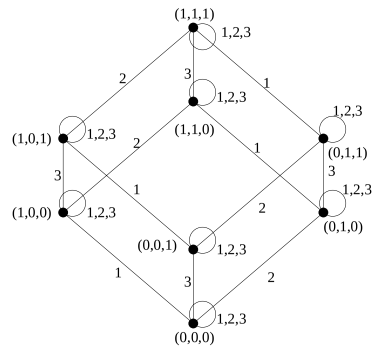
Monen agentin tapauksessa agentit tietävät eri asioita joten ne pitävät eri maailmoja toistensa vaihtoehtoina → kullekin agentille i täytyy määritellä oma saavutettavuusrelaationsa R_i. Alla on kuvattuna kahden agentin tapaus:



Kuvasta nähdään esim. että $K_1(K_2 valk(1) \vee K_2 \neg valk(1))$.

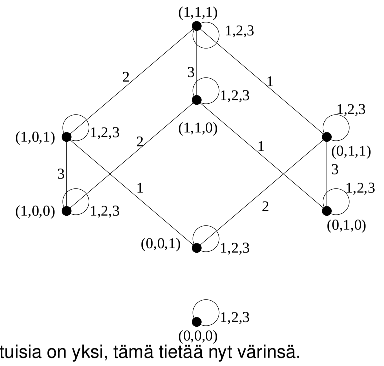
Kolme agenttia ja hatut

Alkutilanne:



Kolme agenttia ja hatut (2)

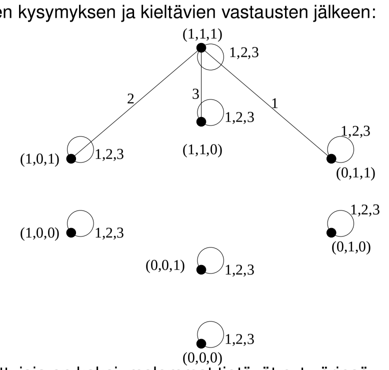
Ilmoituksen jälkeen:



Jos valkohattuisia on yksi, tämä tietää nyt värinsä.

Kolme agenttia ja hatut (3)

Ensimmäisen kysymyksen ja kieltävien vastausten jälkeen:



Jos valkohattuisia on kaksi, molemmat tietävät nyt värinsä.

Normaalit modaalilogiikat

Normaaleissa modaalilogiikoissa, joihin edellä tarkasteltuun Kripke-semantiikkaan perustuvat logiikat kuuluvat, ovat voimassa seuraavat ominaisuudet:

(MP) jos $\models \phi$ ja $\models \phi \rightarrow \psi$, niin $\models \psi$

(K) $\models \Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$

(RN) jos $\models \phi$, niin $\models \Box\phi$

Modaalilogiikan aksiomatisoinneissa MP eli modus ponens ja RN eli välttämättömyyssääntö (rule of necessitation) ovat päättelysääntöjä ja skeema K on teoreema.

Muita ominaisuuksia

- Näiden lisäksi logiikkaa täydennetään sopivilla aksioomilla sen mukaan minkälaista modaliteettia ollaan käsittelemässä.
- Esimerkiksi episteemisessä logiikassa mallinnetaan tietoa, joten olisi hyvä jos voimassa olisi aksiooma joka sanoo että jos joku tietää jotakin, sen on oltava totta: $\models K_i\phi \rightarrow \phi$. Saman tulisi olla voimassa myös aleettisessa logiikassa: jos jokin on välttämättä totta, sen on oltava myös totta, siis $\models \Box\phi \rightarrow \phi$.
- Mahdollisten maailmojen semantiikassa aksiooma saadaan päteväksi rajoittamalla Kripke-kehykset sellaisiin, joissa saavutettavuusrelaatiot ovat refleksiivisiä. Siis kunkin agentin i saavutettavuusrelaation R_i on oltava sellainen että $\forall w \in W\ (w, w) \in R_i$.
- Jos sen sijaan tehdään logiikkaa uskomuksista, edellistä skeemaa ei pidä ottaa aksioomaksi eikä uskomuksia vastaavalta saavutettavuusrelaatiolta vaadita refleksiivisyyttä.

Modaalilogiikan korrespondenssiteoriaa

Modaalilogiikan yhteydessä *korrespondenssiteorialla* tarkoitetaan vastaavuutta logiikan aksioomien ja saavutettavuusrelaation ominaisuuksien välillä:

Nimi	Aksiooma	Ominaisuus	Karakterisointi
T	$\Box\phi \rightarrow \phi$	refleksiivisyys	$\forall w \in W\ wRw$
D	$\Box\phi \rightarrow \Diamond\phi$	sarjallisuus	$\forall w \in W\ \exists w' \in W\ wRw'$
B	$\phi \rightarrow \Box\Diamond\phi$	symmetrisyys	$\forall w, w' \in W\ wRw' \Rightarrow w'Rw$
4	$\Box\phi \rightarrow \Box\Box\phi$	transitiivisuus	$\forall w, w', w'' \in W\ wRw' \wedge w'Rw'' \Rightarrow wRw''$
5	$\Diamond\phi \rightarrow \Box\Diamond\phi$	euklidisuus	$\forall w, w', w'' \in W\ wRw' \wedge wRw'' \Rightarrow w'Rw''$

Modaalilogiikan systeemejä: pienin normaali systeemi on nimeltään K. Muita tunnetuimpia ovat T eli KT, S4 eli KT4 ja S5 eli KTB4=KT5.

Tiedon ominaisuuksia

Mitkä suraavista aksioomista tulisi hyväksyä episteemisessä logiikassa?

Distribuutioaksiooma $\models (K_i\phi \wedge K_i(\phi \rightarrow \psi)) \rightarrow K_i\psi$

Tiedon yleistyssääntö jos $M \models \phi$, niin $M \models K_i\phi$.

Totuusaksiooma $\models K_i\phi \rightarrow \phi$

Aksiooma (B) $\models \phi \rightarrow K\neg K\neg\phi$

Positiivinen introspektio (4) $\models K_i\phi \rightarrow K_iK_i\phi$

Negatiivinen introspektio $\models \neg K_i\phi \rightarrow K_i\neg K_i\phi$

Hintikka hyväksyy S4:n, Fagin *et al* S5:n.

Ongelmana kaikissa normaaleissa episteemisissä logiikoissa *loogisen kaikkitietävyyden* (logical omniscience) ongelma

Yleinen ja hajautettu tieto

Yleisen ja hajautetun tiedon käsittelyä varten episteemisen logiikan syntaksiin voidaan vielä lisätä operaattorit E_G, C_G ja D_G, missä G on agenttien joukko.

- $\langle M, w\rangle \models E_G\varphi$, joss $\langle M, w\rangle \models K_i\varphi$ kaikilla $i \in G$
- $\langle M, w\rangle \models C_G\varphi$, joss $\langle M, w\rangle \models E_G^k\varphi$ kaikilla $k = 1, 2, \ldots$
- $\langle M, w\rangle \models D_G\varphi$, joss $\langle M, w'\rangle \models \varphi$ kaikilla w' s.e. $(w, w') \in \cap_{i\in G}R_i$

Yleiselle ja hajautetulle tiedolle saadaan seuraavat ominaisuudet:

$$\models C_G\phi \leftrightarrow E_G(\phi \wedge C_G\phi),$$

$$\text{jos } M \models \phi \rightarrow E_G(\psi \wedge \phi), \text{ niin } M \models \phi \rightarrow C_G\psi,$$

$$\models D_{\{i\}}\phi \leftrightarrow K_i\phi$$

$$\models D_G\phi \rightarrow D_{G'}\phi, \text{ jos } G \subseteq G'.$$

Muita modaalilogiikoita

- *Deonttisessa* logiikassa tarkastellaan pakollisen ja sallitun käsitteitä. Operaattorin □ merkitys on "on pakollista että" ja sitä merkitään tavallisesti O:lla. Duaalioperaattoria merkitään P:llä ja sen merkitys on "on sallittua että". Nykyään suosittu tutkimusalue DAI:ssa.
- *Temporaalilogiikassa* mahdolliset maailmat tulkitaan maailmaksi eri ajanhetkinä. Erilaisia temporaalilogiikoita saadaan sen mukaan ovatko ajanhetket diskreettejä vai jatkuvia, tarkastellaanko vain tulevaisuutta vai myös menneisyyttä ja ymmärretäänkö aika lineaariseksi vai haarautuvaksi. Käytetään usein DAI:ssa ja yleisestikin ohjelmien spesifioinnissa ja verifioinnissa.
- *Dynaaminen* logiikka tarkastelee tietokoneohjelmien suoritusta. Mahdolliset maailmat ymmärretään tietokoneen tiloiksi ja ohjelmat muuttavat tilan toiseksi. Kullakin ohjelmalla α_i on omat operaattorinsa $[\alpha_i]$ ja $\langle\alpha_i\rangle$. Merkintä $[\alpha_i]p$ tarkoittaa että "p on tosi ohjelman α_i kaikissa mahdollisissa lopputiloissa".

Lähteitä

- V. Rantala & A. Virtanen: *Johdatus modaalilogiikkaan.* Gaudeamus, 2004.
- R. Fagin et al.: *Reasoning about Knowledge.* MIT Press, 1995.
- Wooldridge, luku 12.
- Y. Shoham & K. Leyton-Brown: *Introduction to Multi-Agent Systems*, tulossa.