Luku 5

Bayes-verkkojen sovelluksia

Bayes-verkkojen sovellusesimerkkejä: lääketieteellinen diagnosointi, prosessikontrolli, vikadiagnostiikka, tiedon analysointi. Millä ohjelmistoilla Bayes-verkkomalleja voidaan rakentaa (Bayes-verkkokehittimet): kaupalliset ohjelmistot, tutkimusprototyypit.

Bays-verkkosovellukset voidaan jakaa kahteen ryhmään: yksittäiset Bayes-verkkotekniikkoihin perustuvat sovellukset, ja yleiset Bayes-verkkokehittimet, jotka ovat geneerisiä työkaluja erillisten sovellusten tuottamista varten. Bayes-verkkotutkimukseen on koko sen vasta noin kymmenvuotisen historian aikana liittynyt selkeä pyrkimys kehitettyjen menetelmien kaupalliseen soveltamiseen, ja sovellusten määrä onkin viime vuosina ollut voimakkaassa kasvussa. Lisäkiinnostusta Bayes-verkkojen hyödyntämismahdollisuuksia kohtaan herätti Bill Gatesin lokakuussa 1996 antama lausunto, jonka mukaan Microsoftin kilpailuetu tulevaisuudessa perustuu heidän eritysosaamiseensa Bayes-verkkojen alueella [56]. Microsoftilla onkin yksi maailman johtavia Bayes-verkkojen tutkimusyksiköitä (katso luku 6.2.1).

Luvussa 5.1 esitellään joitakin yksittäisiä Bayes-verkkosovelluksia. Luvussa ei pyritä antamaan tyhjentävää luetteloa kaikista olemassaolevista Bayes-verkkosovelluksista, vaan tyypillisiä esimerkkejä Bayes-verkkotekniikoita käyttävistä ohjelmistoista erilaisilta sovellusalueilta. Jotkut mainituista sovelluksista on luotu käyttäen kaupallisia Bayes-verkkokehittimiä, joita käsitellään luvussa 5.2, toiset taas ovat käsityönä luotuja räätälöityjä ohjelmistoja. Lisätietoa Bayes-verkkosovelluksista löytyy mm. David Heckermanin laatimasta listasta[1], johon on koottu yli sata viitettä Bayes-verkkojen käytännön sovelluksia raportoiviin artikkeleihin vuosilta 1985–1995. Heckermanin lista on melko laaja, mutta jo hieman vanhentunut. AUAI-järjestö (Association for

[1] URL: ftp://ftp.research.microsoft.com/pub/dtg/david/BN-APPS.TXT

Uncertainty in Artificial Intelligence) on koonnut suppeahkon listan joistakin merkittävimmistä viimeaikaisista sovelluksista. Lista on saatavilla järjeston kotisivulta[2], mistä löytyy myös kokoelma käytännön soveltajilta kerättyjä kommentteja siitä, miksi Bayes-verkot ovat osoittautuneet onnistuneeksi ratkaisuksi heidän sovellusympäristössään.

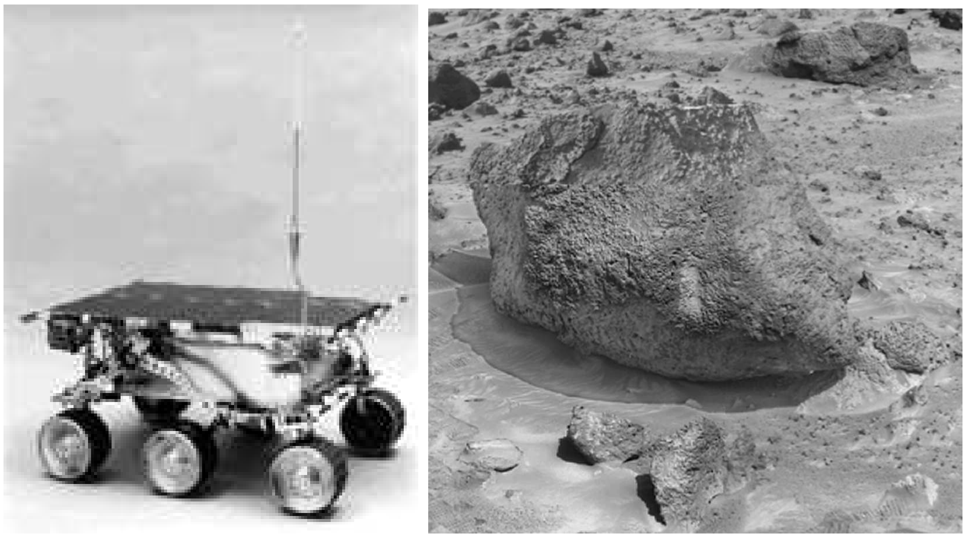
5.1 Sovellusesimerkkejä

Luvussa 4.3 esitettiin, kuinka probabilistisen päättelyn eri muotoja voidaan toteuttaa käytännössä Bayes-verkkomalleja hyväksikäyttäen. Laajat julkisesti saatavilla olevilla luokitteluaineistoilla tehdyt vertailut osoittavat, että bayesiläiset mallit tuottavat johdonmukaisesti erittäin hyviä tuloksia vaihtoehtoisilla menetelmillä (kuten esimerkiksi neuroverkot ja päätöspuut) saavutettuihin tuloksiin verrattuna [138, 75, 140, 76, 82]. Kuten luvussa 2.3 näimme, probabilistista päättelyä voidaan soveltaa luokitteluongelmien lisäksi ratkaisemaan hyvin monentyyppisiä ongelmatilanteita, joten Bayes-verkkojen sovellusalueitakin on lukuisia. Seuraavassa joitakin tyypillisimpiä sovellusesimerkkejä sovellusalueittain luokiteltuna.

Lääketieteellinen diagnosointi. Lääketieteellinen diagnosointi on perinteisesti ollut yksi probabilististen menetelmien tärkeimmistä sovellusalueista. Yksi syy tähän varmaankin on se, että probabilististen menetelmien kanssa on tulosten luotettavuutta mahdollista arvioida tekemällä riskianalyysi luvussa 2.4 esitetyllä tavalla. Yksi ensimmäisiä tällaisia ohjelmistoja oli rintasyöpäpotilaiden diagnosointiin suunniteltu PathFinder [52], joka on myöhemmin kaupallistettu BiopSys Medical-yhtiön[3] IntelliPath-nimisenä tuotteena. IntelliPath-ohjelmisto on tällä hetkellä käytössä sadoissa eri sairaaloissa ympäri maailmaa. Muita vastaavia eri sairaaloissa ja tutkimuslaitoksissa aktiivisessa käytössä olleita järjestelmiä ovat mm. ACORN, DXPLAIN ja Iliad[4]. Yksi suurisuuntaisimpia tämän ongelma-alueen hankkeita on Microsoftin OnParenting-ohjelmisto[5], joka on lasten sairauksien diagnosointiin suunniteltu laaja järjestelmä. Järjestelmä käyttää sairauksien diagnosointiin Knowledge Industries-yhtiön kehittämää Bayes-verkkomallia.

Tiedon analysointi (data mining). Koska Bayes-verkkoja käyttäen voidaan tilastollisesta esimerkkiaineistosta konstruoida ongelmakentästä semant-

[2] URL: http://www.auai.org/
[3] URL: http://www.biopsys.cim
[4] URL: http://www-uk.hpl.hp.com/people/ewc/list-main.html
[5] URL: http://www.onparenting.msn.com



Kuva 5.1: NASA:n Pathfinder-projektin Marsin pinnalle lähettämä Sojouner-robotti, ja sen lähettämistä kuvista superresoluutiotekniikalla tuotettu kuva noin metrin kokoisesta kivenlohkareesta ("Yogi") Marsin pinnalla. Kuva löytyy osoitteesta http://mpfwww.jpl.nasa.gov/mpf/high-res.html.

tisesti tulkittavissa oleva malli, voidaan syntyviä malleja käyttää ongelmakentän ominaisuuksien analysoimisessa ja visualisoimisessa. NASAn Ames-tutkimuskeskuksen AutoClass-projektissa[6] on sovellettu bayesiläistä mallintamista menestyksellisesti mm. IRAS tähtikartaston kuvien, Yhdysvaltojen eri lentokenttien, ja eri DNA-sekvenssien analysoimisessa [24].

Kuvankäsittely. NASAn tutkimusryhmät ovat kehittäneet menetelmän ns. *superresoluutiokuvien (super-resolution images)* tuottamiseksi. Menetelmässä on kyse useiden samasta kohteesta otettujen huonolaatuisten kuvien yhdistämisestä yhdeksi hyvälaatuiseksi ("superresoluutio"-) kuvaksi. Kuvassa 5.1 on tällainen Sojourner-robotin tuottamista kuvista bayesiläisellä tekniikalla muodostettu superresoluutiokuva kivenlohkareesta Marsin pinnalla.

Superresoluutiokuvia tuottava tekniikka ei perustu suoraan Bayes-verkkoteoriaan, mutta se seuraa puhtaasti edellä esitettyjä bayesiläisen mallintamisen periatteita (katso [23]). Paitsi Pathfinder-projektin aineistosta, NASAn tutkijat ovat tuottaneet superresoluutiokuvia myös Galileo-luotaimen aineistosta, sekä Viking-luotaimen Marsin pinnasta ottamista kuvista[7].

[6] URL: http:/ic-www.arc.nasa.gov/ic/projects/bayes-group/autoclass/

[7] URL: http://ic.arc.nasa.gov/ic/projects/super-res/

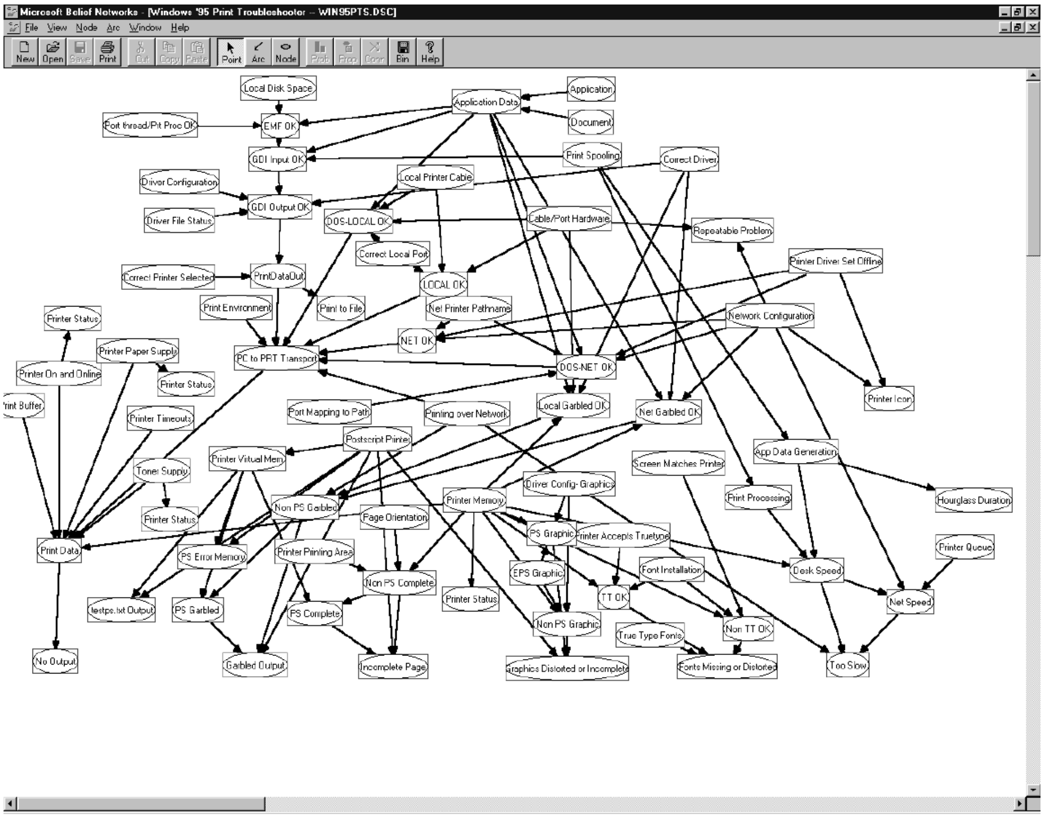
Hahmontunnistus. Bayesiläistä mallinnusta on sovellettu hahmontunnistuksessa menestyksellisesti mm. NASA:n tutkimuskeskuksissa LandSat-satelliittien ottamien kuvien alueiden automaattisessa tunnistamisessa [66]. Käytetty Bayes-verkkomalli oli tässä tapauksessa luvussa 4.5 esitetty äärellinen sekajakaumamalli. Tavanomaisten Bayes-verkkojen käyttöä kuvantunnistuksessa käsitellään mm. lähteissä [63, 43]. Bayes-verkoilla on myös läheinen yhteys ns. *piilo-Markov-malleihin (hidden Markov models)* (katso [131, 30, 111]), joita käytetään yleisesti esimerkiksi puheentunnistukseen liittyvien hahmontunnistusongelmien ratkaisemisessa.

Vikadiagnostiikka. Microsoftin tutkimusosasto on kehittänyt useita kymmeniä Bayes-verkkoihin perustuvia vikadiagnostiikkajärjestelmiä (troubleshooting wizards)[8], jotka auttavat tietokoneen käyttäjää erilaisissa ongelmatilanteissa (ks. kuva5.1). Laitevalmistaja Intel käyttää Bayes-verkkoja prosessoripiirien valmistusprosessin laaduntarkkailussa ja American Airlines lentoyhtiö ohjelmistovirheiden etsimisessä[9].

Prosessikontrolli. Yksi esimerkki prosessikontrolliin liittyvistä sovelluksista on kuvassa 4.4 esitetty ALARM-verkko, jota käytetään sairaalapotilaiden tilaa tarkastelevan mittausprosessin ohjaamisessa [8]. NASA on kehittänyt laajan Bayes-verkkoihin perustuvan VISTA-järjestelmän[10] avaruussukkuloiden lentojen kontrollomiseksi. Jo useita vuosia NASA:n Houstonin avaruussukkuloiden lennonjohtokeskuksessa käytössä ollut järjestelmä paitsi ilmoittaa mahdollisista vikatilanteista, myös ehdottaa päätösteorian mukaisesti parhaat mahdolliset toimenpiteet kriisitilanteissa. Lisäksi ohjelmisto muokkaa aktiivisesti lentojen valvonnassa käytettäviä graafisia näyttöjä siten, että kullakin hetkellä näytetään vain tärkeimmät sukkuloiden tilaa kuvaavat mittaustulokset lentojen kulkua seuraavalle henkilökunnalle, mikä helpottaa teknikkojen työtä, ja nopeuttaa päätöksentekoa ongelmatilanteissa (katso kuva 5.3).

Bayes-verkkoja on sovellettu menestyksellisesti myös mm. lentokoneiden suihkumoottorien suunnittelussa (General Electric[11]), sekä erilaisten asejärjestelmien (Mitre/US Navy[12]) ja miehittämättömien vedenalaisten ajoneuvojen (Lockheed/DARPA[13]) ohjausprosesseissa.

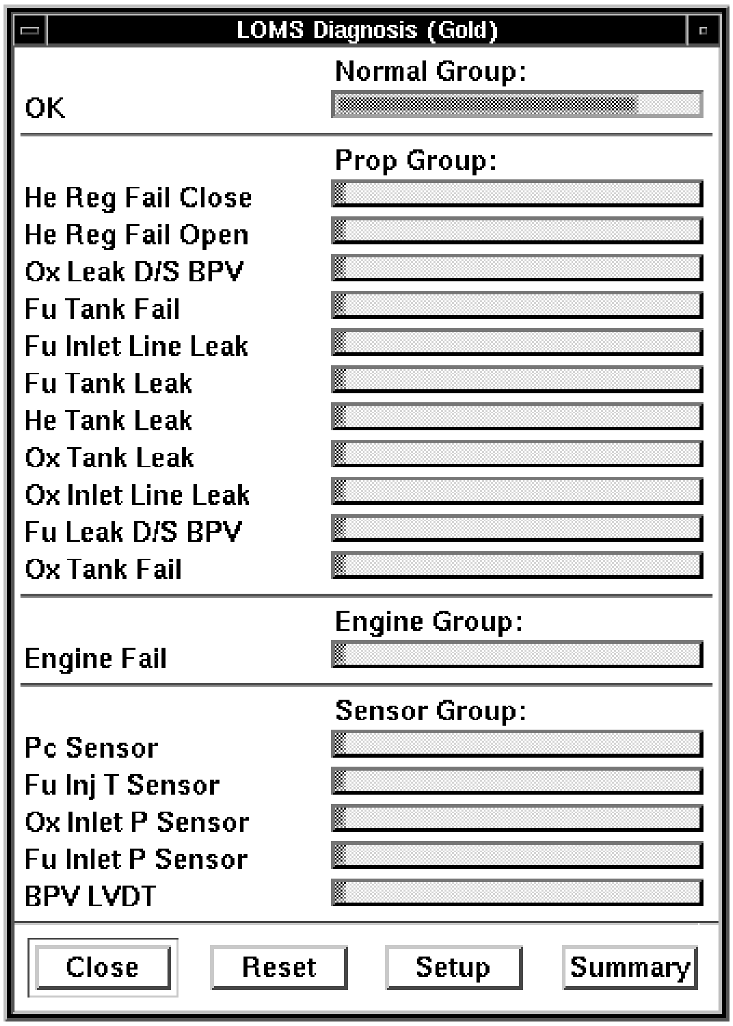
[8] URL: http://support.microsoft.com/support/
[9] URL: http://www.auai.org/BN-Routine.html
[10] URL: http://www.research.microsoft.com/research/dtg/horvitz/vista.htm
[11] URL: http://www3.us.com/kic/kic/
[12] URL: http://www.afit.af.mil/Schools/EN/ENG/LABS/AI/BayesianNetworks/fieldedSystems.html
[13] URL: http://hugin.dk/lockheed-9604.html



Kuva 5.2: Microsoftin kehittämä, Windows-käyttöjärjestelmän käyttämä Bayes-verkko tulostimien vikojen diagnosointiin (printer troubleshooting wizard). Kuva on tuotettu Microsoftin MSBN-ohjelmistolla.

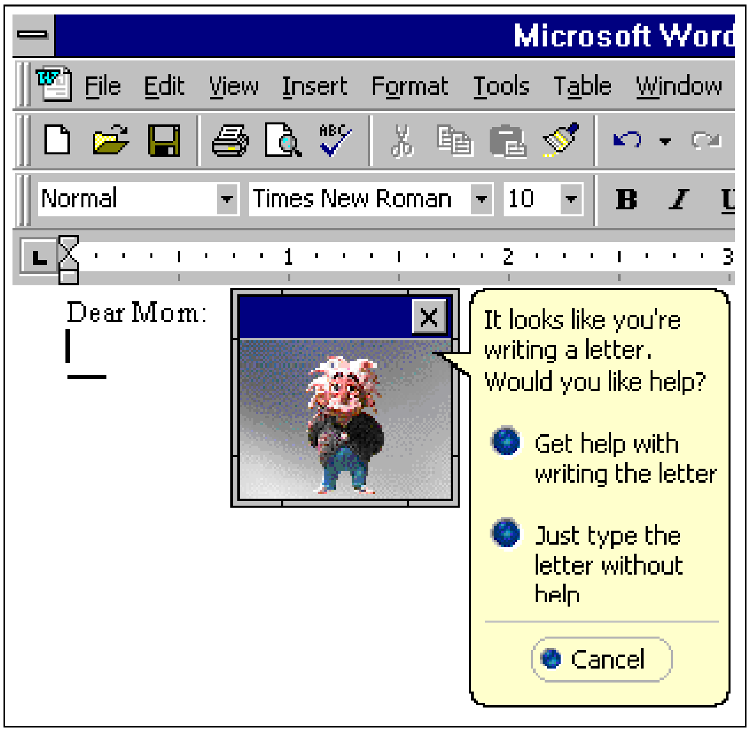
Älykkäät agentit. Älykkäät agentit ovat perussovellusohjelmiin liitettäviä itsenäisiä moduleita, jotka seuraavat järjestelmän toimintaa, ja helpottavat sovellusten käyttöä tarjoamalla aktiivisesti apua silloin kun katsovat sen käyttäjän kannalta tarpelliseksi. Microsoft on soveltanut Bayes-verkkotekniikkaa älykkäiden agenttien toteuttamisessa, ja tekniikka on jo käytössä Office'97-ohjelmistossa *Office Assistant*-nimisenä tuotteena (katso kuva 5.4). Microsoftin pitkän tähtäyksen suunnitelmana on toteuttaa Bayes-verkkotekniikoihin perustuva *adaptiivinen käyttöjärjestelmä*, koodinimeltään *Lumiere*, joka seuraa tietokoneen käyttäjän toimintaa, ja adaptoituu toimimaan siten, että tietokoneen käyttö on mahdollisimman tehokasta. Lisätietoja Lumiere-projektista löytyy projektin WWW-sivulta[14].

[14]URL: http://www.research.microsoft.com/research/dtg/horvitz/lum.htm



Kuva 5.3: Esimerkki NASA:n VISTA-järjestelmään kuuluvista adaptiivisista graafisista käyttöliittymistä, joiden perusteella avaruussukkuloiden valvontakeskus ohjaa lentojen kulkua.

Tiedon kompressointi. Kuten luvussa 3.4 näimme, bayesiläiselle lähestymistavalle voidaan antaa myös informaatioteoreettinen tulkinta todennäköisyysjakaumien ja koodinpituuksien läheisen yhteyden ansiosta. Niinpä ei olekaan yllättävää, että Bayes-verkkoteoriaa voidaan soveltaa myös erilaisten koodausmenetelmien kehittämisessä, ja kehitettyjä koodeja voidaan soveltaa esimerkiksi tiedon kompressoinnissa. Vuonna 1993 esiteltiin uusi vallankumouksellinen koodausmenetelmä, *turbo-koodaus* [11], jonka on empiirisesti havaittu olevan huomattavasti tehokkaampi koodausmenetelmä kuin aikai-



Kuva 5.4: Microsoftin Office Assistant-agentti.

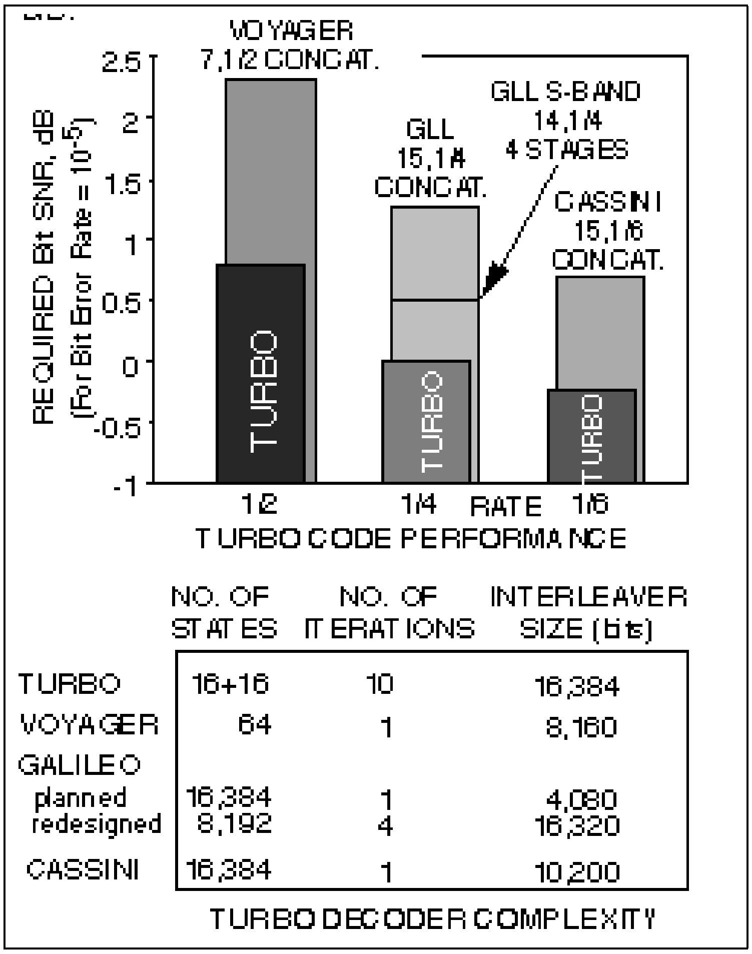
semmin käytetyt koodit (katso kuva 5.5[15]).

Teoksessa [93] osoitetaan kuinka turbo-koodaaja voidaan esittää Bayes-verkkomallina. Turbo-koodausta pidetään yhtenä tärkeimmistä koodausteorian tuottamista tuloksista käytännön sovellusten kannalta. Menetelmä on erityisen kiinnostava avaruustutkimuksen kannalta, koska avaruusluotaimien lähetyskapasiteetti on melko pieni, jolloin tarvitaan tehokkaita tiedon pakkausalgoritmeja.

5.2 Bayes-verkkokehittimet

Koska uusia Bayes-verkko-ohjelmistoja syntyy kiihtyvässä tahdissa, oheinen lista saatavilla olevista Bayes-verkkokehittimistä vanhenee nopeasti. Elektronisesti saatavilla oleva Russell Almondin ylläpitämä lista löytyy Almondin kotisivulta[16], ja toinen vastaavanlainen lista CoSCo-ryhmän kotisivulta[17].

[15] Kuva löytyy WWW-sivulta URL: http://www331.jpl.nasa.gov/public/JPLtcodes.html
[16] URL: http://bayes.stat.washington.edu/almond/belief.html
[17] URL: http://www.cs.Helsinki.FI/research/cosco/



Kuva 5.5: Turbo-koodin vertailu avaruusluotaimissa nykyisin käytössä oleviin koodausmenetelmiin. Huomaa että histogrammikuvan asteikko on logaritminen, joten erot eri menetelmien välillä ovat huomattavia.

Ohjelmistot on seuraavassa jaoteltu kaupallisiin tuotteisiin ja tutkimusprototyyppeihin, mutta on huomattava että joistakin kaupallisista ohjelmistoista on saatavilla ilmainen versio tutkimus- tai opetuskäyttöön. Lisäksi useimmat kaupalliset yritykset tarjoavat tuotteistaan vapaasti kokeiltavia demonstraatioversioita, jotka on saatavilla alla annettujen WWW-linkkien kautta. On myös todettava, että jotkut ilmaiseksi jaettavat, ja siksi tutkimusprototyypeiksi luokitellut ohjelmistot ovat ohjelmistoteknisesti kaupallisten ohjelmistojen tasolla (esim. Microsoftin MSBN).

Alla olevat ohjelmistot on jaoteltu sen mukaan, voidaanko ohjelmistoja käyttää Bayes-verkkojen oppimiseen datasta, vai tukevatko ohjelmat ainoastaan verkkojen manuaalista konstruomista. Oppimista tukevien ohjelmien lukumäärä ei ole toistaiseksi kovin suuri, mutta tilanne tullee muuttumaan lähitulevaisuudessa merkittävästi kun tuoreet aiheeseen liittyvät tutkimustulokset saadaan siirrettyä sovellusasteelle.

5.2.1 Kaupalliset ohjelmistot

Oppimista tukevat ohjelmistot

- Ergo/Cogito
  Lisätietoja: Noetic Systems,
  URL: http://www.noeticsystems.com/ergo.shtml

- MIM
  Lisätietoja: David Edvards,
  URL: http://www.math.auc.dk/~jhb/CoCo/mim.html

- TETRAD
  Lisätietoja: Carnegie-Mellon University,
  URL: http://hss.cmu.edu/HTML/departments/philosophy/TETRAD/ tetrad.html

Manuaaliseen mallin konstruoimiseen perustuvat ohjelmistot

- Baron
  Lisätietoja: KC Associates,
  URL: mailto: kchang@gmu.edu

- Analytica
  Lisätietoja: Lumina Decision Systems,
  URL: http://www.lumina.com/software/index.html

- Decision Tools Suite
  Lisätietoja: Palisade,
  URL: http://www.palisade.com/

- DPL: Decision Programming Language
  Lisätietoja: Applied Decision Analysis,
  URL: http://www.dpl.adainc.com/

- DX Solution Series
  Lisätietoja: Knowledge Industries,
  URL: http://www.kic.com/

- GRAPHICAL-BELIEF
  Lisätietoja: Russell Almond,
  URL: http://bayes.stat.washington.edu/almond/gb/graphical-belief.html

- HUGIN
  Lisätietoja: Hugin Expert,
  URL: http://www.hugin.dk/

- NETICA
  Lisätietoja: Norsys Software,
  URL: http://www.norsys.com/netica.html

- STRATEGIST
  Lisätietoja: Prevision,
  URL: http://www.prevision.com/strategist.html

- DATA
  Lisätietoja: TreeAge Software,
  URL: http://www.treeage.com/

5.2.2 Tutkimusprototyypit

Oppimista tukevat ohjelmistot

- Belief Network Power Constructor
  Lisätietoja: Jie Cheng,
  URL: http://193.61.148.131/jcheng/bnpc.htm

- BIFROST
  Lisätietoja: Bo Thiesson,
  URL: http://www.sp.dk/~sorenh/software.html

- BKD: Bayesian Knowledge Discoverer
  Lisätietoja: Marco Ramoni,
  URL: http://kmi.open.ac.uk/~marco/projects/bkd/software/

- BNG
  Lisätietoja: Peter Haddawy,
  URL: http://www.cs.uwm.edu/public/dsail/research/bng.html

- BUGS
  Lisätietoja: Bugs-projekti,
  URL: http://www.mrc-bsu.cam.ac.uk/bugs/

- COCO
  Lisätietoja: Jens Henrik Badsberg,
  URL: http://www.math.auc.dk/~jhb/CoCo/cocoinfo.html

Manuaaliseen mallin konstruoimiseen perustuvat ohjelmistot

- BAYES
  Lisätietoja: Carnegie-Mellon University,
  URL: http://www.cs.cmu.edu/afs/cs/project/ai-repository/ai/areas/reasonng/probabl/bayes/0.html
- BAYONNET
  Lisätietoja: Yoichi Motomura,
  URL: http://www.etl.go.jp/etl/suri/motomura/BN/bn-java.html
- Bayesian Network Editor
  Lisätietoja: Sreekanth Nagarajan,
  URL: http://eel.engr.orst.edu/~nagarasr/bayesnet/index.html
- BELIEF
  Lisätietoja: Carnegie-Mellon University,
  URL: http://www.cs.cmu.edu/afs/cs/project/ai-repository/ai/areas/reasonng/probabl/belief/0.html
- IDEAL
  Lisätietoja: Rockwell International Science Center,
  URL: http://rpal.rockwell.com/ideal.html
  URL: mailto: ideal-request@rpal.rockwell.com
- JAVABAYES
  Lisätietoja: Fabio Gozman,
  URL: http://www.cs.cmu.edu/~javabayes/Home/
- MacEvidence
  Lisätietoja: Prakash Shenoy,
  URL: http://stat1.cc.ukans.edu:80/~pshenoy/
- Microsoft Bayes Networks (MSBN)
  Lisätietoja: Microsoft Research,
  URL: http://www.research.microsoft.com/research/dtg/msbn/
- Pulcinella
  Lisätietoja: IRIDIA,
  URL: http://iridia.ulb.ac.be/pulcinella/Welcome.html

- S-ElimBel
  Lisätietoja: Nicholas Thiéry,
  URL: http://www.spaces.uci.edu/thiery/elimbel/
- SPI
  Lisätietoja: Bruce D'Ambrosio,
  URL: http://www.cs.orst.edu/~dambrosi/
- TresBel
  Lisätietoja: Hong Xu,
  URL: http://iridia.ulb.ac.be/Projects/imple.html
- XBaies
  Lisätietoja: Robert Cowell,
  URL: mailto:xbaies-list@stats.ucl.ac.uk
- X-pert
  Lisätietoja: AI Group, Cantabria University,
  URL: http://ccaix3.unican.es/~AIGroup/
- XPress
  Lisätietoja: Mark Stitson,
  URL: mailto:M.Stitson@rhbnc.ac.uk

Yllä olevaan listaan on otettu mukaan vain standardeja Bayes-verkkoja käyttävät sovellukset. Sekajakaumamalleja käyttävistä ohjelmistosta tunnetuimpia ovat:

- AutoClass
  Lisätietoja: NASAn AutoClass-projekti,
  URL: http://ic-www.arc.nasa.gov:80/ic/projects/bayes-group/group/autoclass/
- Snob
  Lisätietoja: Snob-kotisivu,
  URL: http://www.cs.monash.edu.au/~dld/Snob.html
- Bayda
  Lisätietoja: Complex Systems Computation Group (CoSCo),
  URL: http://www.cs.Helsinki.FI/research/cosco/

Sekä AutoClass että Snob ovat tutkimusprototyyppejä, joiden toiminta perustuu latentteja muuttujia sisältävien mallien bayesiläiseen oppimiseen datasta. Baydassa opittava malli on odotusarvoparametreja käyttävä

Naiivi Bayes-malli, johon on liitetty automaattinen irrelevanttien muuttujien karsinta-algoritmi. Lisätietoja sekajakaumamallien kaupallisista sovelluksista, ja muista tutkimusprototyypeistä löytyy David Dowen ylläpitämältä WWW-sivulta[18].

Tilastollisista sovelluksistaan tunnettu SPSS-yhtiö on myös ryhtynyt soveltamaan bayesiläisiä tekniikoita kaupallisissa ohjelmistoissaan. Käytettävä malliperhe ei tosin ole Bayes-verkot, vaan neuroverkkomallit. Yhtiön mainoksen mukaan bayesiläisen lähestymistavan avulla voidaan välttää ylioppiminen, ja mallien oppiminen on myös myös yleensä nopeampaa. Lisätietoja bayesiläisiä neuroverkkoja soveltavasta "Neural Connection 2"-tuotteesta löytyy WWW-sivulta URL: http://www.spss.com/software/Neuro/. Vastaavanlainen vapaasti saatavilla olevilla ohjelmisto on raportoitu Radford Nealin väitöskirjassa [99], ja on saatavilla Nealin kotisivulta[19]. Nealin ohjelmistoa on laajennettu hiljakkoin käsittämään neuroverkkojen lisäksi myös äärellisiä sekajakaumamalleja.

[18] URL: http://www.cs.monash.edu.au/~dld/mixture.modelling.page.html
[19] URL: http://www.cs.utoronto.ca/~radford/

Luku 6

Tietolähteitä

Bayes-verkkojen aktiiviset kaupalliset soveltajat ja niitä markkinoivat yritykset. Bayes-verkkoja koskevaa kirjallisuutta. Tutkimusalan tutkimusryhmiä ja järjestöjä.

6.1 Kaupallinen sektori

Bayes-verkkojen kaupallistamisen pioneereja on tanskalainen HUGIN-yhtiö, joka soveltaa Aalborgin yliopistossa toimivan, professori Steffen Lauritzenin johtaman tutkimusryhmän tuloksia kaupallisen Hugin-ohjelmiston kehittämisessä. Kuten Bill Gatesin antamasta lausunnosta [56] saattaa arvata, merkittävin kaupallinen tekijä Bayes-verkkojen sovellusalueella on tällä hetkellä eittämättä Microsoft, jonka Bayes-verkkotutkimusryhmä on tutkimusalan suurimpia.

HUGIN-yhtiö on keskittynyt kehittämään ja myymään yhtä tuotetta, Hugin-kehitintä, kun taas Microsoft tuottaa sisäisiä sovelluksia, jotka liitetään yhtiön kaupallisiin tuotteisiin. Monet Bayes-verkkosovelluksia tekevistä yhtiöistä toimivat HUGIN-yhtiön tapaan tuottaen geneerisiä Bayes-verkkokehittimiä, mutta yhä useammat yhtiöt laajentavat toimintaansa myös konsultointialueelle, ja auttavat asiakkaitaan rakentamaan laajoihin järjestelmiin integroitavia räätälöityjä sovelluksia, samalla tavoin kuin Microsoft liittää oman tutkimuskeskuksensa tulokset tuotteisiinsa. Esimerkkejä tällaisista räätälöidyistä sovelluksista löytyy AUAI-järjestön kotivulta[1], samoin kuin lista niistä syistä jotka ovat kussakin tapauksessa johtaneet Bayes-verkkojen käyttöönottoon[2].

[1] URL: http://www.auai.org/BN-Routine.html
[2] URL: http://www.auai.org/BN-Testimonial.html

Bayes-verkkotekniikoiden aktiivisiin kaupallisiin soveltajiin kuuluvat tällä hetkellä seuraavat yritykset:

- Applied Decision Analysis, URL: http://www.dpl.adainc.com/
- Decisioneering, URL: http://www.decisioneering.com/
- Hugin Expert, URL: http://www.hugin.dk/
- Knowledge Industries, URL: http://www.kic.com/
- Lumina Decision Systems, URL: http://www.lumina.com/
- Microsoft, URL: http://www.research.microsoft.com/research/dtg/
- Noetic Systems, URL: http://www.noeticsystems.com/
- Norsys Software, URL: http://www.norsys.com/home.html
- Palisade, URL: http://www.palisade.com/
- Prevision, URL: http://www.prevision.com/
- ThinkBank, URL: http://www.Thinkbank.COM/
- TreeAge Software, URL: http://www.treeage.com/
- Ultimode Systems, URL: http://www.ultimode.COM/index.html

6.2 Bayes-verkkotutkimus

6.2.1 Tutkimusryhmiä

Yksi merkittävimmistä Bayes-verkkotutkimusryhmistä on perinteisesti ollut professori Steffen Lauritzenin johtama tutkimusryhmä Aalborgin yliopistossa Tanskassa. Ryhmä on osallistunut aktiivisesti myös HUGIN-ohjelmistotuotteen (ks. luku 5.2) rakentamiseen. Muita merkittäviä tutkimusryhmiä Euroopassa löytyy mm. Suomesta, Belgiasta ja Espanjasta (ks. allaoleva lista).

Maailmanlaajuisesti suurin tutkimusryhmä tällä hetkellä on luultavasti Microsoftilla, joka on koonnut koko 90-luvun ajan lupaavia Bayes-verkkotutkijoita tutkimuskeskukseensa. Microsoftin asema lienee vain vahvistumassa tulevaisuudessa, koska tutkimusryhmä on tuottanut jo useita menestyksellisiä Bayes-verkkosovelluksia (ks. luku 5.1), ja yhtiö on lisäksi lisäämässä tutkimusrahoitustaan moninkertaiseksi nykyiseen verrattuna. Yrityksen pääjohtajan Bill Gatesin lausunnot [56] kuvaavat myös yhtiön vahvaa luottamusta Bayes-verkkotekniikkaan. Toinen merkittävä tekijä Bayes-verkkojen tutkimusalueella Yhdysvalloissa on NASA, jolla on ollut lukuisia menestyksellisiä Bayes-verkkoprojekteja. Monash-yliopiston tutkimusryhmä Australiassa on tehnyt pioneerityötä sekajakaumamallien tutkimuksen alueella, samoin kuin CoSCo-ryhmä Helsingin yliopistossa.

Alla oleva lista sisältää joitakin WWW-linkkejä Bayes-verkkotutkimusta tekeviin ryhmiin Suomessa, Euroopassa, ja muualla maailmassa. Lista sisältää luonnollisesti vain muutaman esimerkin Bayes-verkkotutkimusta tekevistä tutkimusryhmistä; lisää hyödyllisiä WWW-linkkejä (mm. yli 100 nimeä käsittävä luettelo bayesiläistä mallintamista tutkivista henkilöistä[3]) löytyy AUAI-järjestön kotisivulta (ks. luku 6.2.4).

- Helsingin yliopisto
  URL: http://www.cs.Helsinki.FI/research/cosco/
- Aalborgin yliopisto
  URL: http://www.cs.auc.dk/general/ES/
- Cantabrian yliopisto/Cornell University
  URL: http://ccaix3.unican.es/~AIGroup/
- IRIDIA
  URL: http://iridia.ulb.ac.be/Projects/imple.html
- NASA
  URL: http://fi-www.arc.nasa.gov/fia/projects/bayes-group/
- Microsoft
  URL: http://www.research.microsoft.com/research/dtg/
- Monash University
  URL: http://www.cs.monash.edu.au/~dld/

6.2.2 Kirjallisuutta

Bayes-verkkokirjallisuuden klassikkona pidetään Pearlin vuonna 1988 julkaisemaa kirjaa [108], joka vielä hieman vanhentuneenakin sopii hyväksi johdatukseksi Bayes-verkkoteoriaan. Pearlin lähestymistapa perustuu pitkälti verkkoteoreettisiin käsitteisiin, kun taas lähes samaan aikaan julkaistu Neapolitanin oppikirja [100] tarjoaa matemaattisesti hieman pragmaattisemman,

[3] URL: http://bayes.stat.washington.edu/bayes_people.html

asiantuntijajärjestelmien rakentamiseen tähtäävän yleisesityksen. Charniakin artikkeli [19] antaa selkeän, ei-matemaattisen johdatuksen Bayes-verkkojen peruskäsitteisiin, samoin kuin artikkeli [55] ja muut samassa CACM-lehden erikoisnumerossa ilmestyneet artikkelit. Hieman teknisempi yleiskatsaus aiheeseen löytyy lähteestä [134].

Edellä mainitut artikkelit, samoin kuin Pearlin ja Neapolitanin kirjat, ovat hieman vanhentuneita mm. siinä mielessä, että niissä ei juurikaan käsitellä Bayes-verkkojen oppimista. Uudempia johdatuksia Bayes-verkkoteoriaan tarjoavat Spirtesin et al. [135], Jensenin [62], Schaferin [123] ja Castillon et al. [18] kirjat, sekä Heckermanin artikkeli [50]. Spirtesin kirja käsittelee kausaalisia Bayes-verkkoja, ja niiden ei-bayesiläistä oppimista. Jensenin ja Schaferin kirjat keskittyvät kuvaamaan probabilistisen päättelyn toteuttamista Bayes-verkoissa, eivätkä käsittele juuri lainkaan Bayes-verkkojen oppimista. Castillon kirja on selkeää luettavaa, mutta Bayes-verkkojen oppimista käsittelevä luku sisältää joitakin epätarkkuuksia. Heckermanin artikkeli on hyvä yleiskatsaus Bayes-verkkotutkimuksen tuoreimpiin tuloksiin. Lauritzenin [86] ja Whittakerin [146] kirjat käsittelevät yleisemmin graafisiin verkkoesityksiin perustuvan mallintamisen teoriaa, josta Bayes-verkkotutkimus on vain yksi osa-alue.

Bayes-verkkojen oppimiseen keskittyvää materiaalia löytyy lähteistä [51, 49, 16, 17]. Suomenkielisiä katsauksia tutkimusalueeseen löytyy tutkielmista [127, 5]. Yleisemmin bayesiläistä mallinmuodostusta käsitellään lähteissä [138, 22, 90, 99]. Näistä Mackay [90] ja Neal [99] keskittyvät bayesiläisen lähestymistavan soveltamiseen neuroverkkojen oppimisessa. Muita neuroverkkojen ja bayesiläisen mallintamisen välisiä yhteyksiä käsitteleviä teoksia ovat [12, 112, 14, 96, 122].

Teoreettisempia katsauksia todennäköisyyslaskennan soveltamistekniikoihin löytyy teoksista [42, 10, 61]. Hyviä johdatuksia päätösteoriaan löytyy lähteistä [32, 9, 39, 88]. Epätäsmällisyyden käsitettä yleisemmin käsitellään teoksissa [124, 105, 143]. Todennäköisyyslaskennan ja sumean logiikan suhdetta tarkastellaan artikkeleissa [21, 20].

Lisää viitteitä Bayes-verkkoihin liittyviin teoksiin löytyy seuraavilta WWW-sivuilta:

- URL: http://www.cs.pitt.edu/~tsamard/bnpointers.html
- URL: http://www.cs.cmu.edu/afs/cs.cmu.edu/user/ldc/WWW/bayes-net-research/bayes-net-research.html
- URL: http://www.Ultimode.com/~wray/graphbib/
- URL: http://www.lis.pitt.edu/~dsl/da-books.html

Elektronisesti luettavia (HTML-formaatissa olevia) johdatuksia Bayes-verkkoteoriaan löytyy mm. seuraavilta WWW-sivuilta:

- URL: http://www.afit.af.mil/Schools/EN/ENG/LABS/AI/BayesianNetworks/
- URL: http://www.cm.cf.ac.uk/Dave/AI2/AI_notes.html
- URL: http://www.research.att.com/~dmac/notes/bayesian.html
- URL: http://yoda.cis.temple.edu:8080/UGAIWWW/lectures97/uncertainty/bindex.html

Lisäksi useiden Bayes-verkkotutoriaalien/kurssien kalvot ovat myös saatavilla elektronisessa muodossa. Näitä löytyy mm. seuraavilta WWW-sivuilta:

- URL: http://www.cs.Helsinki.FI/~myllymak
- URL: http://www.cs.Helsinki.FI/~tirri
- URL: http://www.research.microsoft.com/research/dtg/heckerma/heckerma.html
- URL: http://www.ultimode.COM/~wray/
- URL: http://www.auai.org/auai-tutes.html

6.2.3 Konferensseja ja lehtiä

Yksinomaan Bayes-verkkotutkimukseen keskittynyttä jokavuotista tieteellistä konferenssia ei toistaiseksi ole perustettu, joskin *Uncertainty in Artificial Intelligence (UAI)*-konferenssit ovat käytännössä muodostuneet sellaiseksi — lähes kaikki konferenssissa nykyään julkaistavat artikkelit käsittelevät Bayes-verkkoja. Bayesiläiseen mallinnukseen yleisesti liittyviä asioita käsitteleviä konferensseja ovat mm. *International Workshop on Artificial Intelligence and Statistics*, *Valencia International Meeting on Bayesian Statistics*, *International FLAIRS Conference (Special Track on Uncertain Reasoning)*, ja *International Workshop on Maximum Entropy and Bayesian Methods*. Koska bayesiläinen mallinnus liittyy läheisesti koneoppimiseen ja muihin tekoälyn osa-alueisiim, julkaistaan alueeseen liittyviä artikkeleita paljon mm. seuraavissa konferensseissa: *Knowledge Discovery and Data mining (KDD)*, *International Conference on Machine Learning (ICML)*, *International Joint Conference on Artificial Intelligence (IJCAI)*, ja *National Conference on Artificial Intelligence (AAAI)*. Neuroverkkokonferensseista ovat jokavuotiset *Neural Information Processing Systems (NIPS)*-konferenssit julkaisseet yhä kasvavassa määrin bayesiläiseen mallinnukseen liittyviä artikkeleita. Eurooppalaiset epätäsmällisen tiedon käsittelyyn liittyvät konferenssit, kuten esimerkiksi *International Joint Conference on Qualitative and Quantitative Practical Reasoning (ECSQARU)*, *European Symposium on Intelligent Techniques*, ja *European Congress on Intelligent Techniques and Soft Computing (EUFIT)*, ovat keskittyneet pääasiassa ei-bayesiläisten lähestymistapojen käsittelemiseen.

Bayes-verkkotutkijat eivät toistaiseksi ole perustaneet omaa, Bayes-verkkotutkimukselle omistautunutta tieteellistä julkaisua, vaan alan tutkimustulokset ilmestyvät monissa eri tekoälyn osa-alueita käsittelevissä lehdissä, joita ovat mm. *Artificial Intelligence*, *Journal of Artificial Intelligence Research*, *Journal of Approximate Reasoning* ja *Machine Learning*.

6.2.4 Järjestöjä

Koska edellä mainittu UAI-konferenssi on muodostunut Bayes-verkkotutkimusalueen tärkeimmäksi vuosittaiseksi tapaamiseksi, on konferenssin järjestävästä *Association for Uncertainty in Artificial Intelligence (AUAI)*-organisaatiosta muovautunut eräänlainen Bayes-verkkotutkijoiden kattojärjestö. Muita merkittäviä alan tutkimukseen liittyviä järjestöjä ovat *International Society for Bayesian Analysis (ISBA)*, joka sponsoroi *Valencia International Meeting on Bayesian Statistics*- ja *International Workshop on Maximum Entropy and Bayesian Methods*-kokouksia, sekä *Society for Artificial Intelligence and Statistics*, joka järjestää joka toinen vuosi pidettävän *International Workshop on Artificial Intelligence and Statistics*-kokouksen. Muita tutkimusalueeseen liittyviä järjestöjä on lueteltu alla. Eurooppalaisista organisaatioista on huomautettava, että kuten eurooppalaiset alan konferenssit, myös tieteelliset organisaatiot (kuten esim. ERUDIT, NeuroCOLT, ja ML-net) ovat keskittyneet pääasiassa ei-bayesiläisiin epätäsmällisyyden käsittelytapoihin.

- Association for Uncertainty in Artificial Intelligence (AUAI),
  URL: http://www.auai.org/
- Society for Artificial Intelligence and Statistics,
  URL: http://www.vuse.vanderbilt.edu/~dfisher/ai-stats/society.html
- International Society for Bayesian Analysis (ISBA)
  URL: http://omega.albany.edu:8008/isba/
- ASA Section on Bayesian Statistical Sciences
  URL: http://www.isds.duke.edu/sbss/sbss.html

- European Association for Decision Making
  URL: http://huizen.dds.nl/~eadm/
- Decision Analysis Society
  URL: http://www.fuqua.duke.edu/faculty/daweb/
- Society for Risk Analysis (SRA)
  URL: http://www.sra.org/
- Classification Society of North America (CSNA)
  URL: http://www.pitt.edu/~csna/
- Network of Excellence for Uncertainty Modeling and Fuzzy Technology in the European Union (ERUDIT)
  URL: http://ss-m3.mitgmbh.de/erudit/
- ESPRIT Working Group in Neural and Computational Learning (NeuroCOLT)
  URL: http://www.cs.rhbnc.ac.uk/research/compint/neurocolt/
- ESPRIT Network of Excellence in Machine Learning (MLnet)
  URL: http://omega.gmd.de/ml-archive/MLnet/MLnet.html

Kirjallisuutta

[1] Aha, D., Ed. *Lazy Learning.* Kluwer Academic Publishers, Dordrecht, 1997. Reprinted from Artificial Intelligence Review, 11:1–5.

[2] Aha, D., Kibler, D., and Albert, M. Instance-based learning algorithms. *Machine Learning 6* (1991), 37–66.

[3] Alag, S., and Agogino, A. Inference using message propagation and topology transformation in vector gaussian continuous networks. In *Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence (UAI–96)* (Portland, Oregon, 1996), pp. 20–27.

[4] Amari, S., Murata, N., Müller, K.-R., Finker, M., and Hua Yang, H. Asymptotic statistical theory of overtraining and cross-validation. *IEEE Transactions on Neural Networks 8*, 5 (February 1997), 985–996.

[5] Antti-Poika, T. Bayes-verkkojen käyttö luokittelussa. Master's thesis, Report C-1997-68, Deparment of Computer Science, University of Helsinki, 1997.

[6] Baxter, R., and Oliver, J. MDL and MML: Similarities and differences. Tech. Rep. 207, Department of Computer Science, Monash University, 1994.

[7] Becker, A., and Geiger, D. A sufficiently fast algorithm for finding close to optimal junction trees. In *Proceedings of the Twelfth Annual Conference on Uncertainty in Artificial Intelligence (UAI–96)* (Portland, Oregon, 1996), pp. 81–89.

[8] Beinlich, I., Suermondt, H., Chavez, R., and Cooper, G. The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks. In *Proceedings of the Second European Conf. on Artificial Intelligence in Medicine* (London, UK, August 1989), pp. 247–256.

[9] Berger, J. *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, New York, 1985.

[10] Bernardo, J., and Smith, A. *Bayesian theory.* John Wiley, 1994.

[11] Berrou, G., Glavieux, A., and Thitimajshima, P. Near Shannon limit error-correcting coding: Turbo codes. In *Proceedings of the 1993 International Conference in Communications* (geneve, May 1993), pp. 1064–1070.

[12] Bishop, C. *Neural Networks for Pattern Recognition.* Oxford University Press, 1995.

[13] Bishop, C., Svensén, M., and Williams, C. EM optimization of latent-variable density models. In *Advances in Neural Information Processing Systems 8*, D. Touretzky, M.C.Mozer, and M.E.Hasselmo, Eds. MIT Press, 1996.

[14] Bishop, C., Svensén, M., and Williams, C. GTM: A principled alternative to the self-organizing map. In *Advances in Neural Information Processing Systems 9, Proceedings of the 1996 Conference*, M. Mozer, M. Jordan, and J. Petsche, Eds. MIT Press, 1997.

[15] Breiman, L. Bias, variance, and arcing classifiers. Tech. Rep. TR 460, University of California, Statistics Department, April 1996.

[16] Buntine, W. Operations for learning with graphical models. *Journal of Artificial Intelligence Research 2* (1994), 159–225.

[17] Buntine, W. A guide to the literature on learning graphical models. *IEEE Transactions on Knowledge and Data Engineering 8* (1996), 195–210.

[18] Castillo, E., Gutiérrez, J., and Hadi, A. *Expert Systems and Probabilistic Network Models.* Monographs in Computer Science. Springer-Verlag, New York, NY, 1997.

[19] Charniak, E. Bayesian networks without tears. *AI Magazine 12*, 4 (1991), 50–63.

[20] Cheeseman, P. Probabilistic versus fuzzy reasoning. In Kanal and Lemmer [65], pp. 85–102.

[21] Cheeseman, P. In defense of probability. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence* (1995), vol. 2, pp. 1002–1009.

[22] Cheeseman, P. On Bayesian model selection. In *The Mathematics of Generalization*, D. Wolpert, Ed., vol. XX of *SFI Studies in the Sciences of Complexity*. Addison-Wesley, 1995, pp. 315–330.

[23] Cheeseman, P., Kanefsky, B., Hanson, R., and Stutz, J. Super-resolved surface reconstruction from multiple images. Tech. Rep. FIA-94-12, NASA Ames Research Center, Artificial Intelligence Branch, October 1994.

[24] Cheeseman, P., and Stutz, J. Bayesian classification (AutoClass): Theory and results. In *Advances in Knowledge Discovery and Data Mining*, U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, Eds. AAAI Press, Menlo Park, 1996, ch. 6.

[25] Chickering, D., Geiger, D., and Heckerman, D. Learning Bayesian networks is NP-hard. Tech. Rep. MSR-TR-94-17, Microsoft Research, 1994.

[26] Chickering, D., and Heckerman, D. Efficient approximations for the marginal likelihood of incomplete data given a Bayesian network. In *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence* (Portland, Oregon, August 1996), E. Horvitz and F. Jensen, Eds., Morgan Kaufmann Publishers, pp. 158–168.

[27] Cooper, G. The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence 42*, 2–3 (March 1990), 393–405.

[28] Cooper, G., and Herskovits, E. A Bayesian method for the induction of probabilistic networks from data. *Machine Learning 9* (1992), 309–347.

[29] Dagum, P., and Luby, M. Approximating probabilistic inference in Bayesian belief networks is NP-hard. *Artificial Intelligence 60* (1993), 141–153.

[30] Dean, T., and Wellman, M. *Planning and Control.* Morgan Kaufmann Publishers, San Mateo, CA, 1991.

[31] Dechter, R., and Rish, I. A scheme for approximating probabilistic inference. In *Proceedings of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI–97)* (San Francisco, CA, 1997), Morgan Kaufmann Publishers, pp. 132–141.

[32] DeGroot, M. *Optimal statistical decisions.* McGraw-Hill, 1970.

[33] Dempster, A., Laird, N., and Rubin, D. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B 39*, 1 (1977), 1–38.

[34] Druzdzel, M., and Henrion, M. Efficient reasoning in qualitative probabilistic networks. In *Proceedings of the Eleventh National Conference on Artificial Intelligence* (Washington, DC, July 1993), AAAI Press/MIT, Menlo Park, CA, pp. 548–553.

[35] Everitt, B., and Hand, D. *Finite Mixture Distributions.* Chapman and Hall, London, 1981.

[36] Friedman, J. On bias, variance, 0/1–loss, and the curse-of-dimensionality. *Data Mining and Knowledge Discovery 1*, 1 (1997), 55–78.

[37] Friedman, N., and Goldszmidt, M. Discretizing continuous attributes while learning Bayesian networks. In *Machine Learning: Proceedings of the Thirteenth International Conference* (1996), L. Saitta, Ed., Morgan Kaufmann Publishers, pp. 157–165.

[38] Friedman, N., Goldszmidt, M., Heckeman, D., and Russel, S. What is the impact of bayesian networks on learning? In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence* (Nagoya,Japan, August 1997), Morgan Kaufmann Publishers, pp. 10–15.

[39] Gärdenfors, P., and Sahlin, N.-E., Eds. *Decision, Probability, and Utility.* Cambridge University Press, New York, 1988.

[40] Geiger, D., and Heckerman, D. A characterization of the Dirichlet distribution through global and local independence. Tech. Rep. MSR-TR-94-16, Microsoft Research, November (revised February 1995) 1994.

[41] Geiger, D., and Heckerman, D. Learning bayesian networks. Tech. Rep. MSR-TR-95-02, Microsoft Research, December 1994.

[42] Gelman, A., Carlin, J., Stern, H., and Rubin, D. *Bayesian Data Analysis.* Chapman & Hall, 1995.

[43] Gong, S., and Buxton, H. Bayesian nets for mapping contextual knowledge to computational constraints in motion segmentation and tracking. In *British machine vision conference* (Guildford, Surrey, UK, September 1993).

[44] Gyllenberg, M., Koski, T., and Verlaan, M. Classification of binary vectors by stochastic complexity. Tech. Rep. A5, University of Turku, Institute for Applied Mathematics, November 1994.

[45] Hassoun, M. *Fundamentals of Artificial Neural Networks.* MIT Press, Cambridge, Massachusetts, 1995.

[46] Hecht-Nielsen, R. *Neurocomputing.* Addison-Wesley Publishing Company, Reading, MA, 1990.

[47] Heckerman, D. Probabilistic interpretation for MYCIN's certainty factors. In Kanal and Lemmer [65], pp. 167–196.

[48] Heckerman, D. A Bayesian approach to learning causal network. In *Uncertainty in Artificial Intelligence: Proceedings of the Eleventh Conference* (Montreal, Canada, 1995), P. Besnard and S. Hanks, Eds., pp. 285–295.

[49] Heckerman, D. A tutorial on learning with Bayesian networks. Tech. Rep. MSR-TR-95-06, Microsoft Research, Advanced Technology Division, One Microsoft Way, Redmond, WA 98052, 1996.

[50] Heckerman, D. Bayesian networks for data mining. *Data Mining and Knowledge Discovery 1*, 1 (1997), 79–119.

[51] Heckerman, D., Geiger, D., and Chickering, D. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning 20*, 3 (September 1995), 197–243.

[52] Heckerman, D., and Nathwani, B. An evaluation of the diagnostic accuracy of Pathfinder. *Computers and Biomedical Research 25* (1992), 56–74.

[53] Heckerman, D., and Schachter, R. Decision-theoretic foundations for causal reasoning. *Journal of Artificial Intelligence Research 3* (1995), 405–430.

[54] Heckerman, D., and Shwe, M. Diagnosis of multiple faults: A sensitivity analysis. In *Uncertainty in Artificial Intelligence 9*, D. Heckerman and A. Mamdani, Eds. Morgan Kaufmann Publishers, San Mateo, CA, 1993, pp. 80–87.

[55] Heckerman, D., and Wellman, M. Bayesian networks. *Communications of the ACM 38*, 3 (1995), 27–30.

[56] Helm, L. The future of software may lie in the obscure theories of an 18th century cleric called Thomas Bayes. *Los Angeles Times* (Monday, October 28 1996).

[57] Henrion, M. Uncertainty in artificial intelligence: Is probability epistemologically and heuristically adequate? In *Expert Judgements and Expert Systems*, J. Mumpower, L. Philipps, O. Renn, and U. V., Eds., NATO ASI Series F: Computer and Systems Science. Springer-Verlar, Berlin, 1987, pp. 106–129.

[58] Henrion, M. An introduction to algorithms for inference in belief nets. In *Uncertainty in Artificial Intelligence 5*, M. Henrion, R. Shachter, L. Kanal, and J. Lemmer, Eds. Elsevier Science Publishers B.V. (North-Holland), Amsterdam, 1990, pp. 129–138.

[59] Henrion, M., Pradhan, M., Del Favero, B., Huang, K., Provan, G., and O'Rorke, P. Why is diagnosis using belief networks insensitive to imprecision in probabilities? In *Uncertainty in Artificial Intelligence, Proceedings of the Twelfth Conference* (Portland, Oregon, August 1996), E. Horvits and F. Jensen, Eds., Morgan Kaufmann Publishers, San Francisco, CA, pp. 307–314.

[60] Isomursu, P., Niskanen, V., Carlsson, C., and Eklund, P. Sumean logiikan mahdollisuudet. Tech. Rep. 34/93, Teknologian kehittämiskeskus (TEKES), 1993.

[61] Jaynes, E. Probability theory: The logic of science. Under construction. A draft copy can be dowloaded from 'ftp://bayes.wustl.edu/Jaynes.book'., 1996.

[62] Jensen, F. *An Introduction to Bayesian Networks.* UCL Press, London, 1996.

[63] Jensen, F., Christensen, H., and Nielsen, J. Bayesian methods for interpretation and control in multi-agent vision systems. *Applications of Artificial Intelligence X: Machine Vision and Robotics, SPIE Proceedings Series 1708* (1992).

[64] Jordan, M., and Jacobs, R. Hierarchical mixtures of experts and the EM algorithm. *Neural Computation 6* (1994), 181–214.

[65] Kanal, L., and Lemmer, J., Eds. *Uncertainty in Artificial Intelligence 1.* Elsevier Science Publishers B.V. (North-Holland), Amsterdam, 1986.

[66] Kanefsky, B., Stutz, J., Cheeseman, P., and Taylor, W. An improved automatic classification of a Landsat/TM image from Kansas (FIFE). Tech. Rep. FIA-94-01, NASA Ames Research Center, Artificial Intelligence Branch, May 1994.

[67] Kass, R., and Raftery, A. Bayes factors. Tech. Rep. 254, Department of Statistics, University of Washington, 1994.

[68] Kjærulff, U. Optimal decomposition of probabilistic networks by simulated annealing. *Statistics and Computing 2* (1992), 7–17.

[69] Kohonen, T. *Self-Organizing Maps.* Springer-Verlag, Berlin, 1995.

[70] Kolodner, J. *Case-Based Reasoning.* Morgan Kaufmann Publishers, San Mateo, 1993.

[71] Kontkanen, P., Myllymäki, P., Silander, T., and Tirri, H. A Bayesian approach for retrieving relevant cases. In *Artificial Intelligence Applications (Proceedings of the EXPERSYS-97 Conference)* (Sunderland, UK, October 1997), P. Smith, Ed., IITT International, pp. 67–72.

[72] Kontkanen, P., Myllymäki, P., Silander, T., and Tirri, H. A Bayesian approach to discretization. In *Proceedings of the European Symposium on Intelligent Techniques* (Bari, Italy, March 1997), pp. 265–268.

[73] Kontkanen, P., Myllymäki, P., Silander, T., and Tirri, H. Comparing stochastic complexity minimization algorithms in estimating missing data. In *Proceedings of WUPES'97, the 4th Workshop on Uncertainty Processing* (Prague, Czech Republic, January 1997), pp. 81–90.

[74] Kontkanen, P., Myllymäki, P., Silander, T., and Tirri, H. On the accuracy of stochastic complexity approximations. In *Proceedings of the Causal Models and Statistical Learning Seminar* (London, UK, March 1997), pp. 103–117.

[75] Kontkanen, P., Myllymäki, P., Silander, T., and Tirri, H. Bayes optimal instance-based learning. In *Proceedings of the 10th European Conference on Machine Learning (ECML-98)* (Chemnitz, Germany, April 1998). (To appear).

[76] Kontkanen, P., Myllymäki, P., Silander, T., Tirri, H., and Grünwald, P. Comparing predictive inference methods for discrete domains. In *Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics* (Ft. Lauderdale, Florida, January 1997), pp. 311–318.

[77] Kontkanen, P., Myllymäki, P., Silander, T., Tirri, H., and Grünwald, P. On predictive distributions and Bayesian networks. In *Proceedings of the Seventh Belgian-Dutch Conference on Machine Learning (BeNeLearn'97)* (Tilburg, the Netherlands, October 1997), W. Daelemans, P. Flach, and A. van den Bosch, Eds., pp. 59–68.

[78] Kontkanen, P., Myllymäki, P., Silander, T., Tirri, H., and Grünwald, P. On predictive distributions and Bayesian networks. Tech. Rep. NC-TR-97-032, ESPRIT Working Group on Neural and Computational Learning (NeuroCOLT), 1997.

[79] Kontkanen, P., Myllymäki, P., Silander, T., Tirri, H., and Grünwald, P. Bayesian and information-theoretic priors for Bayesian network parameters. In *Proceedings of the 10th European Conference on Machine Learning (ECML-98)* (Chemnitz, Germany, April 1998). (To appear).

[80] Kontkanen, P., Myllymäki, P., and Tirri, H. Comparing Bayesian model class selection criteria by discrete finite mixtures. In *Information, Statistics and Induction in Science* (Proceedings of the ISIS'96 Conference, Melbourne, Australia, August 1996), D. Dowe, K. Korb, and J. Oliver, Eds., World Scientific, Singapore, pp. 364–374.

[81] Kontkanen, P., Myllymäki, P., and Tirri, H. Constructing Bayesian finite mixture models by the EM algorithm. Tech. Rep. NC-TR-97-003, ESPRIT Working Group on Neural and Computational Learning (NeuroCOLT), 1996.

[82] Kontkanen, P., Myllymäki, P., and Tirri, H. Predictive data mining with finite mixtures. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (Portland, Oregon, August 1996), E. Simoudis, J. Han, and U. Fayyad, Eds., pp. 176–182.

[83] Kontkanen, P., Myllymäki, P., and Tirri, H. Experimenting with the Cheeseman-Stutz evidence approximation for predictive modeling and data mining. In *Proceedings of the Tenth International FLAIRS Conference* (Daytona Beach, Florida, May 1997), D. Dankel, Ed., pp. 204–211.

[84] Langley, P., Iba, W., and Thompson, K. An analysis of Bayesian classifiers. In *Proceedings of the 10th National Conference on Artificial Intelligence* (San Jose, CA, July 1992), MIT Press, pp. 223–228.

[85] Lauritzen, S. Propagation of probabilities, means and variances in mixed graphical association models. *Journal of the American Statistical Association 87* (1992), 1098–1108.

[86] Lauritzen, S. *Graphical Models.* Oxford University Press, 1996.

[87] Li, Z., and D'Ambrosio, B. Efficient inference in bayes nets as a combinatorial optimization problem. *International Journal of Approximate Reasoning 11*, 1 (1994), 55–81.

[88] Lindley, D. *Making Decisions*, second ed. John Wiley & Sons, London, 1992.

[89] M., W., and Holyoak, K. Predictive and diagnostic learning within causal models: Asymmetries of cue competition. *Journal of Experimental Psychology: General 121* (1992), 222–236.

[90] Mackay, D. *Bayesian Methods for Adaptive Models.* PhD thesis, California Institute of Technology, 1992.

[91] MacKay, D. J. C. Bayesian non-linear modelling for the prediction competition. In *ASHRAE Transactions, V.100, Pt.2* (Atlanta Georgia, 1994), ASHRAE, pp. 1053–1062.

[92] Matthews, R. Faith, hope and statistics. *New Scientist 156*, 2109 (22 November 1997), 36–39.

[93] McEliece, R., Mackay, D., and Cheng, J.-F. Turbo decoding as an instance of Pearl's "belief propagation" algorithm. *International Journal of Selected areas in Communication* (1997). (to appear).

[94] McLachlan, G., and Thriyambakam, K., Eds. *The EM Algorithm and Extensions.* John Wiley & Sons, New York, 1997.

[95] Moody, J., and Darken, C. Fast learning in networks of locally-tuned processing units. *Neural Computation 1* (1989), 281–294.

[96] Myllymäki, P. *Mapping Bayesian Networks to Stochastic Neural Networks: A Foundation for Hybrid Bayesian-Neural Systems.* PhD thesis, Report A-1995-1, Department of Computer Science, University of Helsinki, December 1995.

[97] Myllymäki, P., and Tirri, H. Bayesian case-based reasoning with neural networks. In *Proceedings of the IEEE International Conference on Neural Networks* (San Francisco, March 1993), vol. 1, IEEE, Piscataway, NJ, pp. 422–427.

[98] Myllymäki, P., and Tirri, H. Massively parallel case-based reasoning with probabilistic similarity metrics. In *Topics in Case-Based Reasoning*, S. Wess, K.-D. Althoff, and M. Richter, Eds., vol. 837 of *Lecture Notes in Artificial Intelligence.* Springer-Verlag, 1994, pp. 144–154.

[99] Neal, R. *Bayesian Learning for Neural Networks.* No. 118 in Lecture Notes in Statistics. Springer-Verlag, New York, NY, 1996.

[100] Neapolitan, R. *Probabilistic Reasoning in Expert Systems.* John Wiley & Sons, New York, NY, 1990.

[101] Oliver, J., and Baxter, R. MML and Bayesianism: Similarities and differences. Tech. Rep. 206, Department of Computer Science, Monash University, December 1994.

[102] Oliver, J., and Hand, D. Introduction to minimum encoding inference. Tech. Rep. 205, Department of Computer Science, Monash University, July 1994.

[103] Oliver, J., and Hand, D. On pruning and averaging decision trees. In *Machine Learning: Proceedings of the Twelfth International Conference* (1995), A. Prieditis, Ed., Morgan Kaufmann Publishers, pp. 430–437.

[104] Oliver, R., and Smith, J. *Influence diagrams, belief nets and decision analysis.* John Wiley & Sons, 1990.

[105] Paris, J. *The Uncertain Reasoner's Companion: A Mathematical Prespective*, vol. 39 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, Cambridge, 1994.

[106] Parsa, I. KDD-97 knowledge discovery and data mining tools competition. *Knowledge Discovery Nuggets 97*, 19 (1997). An electronic journal, available at `http://www.kdnuggets.com/`.

[107] Pearl, J. Fusion, propagation and structuring in belief networks. *Artificial Intelligence 29* (1986), 241–288.

[108] Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann Publishers, San Mateo, CA, 1988.

[109] Pearl, J., and Verma, T. A theory of inferred causation. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference* (1991), A. Allen, R. Fikes, and E. Sandewall, Eds., Morgan Kaufmann Publishers, San Mateo, CA, pp. 441–452.

[110] Poggio, T., and Girosi, F. Networks for approximation and learning. *Proceedings of the IEEE 78*, 9 (1990), 1481–1497.

[111] Poland, W. *Decision Analysis with Continuous and Discrete Variables: A Mixture Distribution Approach.* PhD thesis, Department of Engineering Economic Systems, Stanford University, 1994.

[112] Ripley, B. *Pattern Recognition and Neural Networks.* Cambridge University Press, 1996.

[113] Rissanen, J. Modeling by shortest data description. *Automatica 14* (1978), 445–471.

[114] Rissanen, J. Stochastic complexity. *Journal of the Royal Statistical Society 49*, 3 (1987), 223–239 and 252–265.

[115] Rissanen, J. *Stochastic Complexity in Statistical Inquiry.* World Scientific Publishing Company, New Jersey, 1989.

[116] Rissanen, J. Fisher information and stochastic complexity. *IEEE Transactions on Information Theory 42*, 1 (January 1996), 40–47.

[117] Rissanen, J. Information theory and neural nets. In *Mathematical Perspectives on Neural Networks*, P. Smolensky, M. Mozer, and D. Rumelhart, Eds. Lawrence Erlbaum Associates, 1996, ch. 16.

[118] Rissanen, J. Oppivien ja älykkäiden järjestelmien sovellukset 1994–1998, tutkimusprojektien arviointi. Teknologiaohjelmaraportti 2/97, Teknologian kehittämiskeskus (TEKES), 1997.

[119] Robinson, R. Counting unlabeled asyclic graphs. In *Combinatorial Mathematics*, C. Little, Ed., no. 622 in Lecture Notes in Mathematics. Springer-Verlag, 1977.

[120] Roth, D. On the hardness of approximate reasoning. In *Proceedings of the 13th International Joint Conference on Artificial Intelligence* (1993), vol. 1, pp. 613–618.

[121] Santos, E., and Shimony, S. Belief updating by enumerating high-probability independence-based arguments. In *Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence (UAI–94)* (Seattle, Washington, July 1994), R. Lopez de Mantaras and D. Poole, Eds., Morgan Kaufmann Publishers, San Francisco, CA, pp. 506–513.

[122] Saul, L., Jaakkola, T., and Jordan, M. Mean field theory for sigmoid belief networks. *Journal of Artificial Intelligence Research 4* (1996), 61–76.

[123] Schafer, G., Ed. *Probabilistic Expert Systems.* Society for Industrial and Applied mathematics (SIAM), 1996.

[124] Schafer, G., and Pearl, J., Eds. *Readings in Uncertain Reasoning.* Morgan Kaufmann Publishers, San Mateo, CA, 1990.

[125] Schwarz, G. Estimating the dimension of a model. *Annals of Statistics 6* (1978), 461–464.

[126] Scott, D. *Multivariate Density Estimation. Theory, Practice, and Visualization.* John Wiley & Sons, New York, 1992.

[127] Seppänen, M. Bayes-verkon muodostaminen datasta. Master's thesis, Report C-1996-33, Department of Computer Science, University of Helsinki, 1996.

[128] Shachter, R. Probabilistic inference and influence diagrams. *Operations Research 36*, 4 (July-August 1988), 589–604.

[129] Shimony, S. Finding MAPs for belief networks is NP-hard. *Artificial Intelligence 68* (1994), 399–410.

[130] Singh, M., and Valtorta, M. An algorithm for construction of Bayesian network structures from data. In *Uncertainty in Artificial Intelligence, Proceedings of the Ninth Conference* (1993), E. Horvits and F. Jensen, Eds., Morgan Kaufmann Publishers, San Francisco, CA, pp. 259–265.

[131] Smyth, P., Heckerman, D., and Jordan, M. Probabilistic independece networks for hidden Markov probability models. Tech. Rep. A.I. Memo 1565, C.B.C.L Memo 132, M.I.T, February 1996.

[132] Sorsa, M. Kvalitatiiviset Bayesverkot. Master's thesis, Report C-1996-98, Department of Computer Science, University of Helsinki, 1996.

[133] Specht, D. Probabilistic neural networks. *Neural Networks 3* (1990), 109–118.

[134] Spiegelhalter, D., Dawid, P., Lauritzen, S., and Cowell, R. Bayesian analysis in expert systems. *Statistical Science 8*, 3 (1993), 219–283.

[135] Spirtes, P., Glymour, C., and Scheines, R., Eds. *Causation, Prediction and Search.* Springer-Verlag, 1993. Out of print. Can be downloaded from 'URL: http://hss.cmu.edu/html/departments/philosophy/TETRAD.BOOK/book.html'.

[136] Stanfill, C., and Waltz, D. Toward memory-based reasoning. *Communications of the ACM 29*, 12 (1986), 1213–1228.

[137] Tikhonov, A. On solving incorrectly posed problems and method of regularization. *Doklady Akademii Nauk USSR 151* (1963), 501–504.

[138] Tirri, H. *Plausible Prediction by Bayesian Inference.* PhD thesis, Report A-1995-1, Department of Computer Science, University of Helsinki, June 1997.

[139] Tirri, H., Kontkanen, P., and Myllymäki, P. A Bayesian framework for case-based reasoning. In *Advances in Case-Based Reasoning*, I. Smith and B. Faltings, Eds., vol. 1168 of *Lecture Notes in Artificial Intelligence*. Springer-Verlag, Berlin Heidelberg, November 1996, pp. 413–427.

[140] Tirri, H., Kontkanen, P., and Myllymäki, P. Probabilistic instance-based learning. In *Machine Learning: Proceedings of the Thirteenth International Conference* (1996), L. Saitta, Ed., Morgan Kaufmann Publishers, pp. 507–515.

[141] Titterington, D., Smith, A., and Makov, U. *Statistical Analysis of Finite Mixture Distributions.* John Wiley & Sons, New York, 1985.

[142] Vapnik, V. *The Nature of Statistical Learning Theory.* Springer-Verlag, New York, 1995.

[143] Walley, P. Measures of uncertainty in expert systems. *Artificial Intelligence 83* (1996), 1–58.

[144] Watson, I., and Marir, F. Case-based reasoning: A review. *The Knowledge Engineering Review 9*, 4 (1994), 327–354.

[145] Wellman, M. Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence 44*, 3 (August 1990), 257–304.

[146] Whittaker, J. *Graphical Models in Applied Multivariate Statistics.* John Wiley & Sons, 1990.

[147] Yannakis, M. Computing the minimal fill-in is NP-complete. *SIAM Journal of Algebraic Discrete Methods 2* (1981), 77–79.

[148] Zhang, L., and Poole, D. Sidestepping the triangulation problem in Bayesian net computations. In *Proceedings of the Eight Annual Conference on Uncertainty in Artificial Intelligence* (Palo Alto, CA, July 1992), Morgan Kaufmann Publishers, pp. 360–367.