



# JOHDATUS TEKOÄLYYN

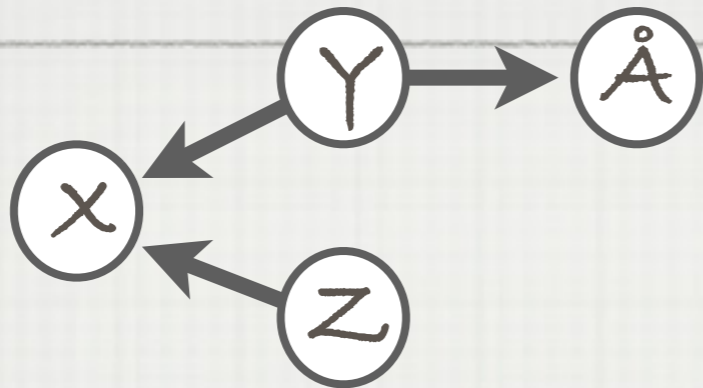
TEEMU ROOS



HELSINGIN YLIOPISTO

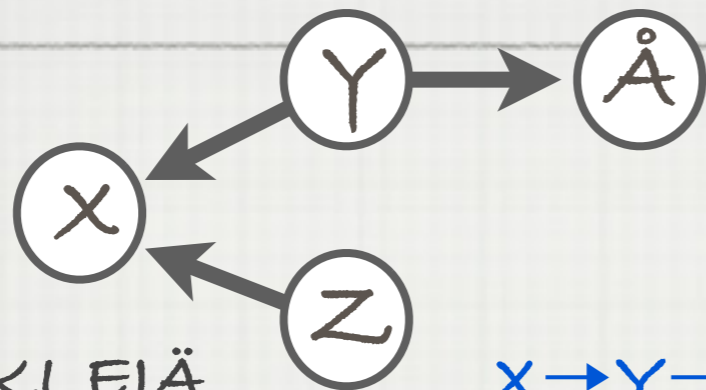


# BAYES-VERKOT



- \* BAYES-VERKKO ON TODENNÄKÖISYYSMALLIN ESITYS
- \* VERKON SOLMUT OVAT SATUNNAISMUUTTUJIA (ESIM. NOPAN SILMÄLUKU)
- \* VERKON KAARET ("NUOLET") VASTAAVAT SUORIA RIIPPUUKSIA: EI KAARTA  $\Leftrightarrow$  EHDOLLINEN RIIPPUMATTOMUUS (TARKKA SELITYS TODENNÄKÖISYYSMALLINNUS-KURSSILLA)
- \* KAARIIN LIITTYY EHDOLLISET TODENNÄKÖISYYDET  $P(S=s \mid \text{VANHEMMAT}_S=...)$  ESIM.  $\text{VANHEMMAT}_X = \{Y, Z\}$

# BAYES-VERKOT



\* EI SYKLEJÄ

$X \rightarrow Y \rightarrow Z \rightarrow X$

\* YHTEISTODENNÄKÖISYYDET SAADAAN LASKEMALLA

$$P(y, z, x, \text{å}) = P(y) \times P(z) \times P(x | y, z) \times P(\text{å} | y)$$

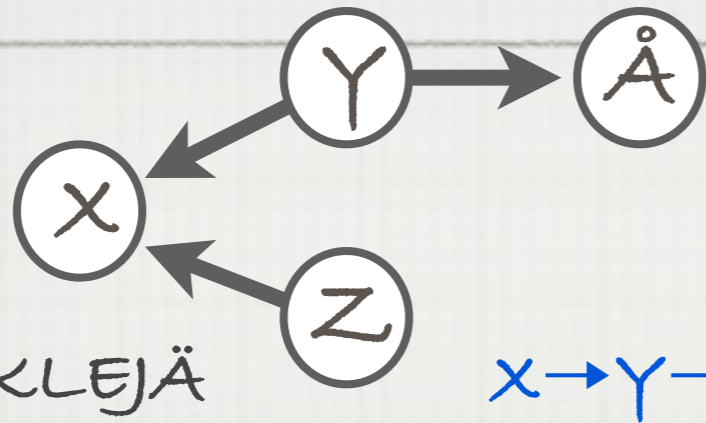
\* VERTAA KETJUSÄÄNTÖÖN:

$\text{VANHEMMAT}_x$

$$P(y, z, x, \text{å}) = P(y) \times P(z | y) \times P(x | y, z) \times P(\text{å} | x, y, z)$$



# BAYES-VERKOT



\* EI SYKLEJÄ

$x \rightarrow y \rightarrow z \rightarrow x$

\* YHTEISTODENNÄKÖISYYDET SAADAAN LASKEMALLA

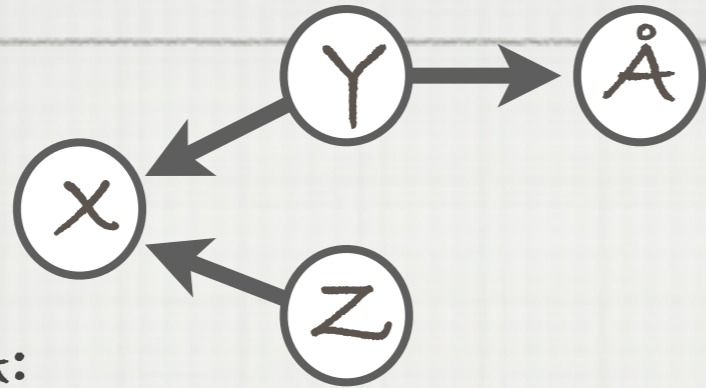
$$P(y, z, x, a) = P(y) \times P(z) \times P(x | y, z) \times P(a | y)$$

"EHDOLLINEN RIIPPUMATTOMUUS"

\* VERTAA KETJUSÄÄNTÖÖN:

$$P(y, z, x, a) = P(y) \times P(z | y) \times P(x | y, z) \times P(a | x, y, z)$$

# BAYES-VERKOT



\* ETUJA:

- \* - EHDOLLISET JAKAUMAT HELPOMPI MÄÄRITELLÄ, KOSKA VÄHEMMÄN MÄÄRITELTÄVÄÄ
- NOPEUTTAA MYÖS PÄÄTTELYÄ

\* VERTAA KETJUSÄÄNTÖÖN:

"EHDOLLINEN RIIPPUMATTOMUUS"

$$P(y, z, x, a) = P(y) \times P(z | y) \times P(x | y, z) \times P(a | x, y, z)$$



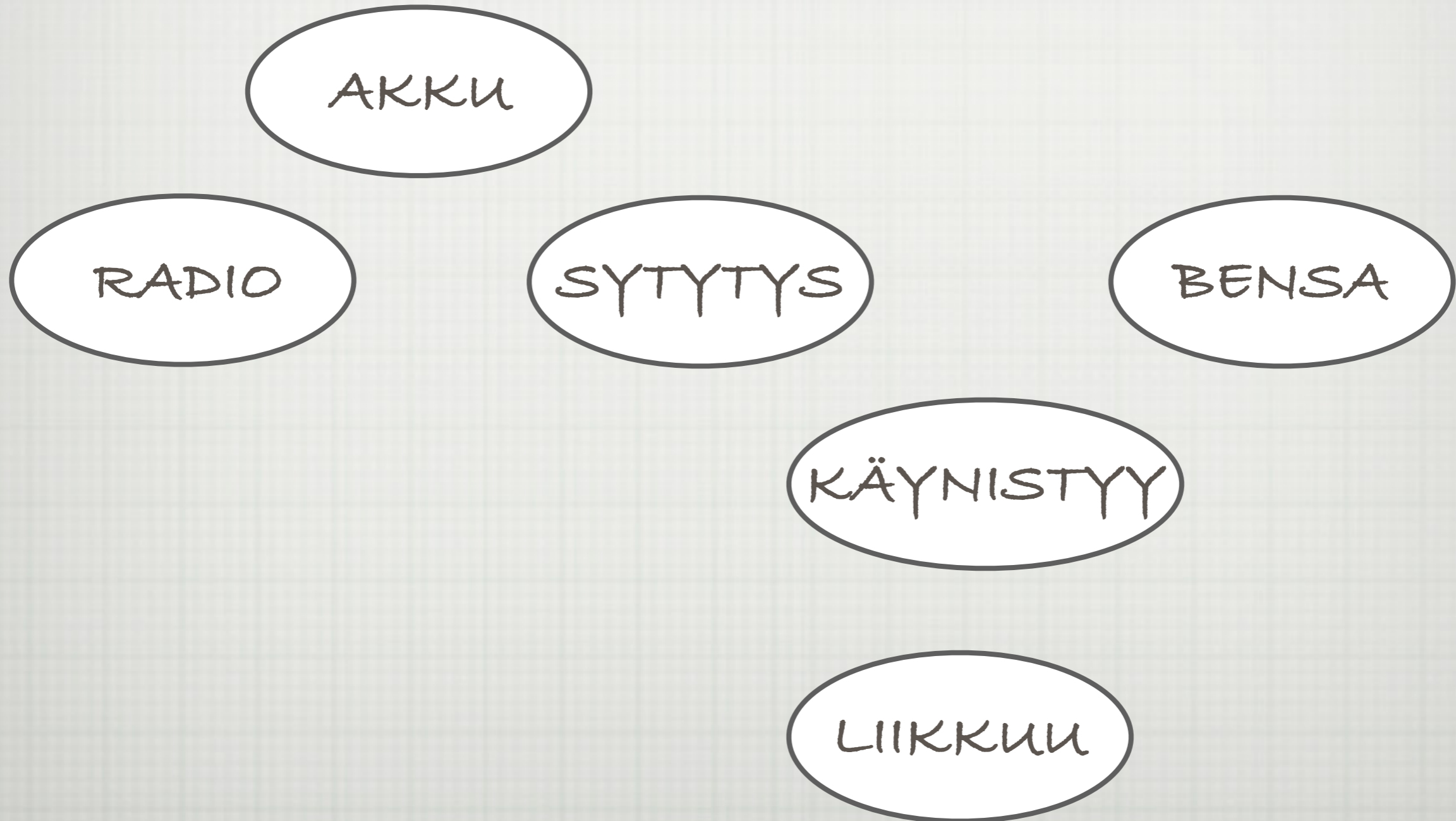
# BAYES-VERKOT

---



# BAYES-VERKOT

---





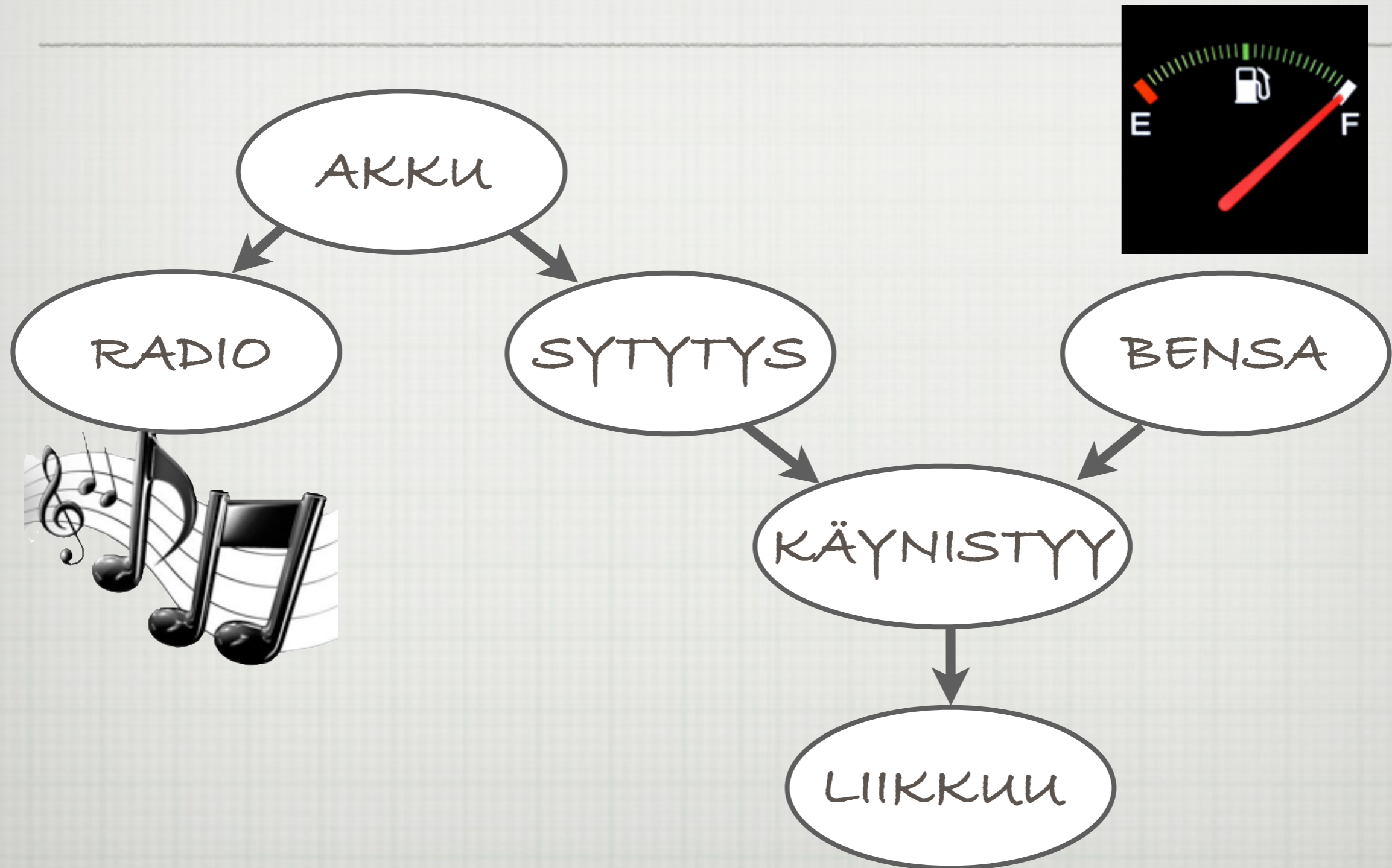
# BAYES-VERKOT

---

- \* JOS AKKU ON TYHJÄ, RADIO EI SOI EIKÄ SYTYTYYS TOIMI.
- \* JOS SYTYTYYS EI TOIMI, AUTO EI KÄYNNISTY.
- \* JOS BENSATANKKI ON TYHJÄ, AUTO EI KÄYNNISTY.
- \* JOS AUTO EI KÄYNNISTY, AUTO EI LIIKU.
  
- \* AUTO EI KÄYNNISTY: MISSÄ VIKA?  $P(\text{TILA} | \text{HAVAINNOT})?$
- \* SOIKO RADIO? ONKO BENSAA? ← HAVAINNOT

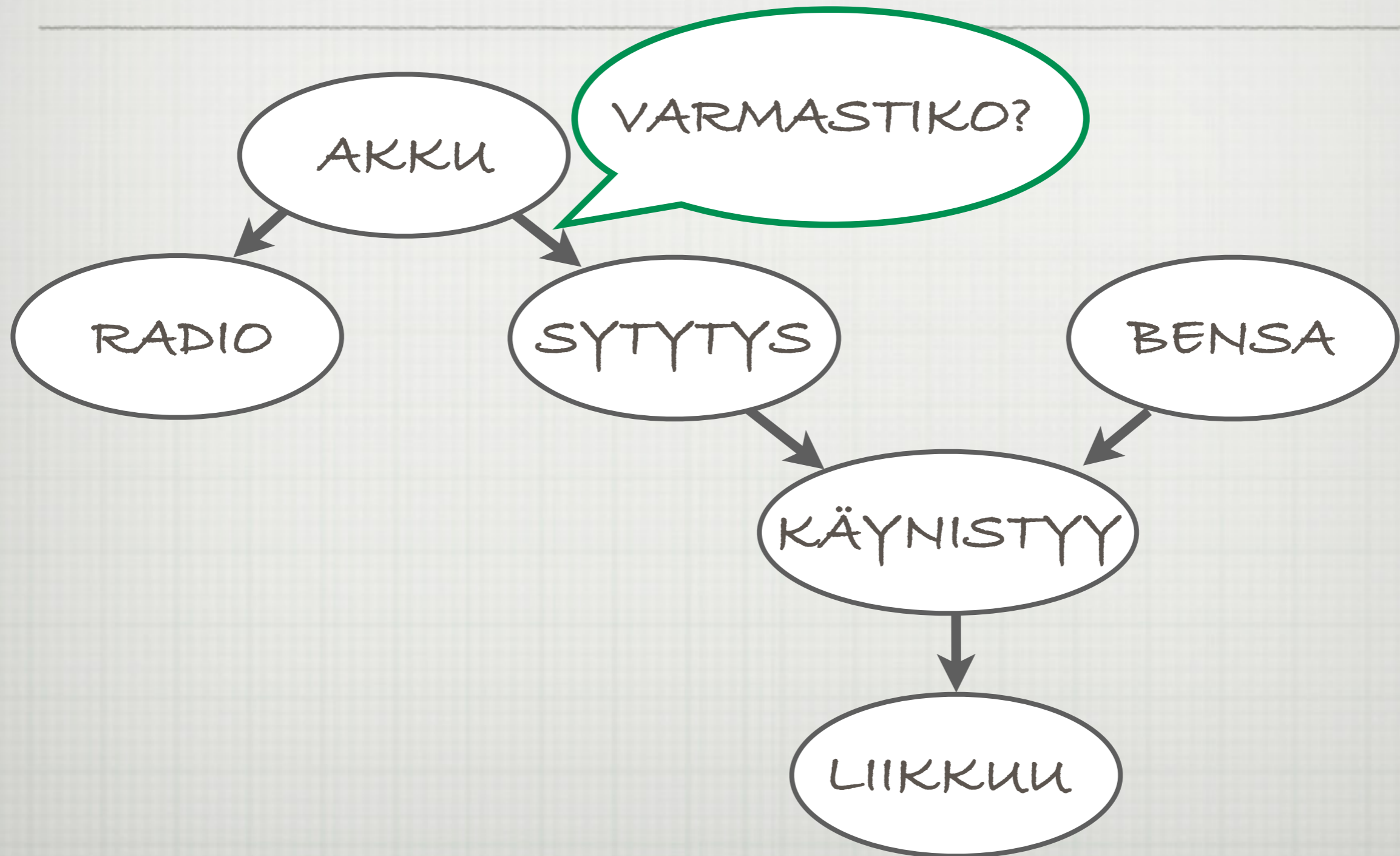


# BAYES-VERKOT



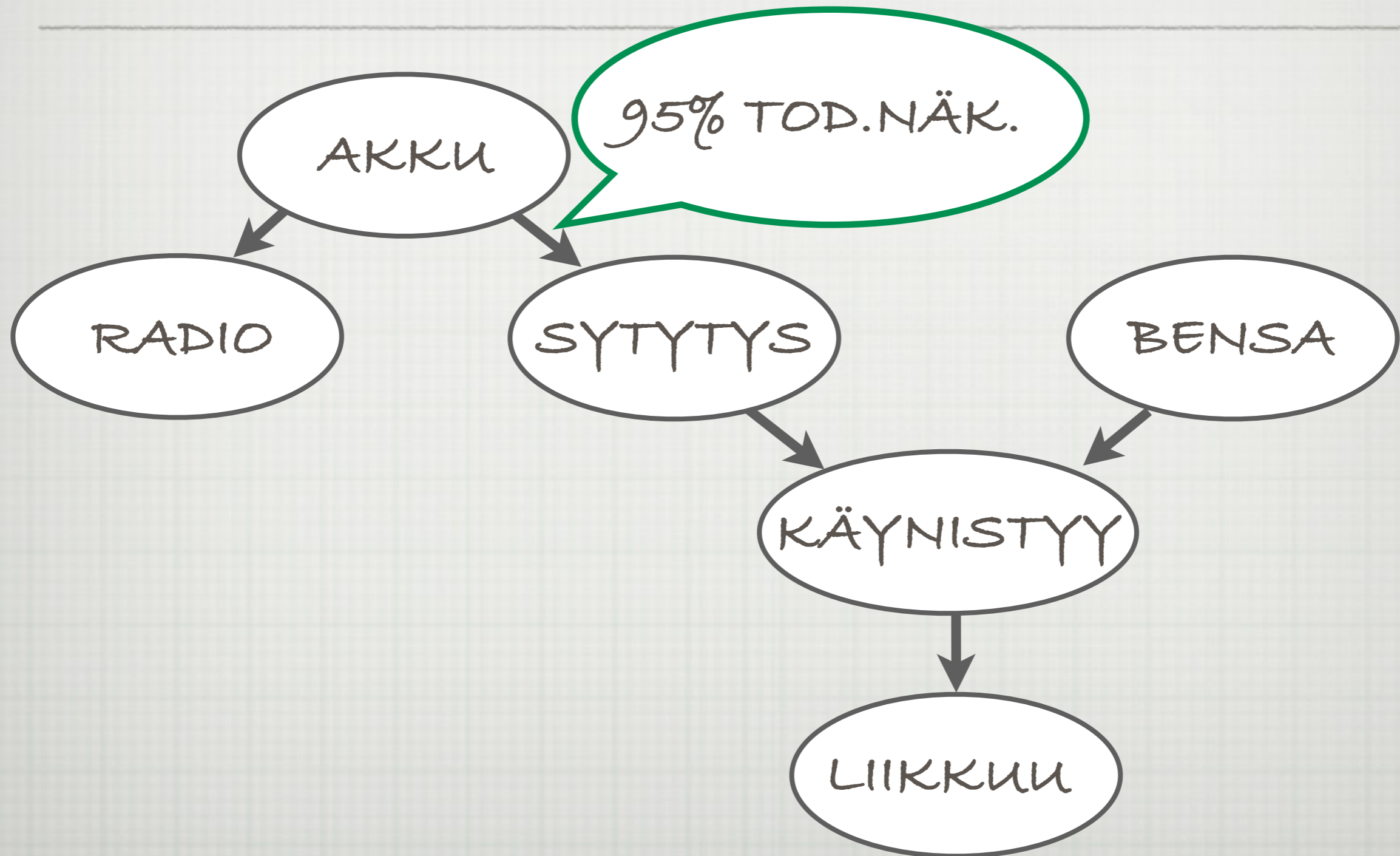
# BAYES-VERKOT

---

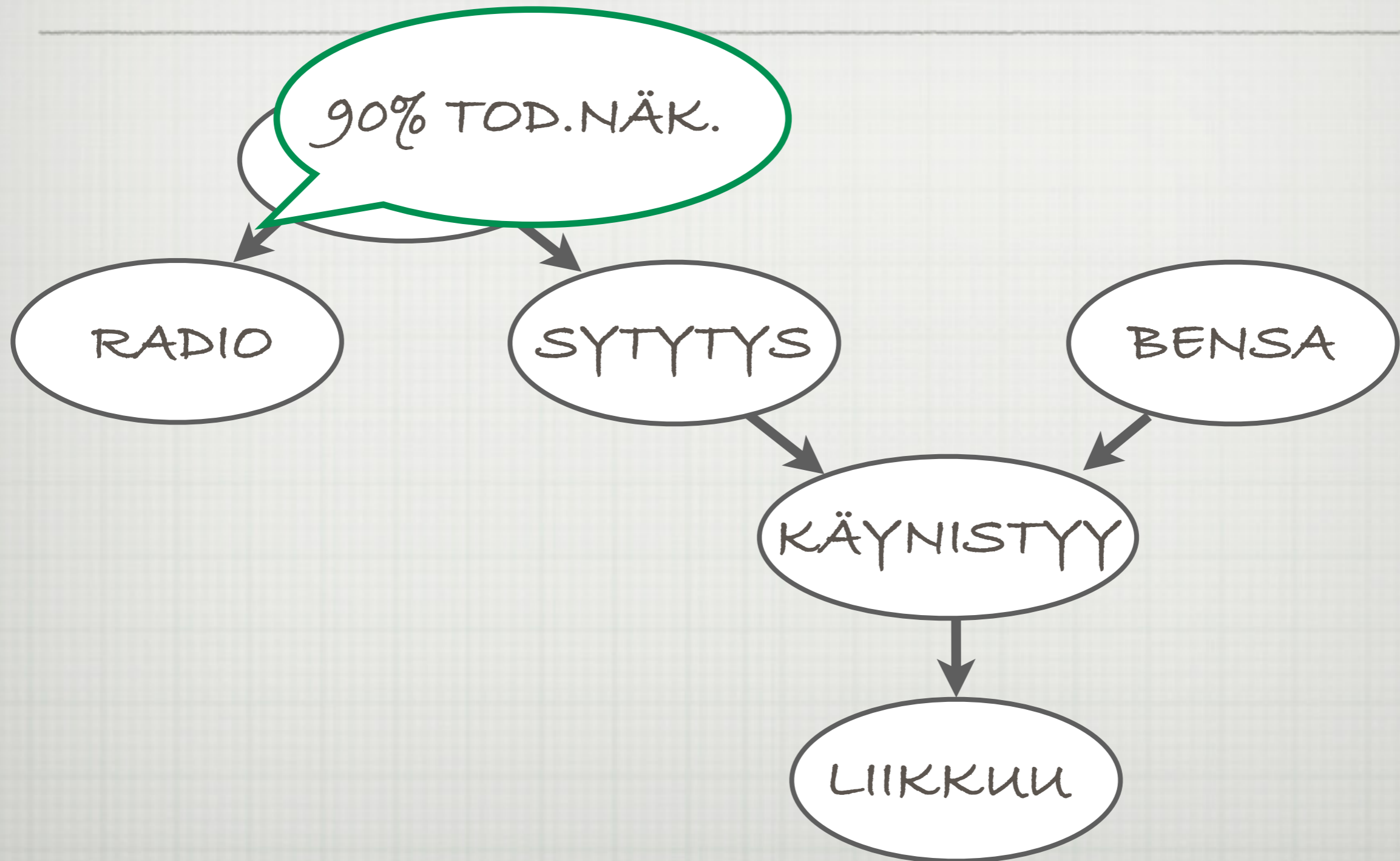




# BAYES-VERKOT



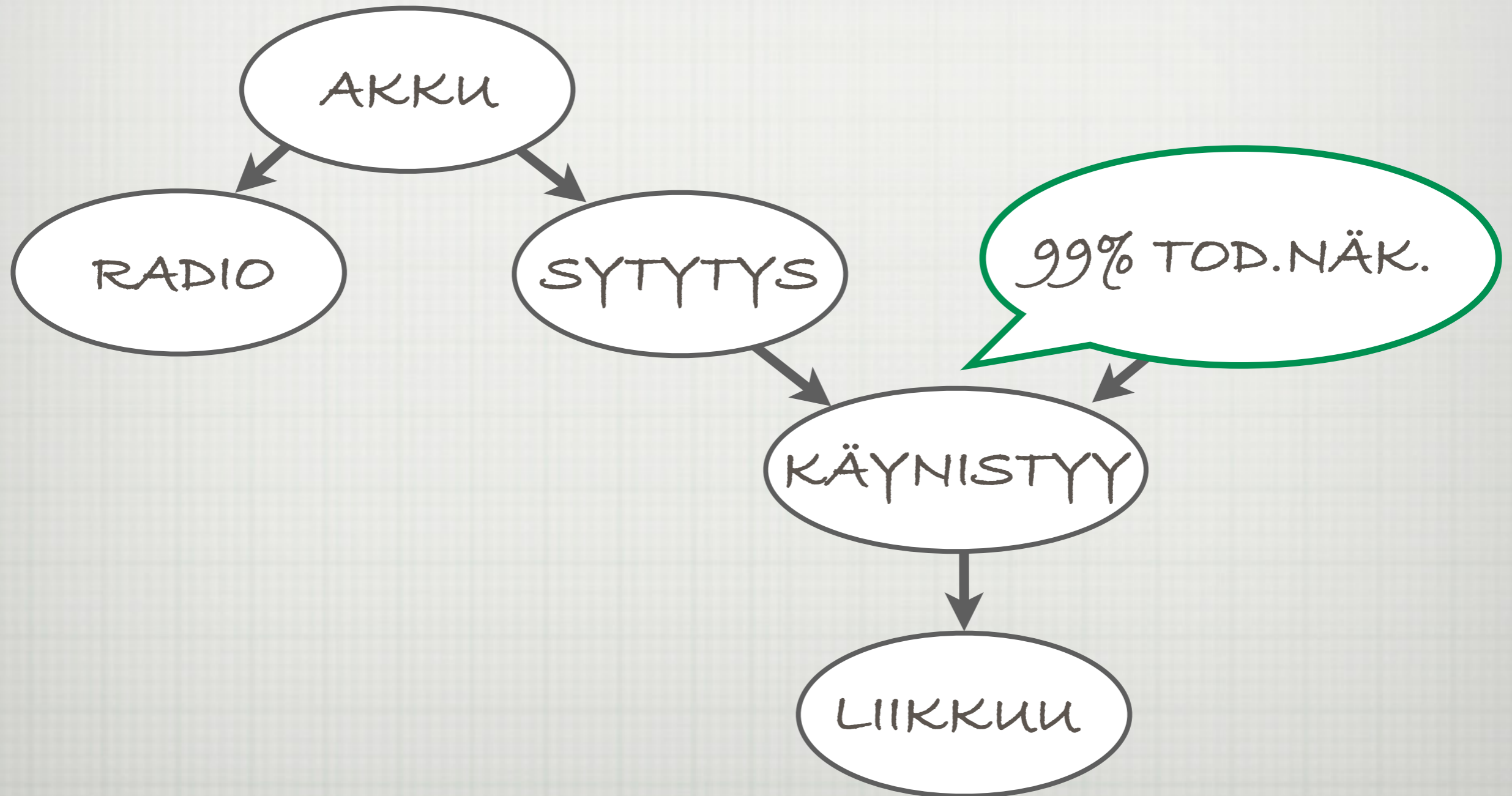
# BAYES-VERKOT





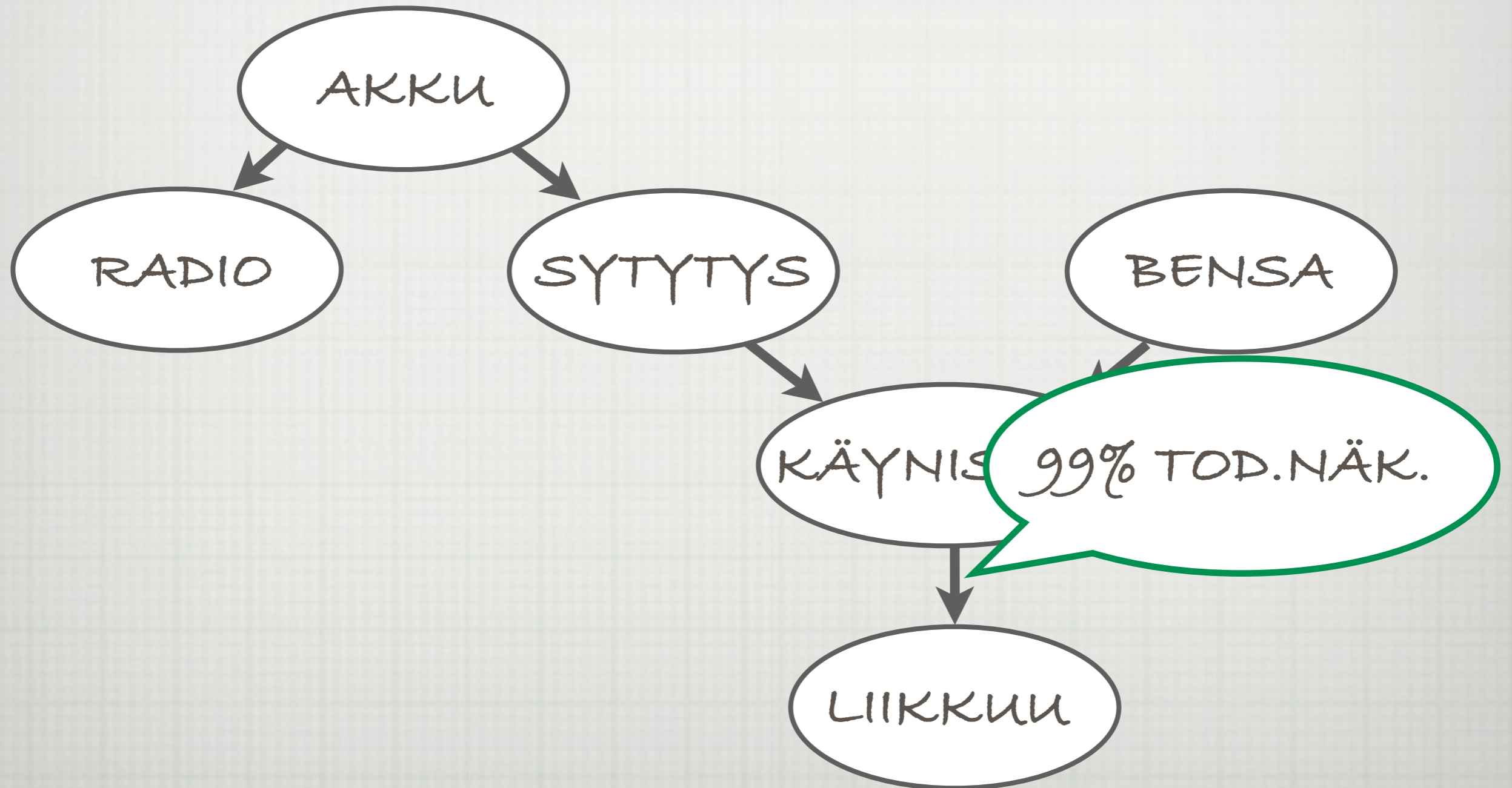
# BAYES-VERKOT

---



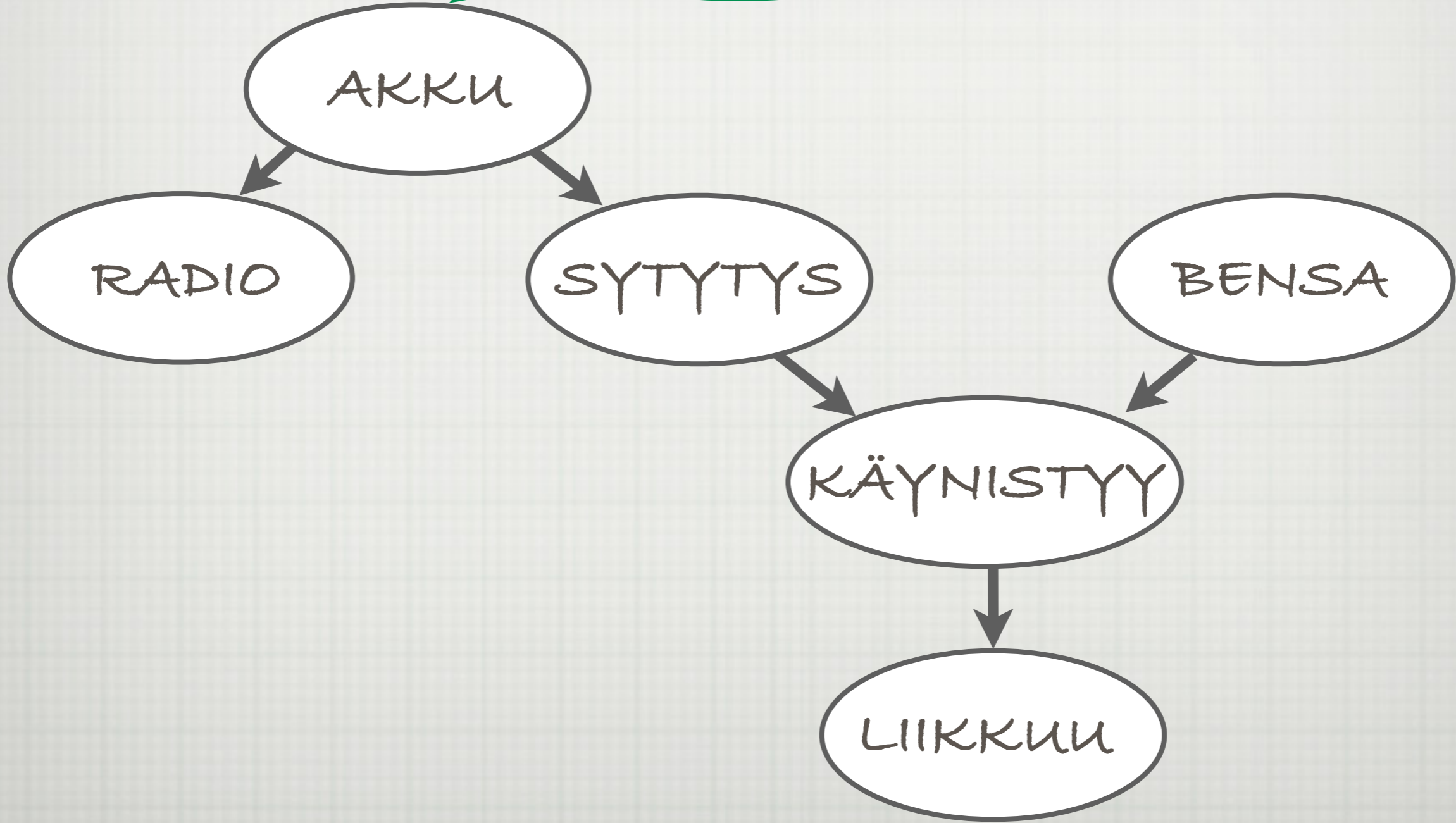
# BAYES-VERKOT

---



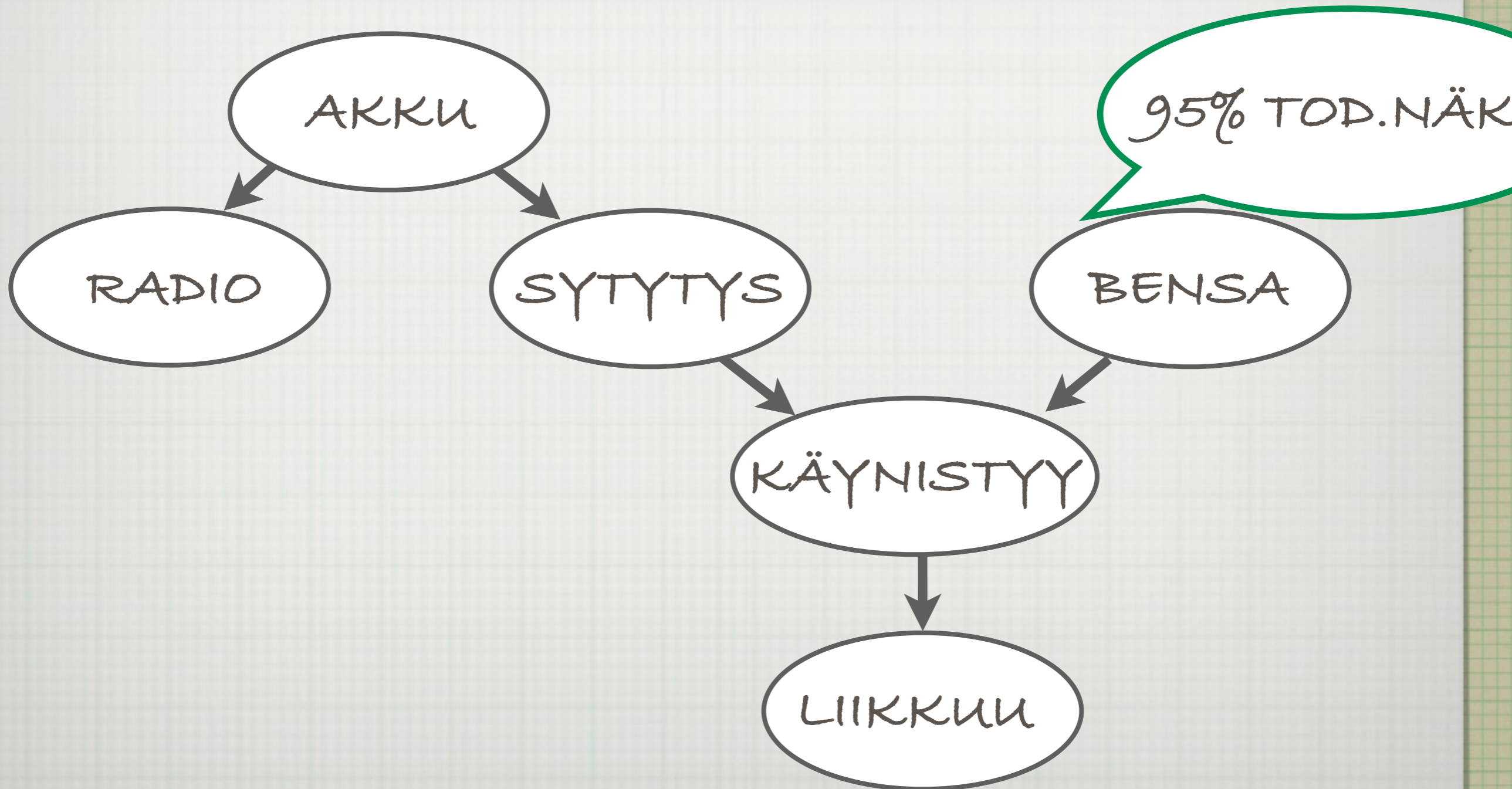


P  
OT  
90% TOD.NÄK.



# BAYES-VERKOT

---





# BAYES-VERKOT

---

- \*  $P(\text{"AKUSSA VIRTAA"}) = 0.9$
- \*  $P(\text{"RADIO"} \mid \text{"AKUSSA VIRTAA"}) = 0.9$   
 $P(\text{"RADIO"} \mid \neg \text{"AKUSSA VIRTAA"}) = 0$
- \*  $P(\text{"SYTYTYYS"} \mid \text{"AKUSSA VIRTAA"}) = 0.95$   
 $P(\text{"SYTYTYYS"} \mid \neg \text{"AKUSSA VIRTAA"}) = 0$
- \*  $P(\text{"BENSAA"}) = 0.95$
- \*  $P(\text{"KÄYNNISTYY"} \mid \text{"SYTYTYYS"} \text{ JA } \text{"BENSAA"}) = 0.99$   
 $P(\text{"KÄYNNISTYY"} \mid \neg \text{"S"} \text{ TAI } \neg \text{"B"}) = 0$
- \*  $P(\text{"LIIKKUU"} \mid \text{"KÄYNNISTYY"}) = 0.99$   
 $P(\text{"LIIKKUU"} \mid \neg \text{"KÄYNNISTYY"}) = 0$

# BAYES-VERKOT

---

\*  $P(\text{"AKKU"} \mid \neg\text{"KÄYNNISTYY"}, \text{"RADIO"}, \text{"BENSAA"}) = ?$

\* "KÄPISTELIJÄN" LÄHESTYMISTAPA:

1. GENEROI MILJOONA TAPAUSTA.
2. VALITSE TAPAUKSET, JOISSA AUTO EI KÄYNNISTY, RADIO SOIJA BENSAA ON.
3. KATSO KUINKA SUURESSA OSASSA AKUSSA ON VIRTAA.



# SATUNNAISAPPROKSIMAATIO

---

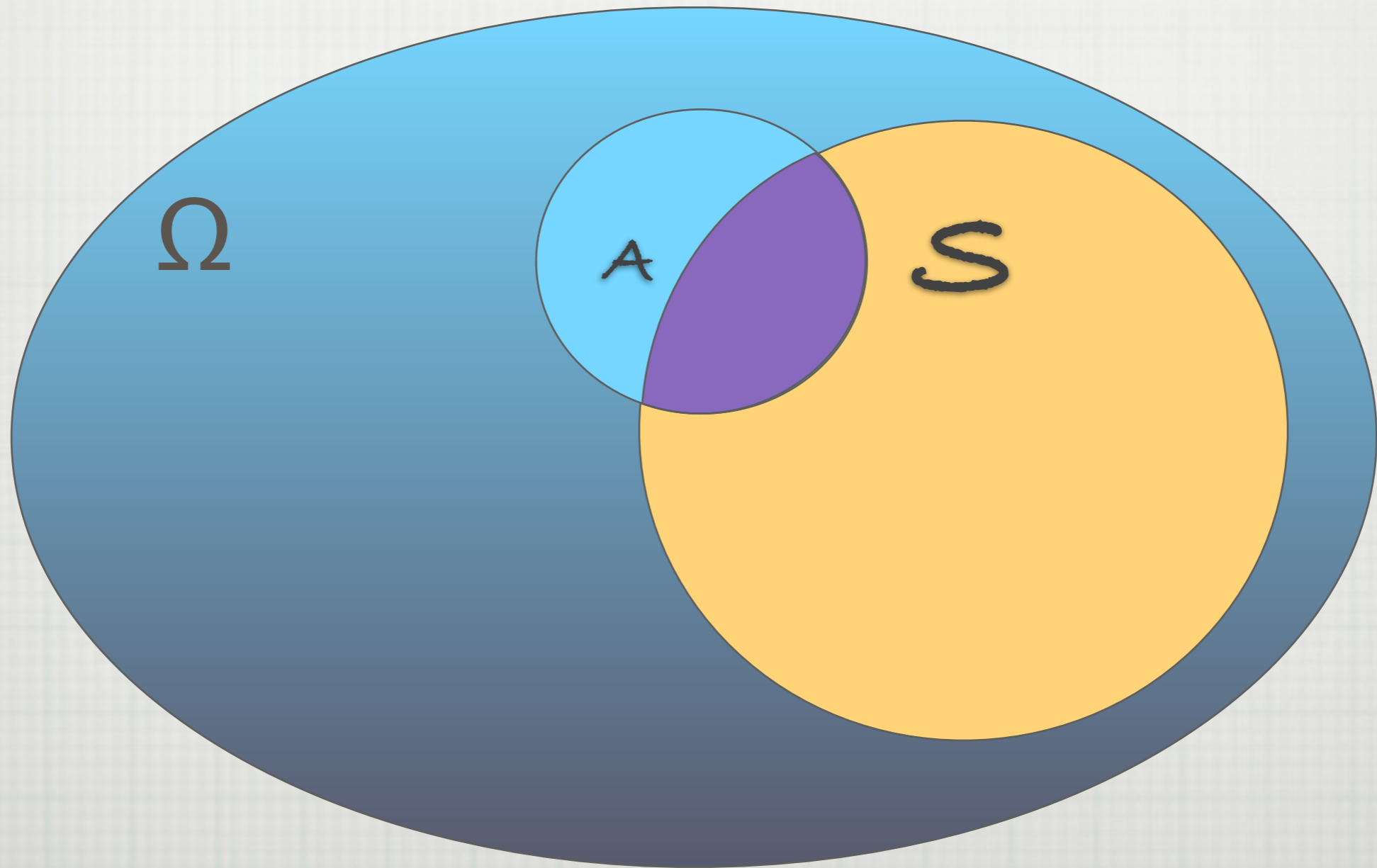
SATUNNAISOTANTAAN PERUSTUVA ("MONTE CARLO-")  
APPROKSIMAATIO:

$$P(S) \approx \frac{\text{"KUINKA MONESSA S"}}{N}$$

$$P(A | S) \approx \frac{\text{"KUINKA MONESSA AJA S"}}{\text{"KUINKA MONESSA S"}}$$

# SATUNNAISAPPROKSIMAATIO

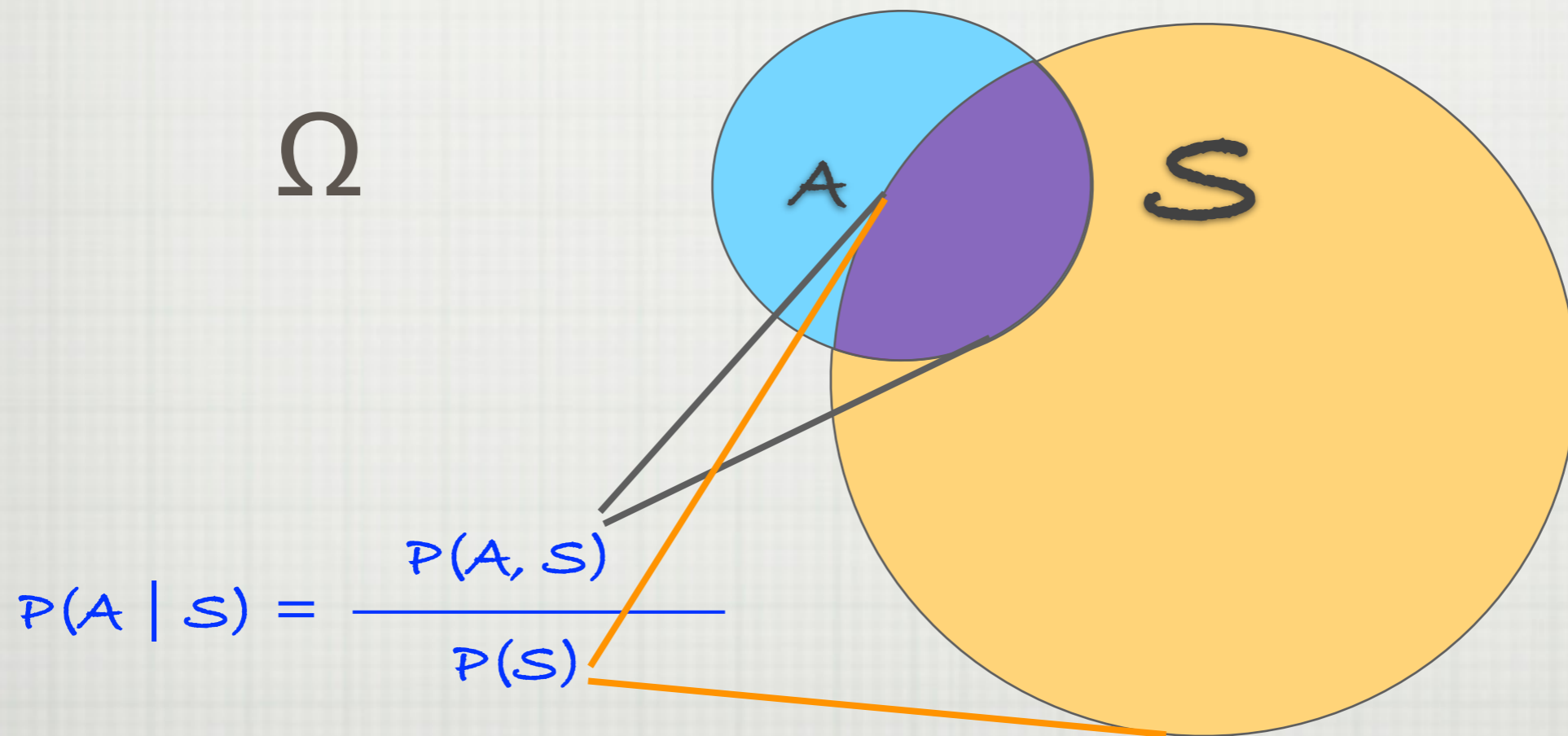
---





# SATUNNAISAPPROKSIMAATIO

---



# SATUNNAISAPPROKSIMAATIO

---

SATUNNAISOTANTAAN PERUSTUVA APPROKSIMAATIO:

$$P(S) \approx \frac{\text{"KUINKA MONESSA S"}}{N}$$

$$P(A | S) \approx \frac{\text{"KUINKA MONESSA AJA S"}}{\text{"KUINKA MONESSA S"}}$$

KUN  $N \rightarrow \infty$ , APPROKSIMAATIO SUPPENEET KOHTI OIKEAA ARVOA.



# BAYES-VERKOT

---

GENEROI-MONIKKOJA(N, Malli):

for i = 1 to N:

for X in Malli.Muuttujat:

X = SATUNNAISLUKU(X.Jakauma(VANHEMMAT(X)))

print X

print '\n'

EHDOLLINEN JAKAUMA



# BAYES-VERKOT

---

GENEROI-MONIKKOJA(N, Malli):

**for** i = 1 to N:

**for** X in Malli.Muuttujat:

X = SATUNNAISLUKU(X.Jakauma(VANHEMMAT(X)))

**print** X

**print** '\n'



SOLMUT, JOISTA  
NUOLI X:ÄÄN



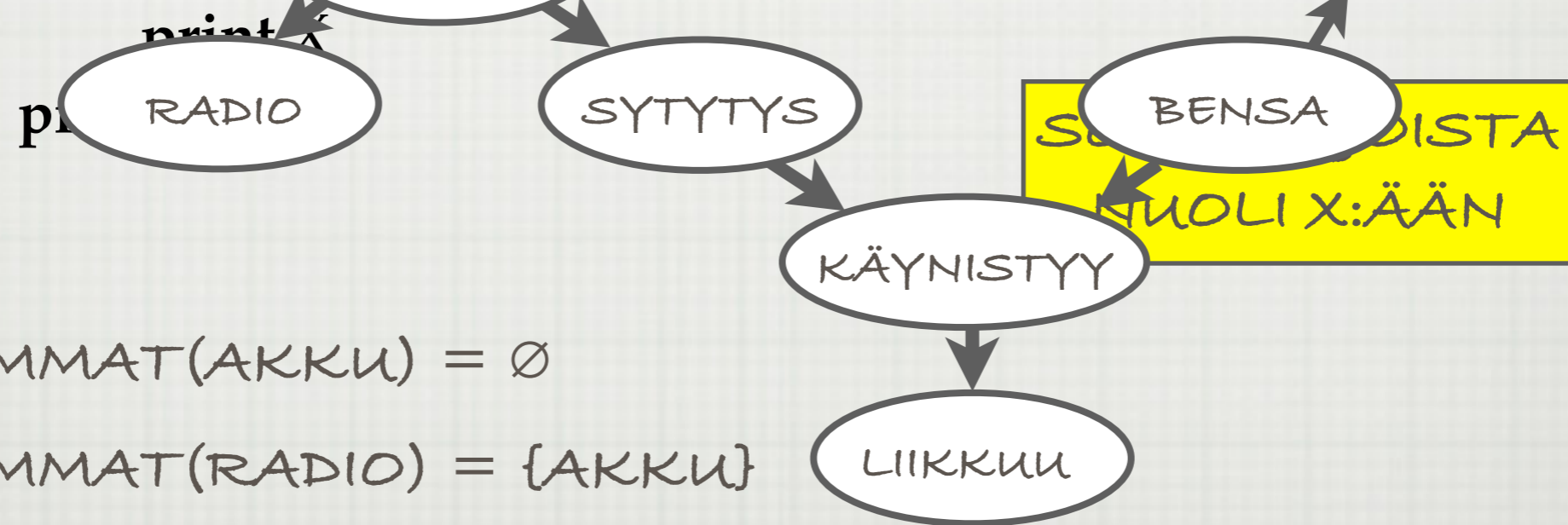
# BAYES-VERKOT

GENEROI-MONIKKOJA(N, Malli):

for i = 1 to N:

for X in Malli Muuttujat:

X = S LUKU(X.Jakauma(VANHEMMAT(X)))



$VANHEMMAT(AKKU) = \emptyset$

$VANHEMMAT(RADIO) = \{AKKU\}$

$VANHEMMAT(SYTYTYYS) = \{AKKU\}$

$VANHEMMAT(KÄYNNISTYY) = \{SYTYTYYS, BENSJA\}$

# BAYES-VERKOT

---

GENEROI-MONIKKOJA(N, Malli):

for i = 1 to N:

for X in Malli.Muuttujat:

X = SATUNNAISLUKU(X.Jakauma(VANHEMMAT(X)))

print X

print '\n'

X = AKKU: X.JAKAUMA( $\emptyset$ ) = (0.1, 0.9)

X = RADIO: X.JAKAUMA(VANHEMMAT(X)) = (1.0, 0.0) JOS  $\neg$ AKKU  
(0.1, 0.9) JOS AKKU

X = K: X.JAKAUMA(VANHEMMAT(X)) = (1.0, 0.0) JOS  $\neg$ S,  $\neg$ B  
(1.0, 0.0) JOS  $\neg$ S, B  
(1.0, 0.0) JOS S,  $\neg$ B  
(0.01, 0.99) JOS S, B



# BAYES-VERKOT

---

\*  $P(\text{"AKKU"} \mid \neg\text{"KÄYNNISTYY"}, \text{"RADIO"}, \text{"BENSAA"}) = ?$

\* EKSAKTI LÄHESTYMISTAPA:

$$P(A \mid \neg K, R, B) = \frac{P(A, \neg K, R, B)}{P(\neg K, R, B)}$$

$$\begin{aligned} P(A, \neg K, R, B) &= P(A, R, S, B, \neg K, L) & \omega_2 \\ &+ P(A, R, S, B, \neg K, \neg L) & \omega_3 \\ &+ P(A, R, \neg S, B, \neg K, L) & \omega_{10} \\ &+ P(A, R, \neg S, B, \neg K, \neg L) & \omega_{11} \end{aligned}$$

\* TAAS SUMMA ALKEISTAPAHTUMIEN TOD.NÄK.!

# BAYES-VERKOT

---

- \*  $P(\text{"AKKU"} \mid \neg\text{"KÄYNNISTYY"}, \text{"RADIO"}, \text{"BENSAA"}) = ?$
- \* EKSAKTI LÄHESTYMISTAPA:
- \* SUMMIEN TERMIT SAADAAN HELPOSTI BAYES-VERKOSTA, ESIM.

$$P(A, R, S, B, \neg K, L) \\ = P(A) P(R|A) P(S|A) P(B) P(\neg K|S, B) P(L|\neg K)$$

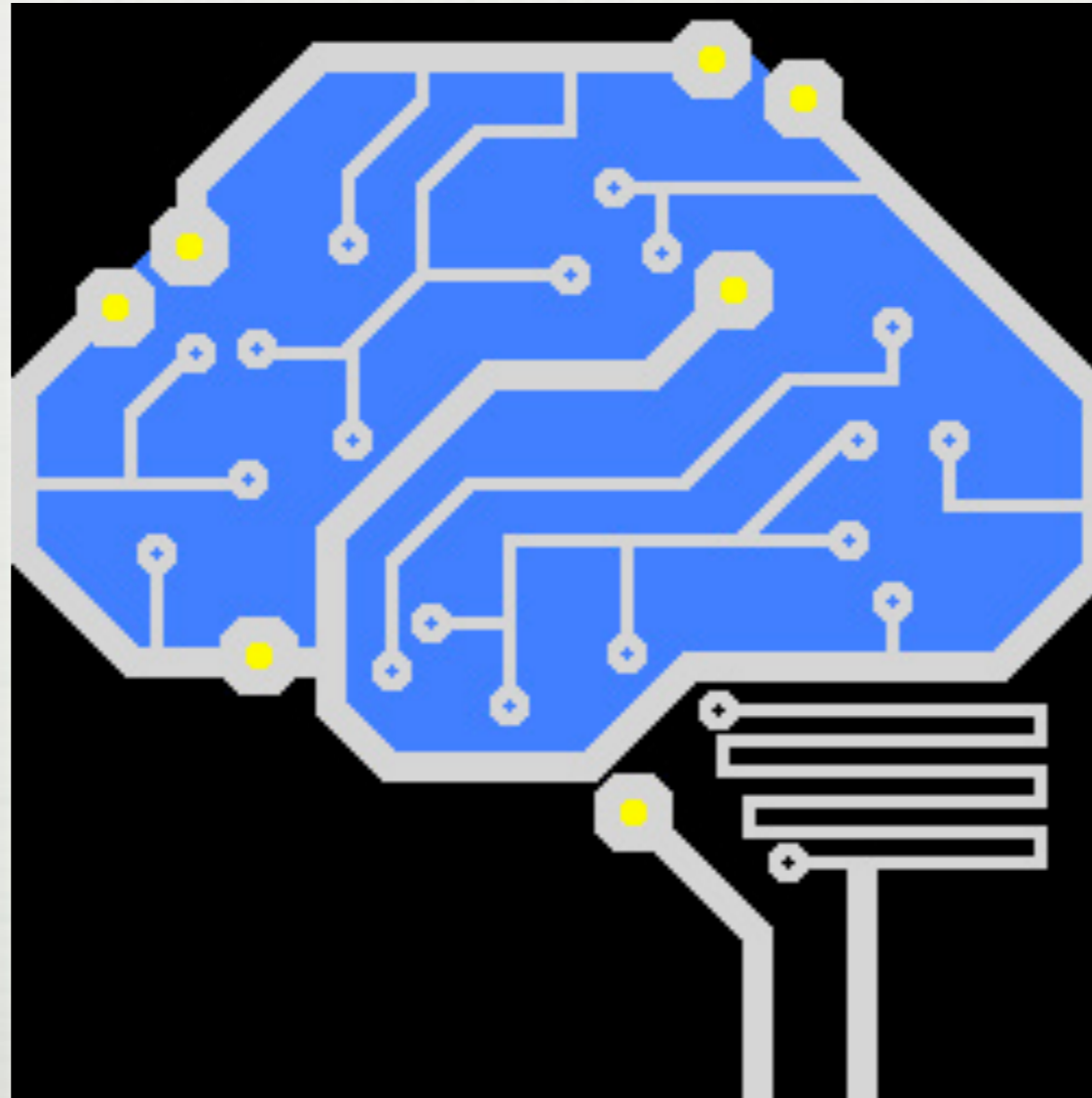
ELI TULO, JONKA TEKIJÖINÄ  $P(\text{MUUTTUJA} \mid \text{VANHEMMAT})$ .

- \* KAIKKI TEKIJÄT ON MÄÄRITELTY ALUSSA.
- \* EI KÄYDÄ LÄPI TARKEMMIN TÄLLÄ KURSSILLA



# KONEOPPIMINEN

---



LINKKI VIDEOON:

INTERVIEW WITH PROF. POGGIO FOR THE MCGOVERN INSTITUTE FACULTY PROFILE. MIT 2009-2010

# KONEOPPIMINEN

---

- \* MÄÄRITELMÄ:

- \* KONE = TIETOKONE, TIETOKONEOHJELMA

- \* OPPIMINEN = ONGELMANRATKAISUKYVYN  
PARANTUMINEN KOKEMUKSEN AVULLA

- \* TOISIN SANOEN...

- \* SEN SIAAN ETTÄ OHJELMOIJA KIRJOITTAÄ TARKAT SÄÄNNÖT  
JONKUN ONGELMAN RATKAISUUN, OHJELMOIJA OHJEISTAA  
TIETOKONETTA OPPIMAAN ESIMERKEISTÄ

- \* MONISSA TAPAUKSISSA TIETOKONEOHJELMA VOI TÄLLÖIN  
OPPIA RATKAISEMAAN JOTAIN TEHTÄVÄÄ PAREMMIN KUIN  
OHJELMOIJA ITSE OSAA!



# MIKSI KONEOPPIMISTA?

---

\* NYKYPÄIVÄNÄ ON HELPPOA JA HALPAA:

- KERÄTÄ DATAA (HALVAT SENSORIT, PALJONJO DIGITAALISTA)
- TALLENTAA DATAA (KOVALEVYTI LA HALPAA)
- LÄHETTÄÄ DATAA (LÄHES ILMAISTA NETISSÄ)

\* JOTEN: KAIKKI KERÄÄVÄT SUURIA MÄÄRIÄ DATAA!

- YRITYKSET: KAUPAT (OSTOTAPAHTUMAT), HAKUKONEET (HAKULAUSEKKEET, VALINNAT), FINANSSISEKTORI (OSAKKEET, VALUUTTAKURSSIT), TEHTAAT (ERILAISET SENSORIT), SOSIAALINEN MEDIA (FACEBOOK, TWITTER, ...), KAIKKI PALVELIMET
- TIEDE: GEENISEKVENSSIT, GEENIEKSPRESSIO, HIUKKASKOKEET FYSIIKASSA, TÄHTITIEDE, ...

□ MUTTA: MITEN HYÖDYNTÄÄ SITÄ?

# ESIMERKKI I

- \* MITEN OHJELMOIDA TIETOKONETTA PELAAMAAN ESIM RISTINOLLAA?

	X	O	X	O	X	O	X	O	X	O	X	O	X
						O		O		O	O	O	O
			X		X		X	X	X	X	X	X	X

- \* VAIHTOEHTO 1: KATSO KOKO PELIPUU LÄPI, VALITSE OPTIMAALISESTI (VAIN HYVIN YKSINKERTAiset PELIT!)
- \* VAIHTOEHTO 2: OHJELMOIJA KIRJOITTAÄ SÄÄNNÖT, TYYLIIN 'JOS VASTUSTAJALLA ON KAHDEN SUORA JA KOLMAS ON VAPAA, ESTÄ SUORA LAITTAMALLA OMA MERKKI SIIHEN', JNE (TYÖLÄSTÄ, VAIKEAA!)
- \* VAIHTOEHTO 3: TIETOKONEOHJELMA KOKEILEE ERILAISIA SÄÄNTÖJÄ, JA PELAA ITSEÄÄN (TAI MUITA) VASTAAN KOKEILLEN MILLÄ SÄÄNNÖILLÄ VOITTAÄ JA MILLÄ SÄÄNNÖILLÄ HÄVITÄÄN, OPPIEN NÄIN VOITTAMAAN ('KONEOPPIMINEN')



# ESIMERKKI I



- \* ARTHUR SAMUEL (50-60 LUVUILLA):
  - \* TIETOKONEOHJELMA OPPII PELAAMAAN TAMMEA
  - \* OHJELMA PELAA ITSEÄÄN VASTAAN TUHANSIA KERTOJA, OPPII MITKÄ POSITIOT OVAT HYVIÄ, MITKÄ HUONOJA (SEN PERUSTEELLA, KUINKA USEIN NE JOHTAVAT VOITTOON/HÄVIÖÖN)
  - \* OHJELMA TULEE LOPULTA PAREMMAKSI KUIN SITÄ 'OPETTAVA' OHJELMOJA

# ASETTELMA

---

ERÄS KONEOPPIMISEN MÄÄRITELMÄ ON ETTÄ TIETOKONE PARANTAA SUORITUSKYKYÄÄN SUORITTAESSAAN JOTAIN TIETTYÄ TEHTÄVÄÄ SITÄ MUKAAN KUN SE NÄKEE ESIMERKKEJÄ.

SIISPÄ TÄYTYY EROTtaa:

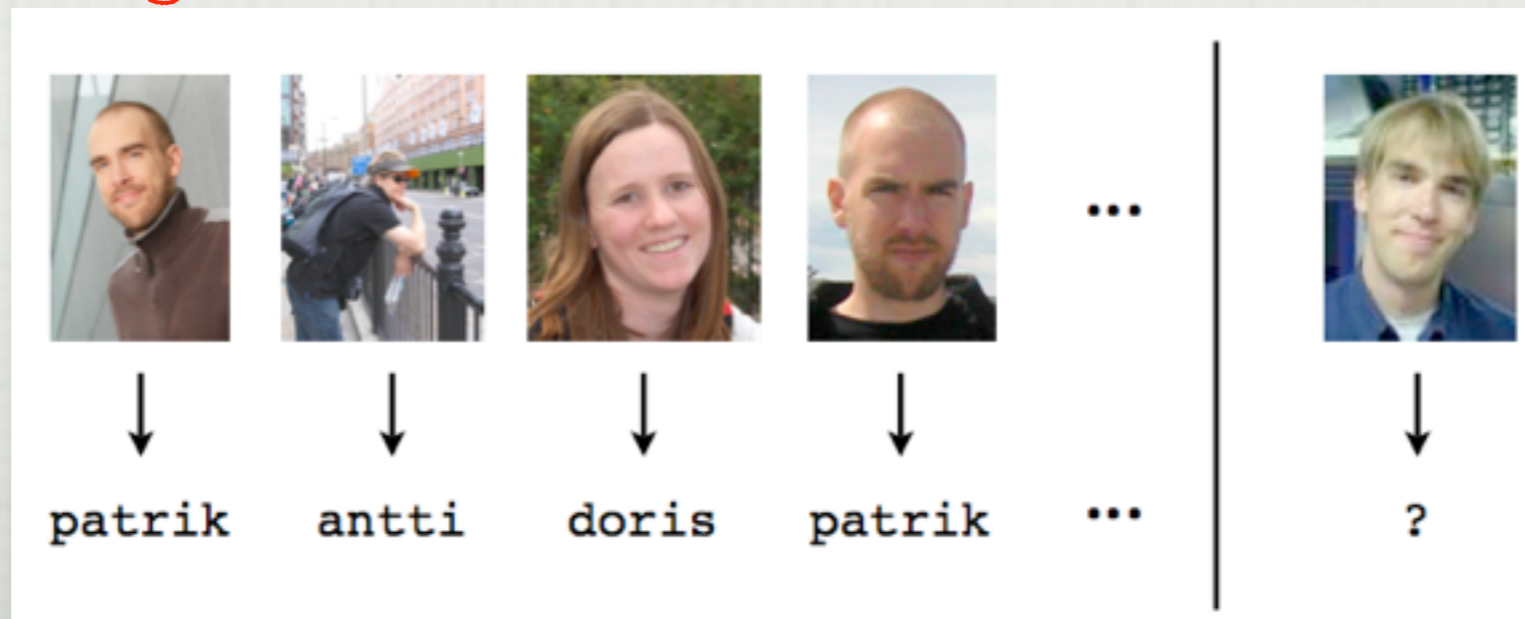
- \* TEHTÄVÄ: MITÄ KONE/OHJELMA YRITTÄÄ TEHDÄ? MINKÄ ONGELMAN SE RATKAISEE?
- \* HYVYYSMITTA: MITEN MITATAAN KUINKA HYVIN KONE/OHJELMA RATKAISEE TEHTÄVÄÄ?
- \* ESIMERKIT/DATA: MIKÄ ON SE KOKEMUSPOHJA/ESIMERKKIDATA JONKA PERUSTEELLA KONE/OHJELMA OPPII?



# ESIMERKKI 2

AUTOMAATTINEN KASVOJENTUNNISTUS (FACEBOOK, APPLE, ...):

- \* OHJELMOJA KIRJOITTAÄ SÄÄNTÖJÄ: 'JOS TUMMA LYHYT TUUKKA, JOSKUS SILMÄLASIT, ISO NENÄ, NIIN SE ON MIKKO' (EI TOIMI! MITEN TUNNISTAA 'TUKAN'? MITEN ARVIOIDA 'NENÄN' KOKO? JNE...)
- \* TIETOKONEELLE NÄYTETÄÄN (KUVA, NIMI)-ESIMERKKIPAREJA, TIETOKONE OPPII ITSE MITKÄ 'PIIRTEET' OVAT RELEVANTTEJA (VAIKEA ONGELMA, MUTTA MAHDOLLISTA JOS RIITTÄVÄSTI ESIMERKKEJÄ!)



# ESIMERKKI 3

ROSKAPOSTIN SUODATUS:

- \* ~~OHJELMOJA KIRJOITTAAN SÄÄNTÖJÄ: 'JOS SISÄLTÄÄ 'VIAGRA', NIIN SE ON ROSKAA'~~ (VAIKEAA, EIKÄ KÄYTTÄJÄLLE RÄÄTÄLÖITY)
- \* KÄYTTÄJÄ OHJEISTAA TIETOKONETTA, MITKÄ POSTIT OVAT ROSKAA, TIETOKONE OPPII ITSE LUOKITTELU-SÄÄNNÖT.

From: medshop@spam.com Subject: viagra cheap meds...	spam
From: my.professor@helsinki.fi Subject: important information here's how to ace the test...	non-spam
:	:
From: mike@example.org Subject: you need to see this how to win \$1,000,000...	?



# ESIMERKKI 4

---

## HAKULAUSEKKEIDEN ENNUSTAMINEN:

- \* ~~OHJELMOJA ANTAA VALMIIN SANAKIRJAN.~~ (SANAT MUUTTUU!)
- \* EDELLISET HAKULAUSEKKEET KÄYTETÄÄN ESIMERKKEINÄ!

## HAKUTULOSTEN JÄRJESTÄMINEN:

- \* ~~OHJELMOJA KIRJOITTA TARKAN KAAVAN JOLLA PISTEYTETÄÄN SANAN ESIINTYMISMÄÄRÄ, LINKKIEN MÄÄRÄ, Y.M. JA SITTEEN JÄRJESTETÄÄN PISTEIDEN MUKAAN.~~ (MISTÄ OHJELMOJA TIETÄÄ MITEN ERI SIVUN OMINAISUUDET PITÄISI PAINOTTA?)
- \* TALLENNETAAN MITKÄ LINKIT VALITAAN MINKÄKIN HAKULAUSEKKEEN JÄLKEEN, JA LAITETAAN SUOSITUIMMAT SIVUT KÄRKEEN (YRITETÄÄN ENNUSTAA MITÄ KÄYTTÄJÄ HALUAA!)

# ESIMERKKI 5

KÄSINKIRJOITETTUIEN MERKKIEN TUNNISTAMINEN:



(POSTIN LAJITTELU, VANHOJEN KIRJOJEN DIGITOINTI, ...)



# ESIMERKKI 6

---

AUTOT JOTKA EIVÄT TARVITSE KULJETTAJAA:



Sebastian Thrun:  
Google's Driverless Car

© TED Conferences, LLC

"...driving itself..."

# ESIMERKKI 7

## SUOSITUSJÄRJESTELMÄT:

\* AMAZON.COM: "JOS OSTIT TÄMÄN KIRJAN, SAATAT MYÖS OLLA KIINNOSTUNUT TÄSTÄ TOISESTA KIRJASTA"

\* NETFLIX.COM: KÄYTTÄJÄT ARVIOIVAT ELOKUVIA, JÄRJESTELMÄ EHDOTTAÄ SEN POHJALTA KÄYTTÄJILLE UUSIA ELOKUVIA ('NETFLIX PRIZE': \$1.000.000 PALKINTO, VUODET 2006-2009)

	Seven	Fargo	Aliens	Leon	Avatar
Linda	4	5	5	1	2
Jack	1	4	1	5	1
Bill		?	4	1	?
Lucy		2	1	5	
John	1		1	4	5
	4			5	5
	2	3			3

[HTTP://WWW.YOUTUBE.COM/WATCH?  
FEATURE=PLAYER\\_DETAILPAGE&V=IMPV7OULXYW](http://www.youtube.com/watch?feature=player_detailpage&v=impv7oulxYW)



# ESIMERKKI 8

---

KONEKÄÄNTÄMINEN:

- \* PERINTEINEN TAPA: SANAKIRJA JA KIELIOPPI
- \* NYKYÄÄN (ENEMMÄN JA ENEMMÄN): TILASTOLLINEN KONEKÄÄNTÄMINEN, JOKA PERUSTUU ESIMERKKEIHIN/DATAAN

Google käänntäjä

Kielest4: suomi ▼ ↕ Kielelle: englantti ▼ Käänn4

Tietojenk4sittelytieteen opinnot antavat erinomaisen pohjan ty4skentelylle kaikkialla, miss4 kehitet4n tai sovelletaan tietotekniikkaa.

K4nn4s (suomi > englantti)

Computer studies provide an excellent foundation for the work, wherever applicable, or to develop information technology.

# ESIMERKKI 9

'KAVEREIDEN' EHDOTTAMINEN:

- \* VOIKO FACEBOOK KAVERIGRAAFIN PERUSTEELLA ARVATA TUNTEVATKO KAKSI HENKILÖÄ TOISENSA VAIKO EI?
- \* OSAAKO TWITTER ARVATA KETÄ OLISIT KIINNOSTUNUT 'SEURAAMAAN'?

