

Tekoäly Syksy 2004

Eero Hyvönen
Helsingin yliopisto ja HIIT
Semanttisen laskennan tutkimusryhmä
SeCo

Kurssin tavoitteet ja sisältö

■ Johdatus tekoälyyn: "Good Old-fashiononed AI"

- ◇ intelligent agents
- ◇ search and game-playing
- ◇ logical systems
- ◇ planning systems
- ◇ uncertainty—probability and decision theory
- ◇ learning
- ◇ language
- ◇ perception
- ◇ robotics
- ◇ philosophical issues

Tekoäly-kurssin fokus

"Kolme käsitettä" -kurssit

Kieliteknologian kurssit
Robotiikan kurssit

Tekoäly-kurssi ym.

Kurssin järjestelyt ja ohjelma

- Esitetty kotisivulla

- <http://www.cs.helsinki.fi/eahyvone/courses/tekoaly2004/>

Tekoäly, Eero Hyvönen, 2004

3

Johdatus tekoälyyn

Mitä on tekoäly?

- Tekoäly (artificial intelligence) on monitieteinen tieteenala, jolla
 - Analysoidaan älykkääksi katsottua toimintaa
 - Syntetisoidaan älykkäitä tietojärjestelmiä
- Tekoälyllä viitataan usein myös itse ilmiöön, ts. koneen älyyn
 - Koneäly (machine intelligence) termiä voi käyttää ilmiön nimenä
 - Keinoäly termiäkin käytetään
 - tekomunuainen vs. keinomunuainen
 - tekosilmä vs. keinosilmä

Tekoäly, Eero Hyvönen, 2004

5

Neljä näkökulmaa tekoälyyn

	Human	Rational
Think T h i n k	“The exciting new effort to make computers think ... <i>machines with minds</i>, in the full and literal sense” (Haugeland, 1985) “[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning ...” (Bellman, 1978)	“The study of mental faculties through the use of computational models” (Charniak and McDermott, 1985) “The study of the computations that make it possible to perceive, reason, and act” (Winston, 1992)
	“The art of creating machines that perform functions that require intelligence when performed by people” (Kurzweil, 1990) “The study of how to make computers do things at which, at the moment, people are better” (Rich and Knight, 1991)	“A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes” (Schalkoff, 1990) “The branch of computer science that is concerned with the automation of intelligent behavior” (Luger and Stubblefield, 1993)

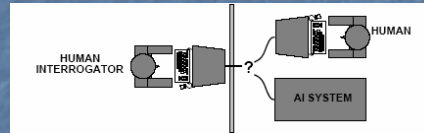
Taulukko 1: 4 käsitystä tekoälystä [RN95, kuva 1.1].

Tekoäly, Eero Hyvönen, 2004

6

1. Tekoäly: Ihmisen toiminnan näkökulma

- Turingin testi (1950)
 - Älykkyyden operationaalinen määrittely
 - Väistää filosofisen pohdiskelun
 - Edellytyksiä (mentaalisia)
 - Luonnollisen kielen ymmärtäminen
 - Tietämyksen esittäminen
 - Päättely
 - Oppiminen
 - Edellytyksiä (fyysisiä, "total Turing test")
 - Konenäkö
 - Robotiikka
 - Rajoitettu Turingin testi
 - Esim. shakin peluuseen tms. rajattuun sovellusalueeseen
 - Ei merkittävä tavoite käytännössä
 - Loebner prize –kilpailu (<http://www.loebner.net>)
 - 1990, Hugh Loebner, \$100,000 Turingin kokeen läpäisijälle
 - Joka vuosi \$2000 palkinto parhaalle yritykselle
 - Esimerkki: Raymond Kurzweilin "Ramona"



Tekoäly, Eero Hyvönen, 2004

7

2. Tekoäly: ihmismäisen ajattelun näkökulma

- Kognitiotieteen näkökulma
 - Lähtökohtana ihmisen ajattelun mallittaminen
 - Tärkeä inspiraation lähde tekoälytutkimuksessa
- Luonnon imitointi vs. tekninen ratkaisu
 - Esim. lentotekniikka
 - Leonardo Da Vincin räpyttelevä lentokone
 - Nykyinen lentokone ei imitoi lintua ja lentää sitä nopeammin
 - Toisaalta lentokone ei osaa laskeutua puun oksalle
- Ajattelun malli $\neq$ hyvä algoritmi

Tekoäly, Eero Hyvönen, 2004

8

3. Tekoäly: rationaalisen ajattelun näkökulma

- Ajattelun lakien kehittäminen
 - Aristoteleen syllogismit
 - Sokrates on ihminen
Ihmiset ovat erehtyväisiä
-> Sokrates on erehtyväinen
 - "Sokrates", "ihminen", "erehtyväinen" voidaan korvata muuttujilla
 - -> sovelluksesta riippumaton päättely!
 - Logiikan kehitys
 - Ongelmia
 - Epätasällisyyden hallinta (epätarkka, virheellinen, puutteellinen tieto)
 - Päättelyn suuntaaminen ja tehokkuus

Tekoäly, Eero Hyvönen, 2004

9

4. Tekoäly: rationaalisen toiminnan näkökulma

- Sisältää rationaalisen loogisen ajattelun
- Toiminta myös tilanteissa, joissa loogista ratkaisua ei ole mahdollista johtaa
- Ei-looginen reaktiivinen rationaalinen toiminta (esim. robotiikassa)
- Oppikirjan näkökulma tekoälyyn

Tekoäly, Eero Hyvönen, 2004

10

Muita määritelmiä

- ”Tekoäly on sitä mitä ei vielä pystytä ohjelmoimaan tietokoneelle”
 - Kuinka monen mielestä taskulaskin on älykäs?
 - Kuinka monen mielestä shakkikone on älykäs?
- ”Tekoälyä ei voi määritellä, mutta kyllä sen tuntee, kun sellainen vastaan tulee”
 - Vrt. pornografian määrittely-yritykset USA:n kongressissa

Tekoäly, Eero Hyvönen, 2004

11

Tutkimuskohteita

- Tehtävätyypit
 - Keinotekoiset tehtävät
 - Pelit, matematiikka, logiikka, ohjelmointi, ...
 - Asiantuntijatehtävät
 - Suunnittelu, diagnosointi, data-analyysi, ...
 - Arkipäiväiset tehtävät (common sense)
 - Näkeminen, kieli, liikkuminen, oppiminen, ...
 - Paradoksaalisesti kaikkein vaikein alue!

Tekoäly, Eero Hyvönen, 2004

12

Tekoälyn tausta ja historia

Filosofia

- Aristoteles (384-322 eKr)
 - Syllogismit
- Wilhelm Schickard, Blaise Pascal (1600-luku)
 - Mekaaniset laskukoneet
- Rene Descartes (1600-luku)
 - Eläin = "Lihakone"
 - Dualismi vs. materialismi
- Leibniz (1600-1700-luvut)
 - Mekaaninen ongelmanratkenta
- Empirismi – miten fyysinen mieli saa tietoa
 - David Hume (1700-luku) , induktioperiaate

Matematiikka

- Formaali päättely
- Algoritmin käsite
 - al-Khowarazmi (900-luku)
 - Esim. Euklides, Boole, ...
- Laskennan rajat
 - Mitä voidaan todistaa/laskea
 - Kurt Gödelin epätäydellisyysteoreema (1931)
 - Church-Turing teesi: Turingin kone (1936) laskee kaiken laskettavissa olevan
 - Kaikkea ei voi laskea. esim. sitä pysähtyykö kone koskaan
- Laskennan kompleksisuus
 - Polynominen (tractable) vs. eksponentiaalinen (intractable)
 - NP-täydellisyys
- Epätasällisen tiedon käsittely
 - Todennäköisyys
 - Gerolamo Cardano (1501-1576)
 - Thomas Bayes (1702-1761): uskomusten päivitys

Tekoäly, Eero Hyvönen, 2004

15

Taloustiede

- Rationaalisuus on hyödyn tavoittelua
 - Miten maksimoida hyöty?
 - Entä jos muut pelaavat sinua vastaan?
 - Tulevan hyödyn diskonttaus?
- Kehitystä
 - Utiliteetin käsite: Leaon Valras (1834-1910)
 - Peliteoria: John von Neumann & Oskar Morgenstern (1944)
 - Päättösteoria: todennäköisyys + utiliteetti (1957)
 - Operaatiotutkimus/systeemianalyysi

Tekoäly, Eero Hyvönen, 2004

16

Neurotiede

- Miten aivot käsittelevät tietoa?
 - Tietoisuus sijoitettiin aivoihin vasta 1700-luvulla
- Aivojen rakenne on ollut malli monelle tietojenkäsittelyn arkkitehtuurille
 - "Sähköaivot"
 - Neuroverkot
 - Semanttiset verkot

Tekoäly, Eero Hyvönen, 2004

17

Tietokone vs. aivot

- | ■ KONE | ■ IHMINEN |
|---|----------------------|
| ■ CPU: 10^8 porttia | ■ 10^{11} neuronia |
| ■ Muisti: 10^{10} bit RAM
10^{11} bit levy | ■ 10^{11} neuronia |
| ■ Taajuus: 10^{-9} s | ■ 10^{-3} s |
| ■ Kaista: 10^{10} bit/s | ■ 10^{14} bit/s |
| ■ Muistipäiv./s: 10^9 | ■ 10^{14} |

Tehovertilau: Aivot 100.000 kertaa (?) konetta tehokkaammat
Tietojen käsittelyn laatu vs. teho?

Tekoäly, Eero Hyvönen, 2004

18

Psykologia

- Miten elävät olennot ajatelevat ja toimivat?
- Hermann von Helmholtz (1821-1894)
 - Tieteellinen psykologia
- Behaviorismi
 - Input-output näkökulma
- Kognitiivinen psykologia
 - Aivot tietojenkäsittelylaitteena
- Kognitiotiede 1956
 - Kognitiivinen teoria on kuin tietokoneohjelma

Tekoäly, Eero Hyvönen, 2004

19

Tietokonetekniikka

- Miten voidaan rakentaa tehokas tietokone?
- Varhaiset laskukoneet
 - 1. ohjelmitava kone (1805): Joseph Maria Jacquard, kutomakone
 - Charles Babbagen (1792-1871) Difference Engine ja Analytical Engine
 - DE toteutettiin 1991 British Museumin toimesta
 - Ada Lovelace – ensimmäinen ohjelmoija (AE)
- Modernin tietokoneen synty 2. maailmansodan aikana
 - Englanti:
 - Alan Turing et al.: 1. toimiva tietokone, Heath Robinson (1940), Colossus (1943)
 - Saksa:
 - Konrad Zuse: 1. ohjelmitava kone Z-3 (1941)
 - USA:
 - John Atanasoff, Clifford Berry: 1. elektroninen tietokone ABC (1940-42)

Tekoäly, Eero Hyvönen, 2004

20

Säätöteoria ja kybernetiikka

- Artefaktin toiminnan itseohjaus?
- Ktesibios Aleksandrialainen (250 eKr)
 - 1. säätöjärjestelmä: veden virtauksen säätö vesikellossa
- Takaisinkytkennän teoria (feedback systems) 1800-luvulla
 - Esim. termostaatti
- Säätöteoria (Control Theory)
 - Norbert Wiener (1894-1964)
 - Toiminnan ohjaus minimoimalla tavoite- ja nykytilan eroa (virhettä)
 - Nykyään: pyrkimyksenä tavoitefunktion arvon maksimointi
 - Idea sama kuin tekoälyssä, mutta menetelmät perinteistä matematiikka ja sovelluskohde fyysiset systeemit

Tekoäly, Eero Hyvönen, 2004

21

Kielitiede

- Kielen ja ajattelun suhde?
- Moderni kielitiede syntyi 1957
 - B. F. Skinnerin behavioristinen teoria kielen oppimisesta
 - Noam Chomskyn kritiikki ja oma teoria (Syntactic Structures)
- Tietokoneлингvistiikka / Kieliteknologia (Computational Linguistics)
 - Formaalit menetelmät ja laskenta kielen mallina
- Luonnollisen kielen ymmärtäminen (Natural Language Understanding)
- Vaikutti mm. tietämyksen esittämisen tutkimukseen

Tekoäly, Eero Hyvönen, 2004

22

Tekoälyn historiaa

- Varhaishistoria (1943-1955)
 - McCulloch & Pitts neuroverkko (1943)
 - Marvin Minsky, Dean Edmonds SNARC
 - Alan Turing: "Computing Machinery and Intelligence", 1950.
- Synty (1956)
 - Dartmouth College työpaja
 - Nimi "Artificial Intelligence" otettiin käyttöön
 - John McCarthy, Herbert Simon, Alan Newell, Arthur Samuel, ...

Tekoäly, Eero Hyvönen, 2004

23

- Varhainen innostus (1952-1969)
 - General Problem Solver (GPS)
 - Logic Theorist
 - Samuelin oppiva tammiohjelma
 - Lisp-kieli (1958)
 - 2. vanhin edelleen käytetty kieli (Fortranin jälk.)!
 - Mikromaailmat, Terry Winogradin SHRDLU
 - Frank Rosenblattin oppiva perceptron

Tekoäly, Eero Hyvönen, 2004

24

■ 1. takatalvi (1966-1973)

- Systeemit eivät skaalautuneetkaan
 - eri sovellusalueille ja
 - isoille aineistolle
 - Lighthill raportti (1973) leikkasi alan tutkimuksen Englannissa
- Sovelluskohtainen tietämys puuttui
- Konekääntämisen vastoinkäymiset
 - "The spirit is winning but the flesh is weak"
 - -> "The vodka is good but the meat is rotten"
- Kone-evoluution (geneettiset algoritmit) vastoinkäymiset
- Neurolaskennan vastoinkäymiset
 - Perceptronin ilmaisuvoiman rajoitus kävi ilmi
 - Ei voinut oppia esim. XOR-funktiota

Tekoäly, Eero Hyvönen, 2004

25

Tekoälyn kevät (1970-luku)

- Uusi nousu: tietämysjärjestelmät (knowledge based systems)
 - Geneerisistä ongelmanratkojista sovelluskohtaisen tietämyksen käyttöön
 - Asiantuntijajärjestelmät (expert systems)
 - DENDRAL, MYCIN, Prospector,...
 - Edward Feigenbaum ym.
 - Kielen ymmärtäminen
 - Roger Schank ym.
 - Säännöt ja kehykset (frame)

Tekoäly, Eero Hyvönen, 2004

26

- Teollinen soveltaminen (1980-)
 - R1, XCON, ym. asiantuntijajärjestelmät
 - Isot kansalliset/kansainväliset hankkeet
 - Japanin 5. polven tietokone –hanke
 - Logiikkaperustaiset rinnakkaiset tietokoneet
 - Pearl Harbour revised
 - USA:n MCC
 - Iso-Britannian Alvey-raportti
 - Sumean logiikan sovellukset
 - Lopputulos: pieniä näppäriä sääntöjärjestelmiä eikä ylivertaista Feigenbaumilaista tekoälyä!

Tekoäly, Eero Hyvönen, 2004

27

- Tekoälyn 2. takatalvi (1988-)
 - Taas petyttiin odotuksissa
 - Tietämyksen siirron ongelma
 - Tietämyksen ylläpidon ongelma
 - Logiikka- ja Lisp-koneet eivät lunastaneet odotuksia
 - Neurolaskenta kuitenkin nousi yhdeksi "pelastajaksi" (1986-)
 - Back-probagation learning algoritmi
 - Teuvo Kohosen SOM
 - Symbolisesta laskennasta konnektionismiin

Tekoäly, Eero Hyvönen, 2004

28

- Vakiintuminen tieteelliseksi alueeksi (1987-)
 - Pelle Peloton –tyyppinen protoilu -> main stream computer science
 - Metodologian kypsyminen
 - Ad hoc –ratkaisuista hyvin määriteltyihin menetelmiin
 - Esim. Judea Pearl ja Bayes-verkot
 - Agentti-perustainen näkemys tekoälystä yhdistää alaa
 - Tekoäly ei ole vain työkalukokoelma

Tekoäly, Eero Hyvönen, 2004

29

Miten älykäs kone on tänään?

Automaattinen suunnittelu ja skedulointi
 DeepSpace
 Palaaminen
 Deep Blue voitti Garry Kasparovin 1997
 Robottiikka & ohjaus
 NavLab ajoi läpi USA:n 98% matkasta
 Diagnosointi
 Paljon sovelluksia esim. lääketieteessä
 Kielen ymmärtäminen
 Proverb ristisanaratkoja

...

Tekoäly, Eero Hyvönen, 2004

30

Tekoälyä on kaikkialla

Sumeat säätimet

Kameran etäisyysmittaus

Auton ABS, pölynimurit, pesukoneet

Pelien tekoäly

Kielenhuolto ja kääntäminen

Matemaattinen ongelmanratkonta

MUTTA KUN ONGELMA OSATAAN RATKAISTA EI
SEN ARVELLA ENÄÄ VAATIVAN (TEKO)ÄLYÄ!

Tekoäly, Eero Hyvönen, 2004

31

Tekoälyn filosofiaa

Keskeinen kysymys: voiko kone "ajatella"?

- Mutta mitä on "ajattelu"
 - ajattelua ihmismielessä
 - jotain yleisempää ongelmanratkontaa tms.
 - Voiko kone lentää?
 - Voiko kone uida?
- Kaksi peruskäsitystä
 - Heikko tekoäly
 - Vahva tekoäly

Tekoäly, Eero Hyvönen, 2004

33

Heikko tekoäly (weak AI)

- Koneet toimivat ikään kuin olisivat älykkäitä
- Älyn kriteeri: konetta ei voida ulkoisesti havainnoimalla erottaa älykkästä ihmisestä
- Turingin koe
 - Operationaalinen määrittely
 - Käytännöllinen insinöörin näkökulma
 - Vrt. esim. ajotaito ja insinööriajo
 - Rajoitettu Turingin koe

Tekoäly, Eero Hyvönen, 2004

34

Kyvyttömyyden vastaväite

- Koneet ovat mekaanisia, eivätkä voi olla "luovia"
 - Esim. taskulaskin ja shakkikoneet
- Koneet eivät voi oppia ajamaan polkupyörää, rakastamaan, nauttimaan oluesta jne.

Tekoäly, Eero Hyvönen, 2004

35

Formaali vastaväite

- Gödelin epätäydellisysteoreema
 - Otetaan mikä tahansa (konsistentti) formaali järjestelmä F joka sisältää luonnollisten lukujen aritmetiikan
 - F:ssä voidaan esittää lause G, joka on tosi, mutta jota ei voida todistaa F:ssä
- Merkitseekö tämä sitä, että kone ei voi olla älykäs samalla tavalla kuin ihminen?
 - Vilkas vuosikymmeniä kestänyt keskustelu
 - Koskeeko rajoitus myös ihmistä?

Tekoäly, Eero Hyvönen, 2004

36

Älykkyyden ei-formaalisuus vastaväite

- Logiikka ei riitä kompleksisen ihmisen toiminnan kuvaamiseen
 - Good Old-Fashioned AI (GOFAI)
 - Qualification problem
 - Tarvitaan tausta/kontekstietoa (intuitio)
 - Esim. shakkimestarin pelaaminen (H. Dreyfus)

Tekoäly, Eero Hyvönen, 2004

37

Vahva tekoäly (Strong AI)

- Kone ei ainoastaan simuloi älykästä toimintaa vaan on "älykäs" vastaavatyypisessä mielessä kuin ihminen
 - Tietoisuus: voiko kone olla tietoinen tilastaan
 - Fenomenologia: voiko kone tuntea reaalimaailman
 - Intentionaalisuus: koneen pyrkimykset ja tavoitteet

Tekoäly, Eero Hyvönen, 2004

38

- Keinotekoisuus vs. aitous
 - Keinokasvi ei ole aito kasvi
 - Keinotekoinen Gallen-Kallela ei ole aito
 - Vaikka olisi tarkka kopio
 - Synteettinen urea on aitoa ureaa
 - Sama kemiallinen koostumus
 - Shakkipelin simulaatio on aito peli
 - Keinotekoinen äly on aitoa?
 - Funktionalismi: input-output käyttäytyminen ratkaisee
 - Biologinen naturalismi: implementaatio ratkaisee

Tekoäly, Eero Hyvönen, 2004

39

Mind-Body Problem

- Dualismi: henki $\neq$ materia
 - Ongelma: miten henkiset tilat voivat aiheuttaa fyysisiä aktiviteetteja
- Monismi (materialismi): henki = materia
 - Ongelma: vapaata tahtoa ei olisi
 - Mieli olisi kuin putoava kivi
 - Ongelma: tietoisuuden palauttaminen materiaan

Tekoäly, Eero Hyvönen, 2004

40

- Hans Moravec'in aivoproteesi-ajatuskoe (1988)

- Funktionalistinen argumentti
- Aivojen solujen korvaaminen yksitellen mekaanisilla soluilla
 - Tietoisuus ilmeisesti säilyy?
- Entä isompien aivonosien korvaaminen
 - Jopa koko aivot kerralla
 - Säilyykö tietoisuus silloin?

Tekoäly, Eero Hyvönen, 2004

41

- John Searlen Kiinalainen huone –ajatuskoe (1980)

- Biologismnaturalistinen argumentti
- Koneen (huoneen) input-output kiinaksi
- Sisällä kiinaa osaamaton henkilö, jolla tarkka sääntökirja vastauksista
 - Turingin koe menee läpi
- Algoritmi (syntaksi) ei siis synnytä mieltä?
- Ydinkysymys: onko mieli koko systeemissä?

Tekoäly, Eero Hyvönen, 2004

42

Tekoäly, etiikka, riskit

- Työ ja teolliset vallankumoukset
 - Koneet ja teollisuuden/maatalouden rakennemuutos
 - Tekoäly ja valkokaulusalojen rakennemuutos
 - Älykkäät järjestelmät avustavat vs. korvaavat ihmisiä
- Käsitys ihmisen asemasta luonnossa muuttuu
 - Kopernikaaninen kumous
 - Darwinin teoria
 - Freudin teoria
 - Tekoäly
 - Tekoelämä

Tekoäly, Eero Hyvönen, 2004

43

- Orwellilainen yhteiskunta
 - Terroristien etsintä
 - Echelon systeemi
 - Yksityisyyden suoja
- Vastuukysymykset
 - Esim. koneen tekemä hoitovirhe?
 - Agentin taloudelliset päätökset
- Älyaseiden edut ja vaarat
 - Nykyisin USA ja YK pystyvät toimimaan "maailmanpoliiseina" teknologisen etumatkan turvin
 - Esim. Irakin valloitus pienin miestappioin
 - Älykkäät joukkotuhoaseet uhka väärissä käsissä
- Koneet ovat evoluution seuraava askel?
 - Esim. R. Kurzweilin visiot
 - v. 2010 1000e:n PC:n teho vastaa rotan aivojen laskentatehoa
 - v. 2025 1000e:n PC:n teho vastaa ihmisen aivojen laskentatehoa
 - v. 2060 1000e:n PC:n teho vastaa ihmiskunnan aivojen laskentatehoa
 - S. Kubrick, S. Spielberg elokuvat (Avaruusseikkailu 2001, A.I.)
 - Koneiden moraaliset oikeudet

Tekoäly, Eero Hyvönen, 2004

44

Lähteet

- Kalvot perustuvat pääosin oppikirjaan: S. Russell, P. Norvig: Artificial Intelligence, a Modern Approach. Prentice-Hall, 2003.
- E. Hyvönen, I. Karanta, M. Syrjänen (toim.): Tekoälyn ensyklopedia. Gaudeamus, 1995.
- R. Kurzweil: The Age of Spiritual Machines. Benguin Books, 1999.
- M. Nykänen: Tekoäly-kurssin kalvot, 2001.