

Koolitus „Andmeteadus on Popp“

# Autonoomsed otsustusesüsteemid

November 2021

Kristjan Eljand



Euroopa Liit  
Euroopa Sotsiaalfond



Eesti  
tuleviku heaks

# Päevakava

---

Teema

Sisu

**Optimeerimine**

Kuidas leida matemaatiliselt parim lahendus probleemile.

**Stimulõpe**

Kuidas luua iseõppivat süsteemi?

---

# Optimeerimine

Ehk kuidas peaksin käituma, et tulemus oleks minimaalse kuluga või maksimaalse tuluga või kiireim või ...

# Optimeerimisülesanded

- **Tootmispinna planeerimine** – kuidas planeerida tehase sisseseade selliselt, et tootmisaeg oleks minimaalne?
- **Tootmise planeerimine** – Mida järgmisel nädalal toota, et maksimeerida kasum, võttes arvesse turunõudlust ja erinevatele toodetele kuluvat ressursi?
- **Ladude asukohtade planeerimine** – Kuhu ehitada laohooned, et kliendid saaksid väikseima kuluga teenindatud.
- **Õige segu leidmine** - Millistest toorainetest segada minimaalse kuluga kokku toode, mis vastab kliendi nõuetele. (lammastele sobiva sööda segamine olemasolevatest toorainetest.)

# Optimeerimisülesanded

- **Jaotusprobleem (*assignment problem*)** – Kuidas jaotada ülesanded ressursside vahel. Näide: kolm taksot ja kolm klienti – kuidas jaotada nii, et kulu oleks minimaalne.
- **Ülejäägi minimeerimine (*cutting stock problem*)** – Kuidas lõigata materjalid, kasutada toorainet selliselt, et jäägid oleksid minimaalsed.
- **Võrgustiku ühendamise probleem (*dominating set*)** – Kuhu seada ühendused, et kogu võrgustik saaks kaetud ja ühenduspunktide arv oleks minimaalne? Näide: kuhu paigaldada vahealajaamad, nii et iga jaotusalajaam oleks ühendatud neist vähemalt ühega!?

# Optimeerimise praktiline näide

Ehk kuidas peaksin käituma, et tulemus oleks minimaalse kuluga või maksimaalse tuluga või kiireim või ...



Euroopa Liit  
Euroopa Sotsiaalfond



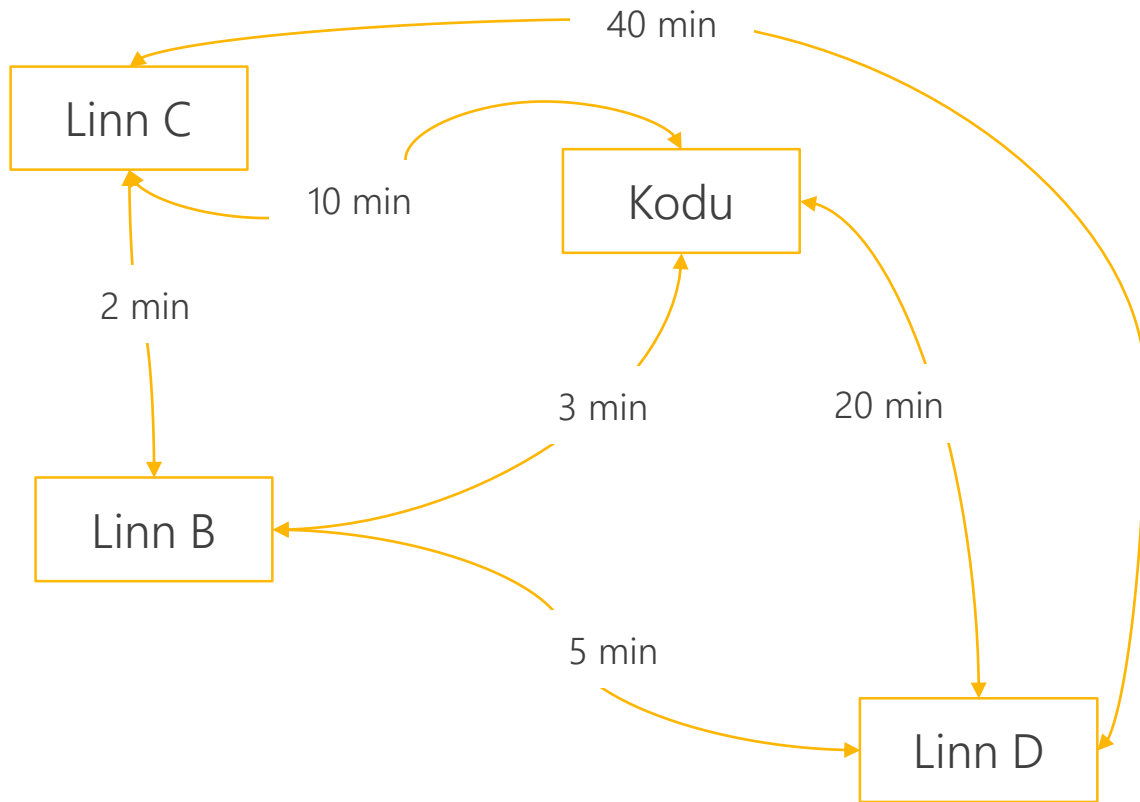
Eesti  
tuleviku heaks

# Stiimulõpe (*reinforcement learning*)

Tehnoloogia, mis õpib tegema otsuseid.

# Reisiva müügimehe ülesanne

Anname ette a.) keskkonna, b.) tegevused, mida keskkonnas saab teostada ning palume mudelil leida optimaalse tegevusplaani.

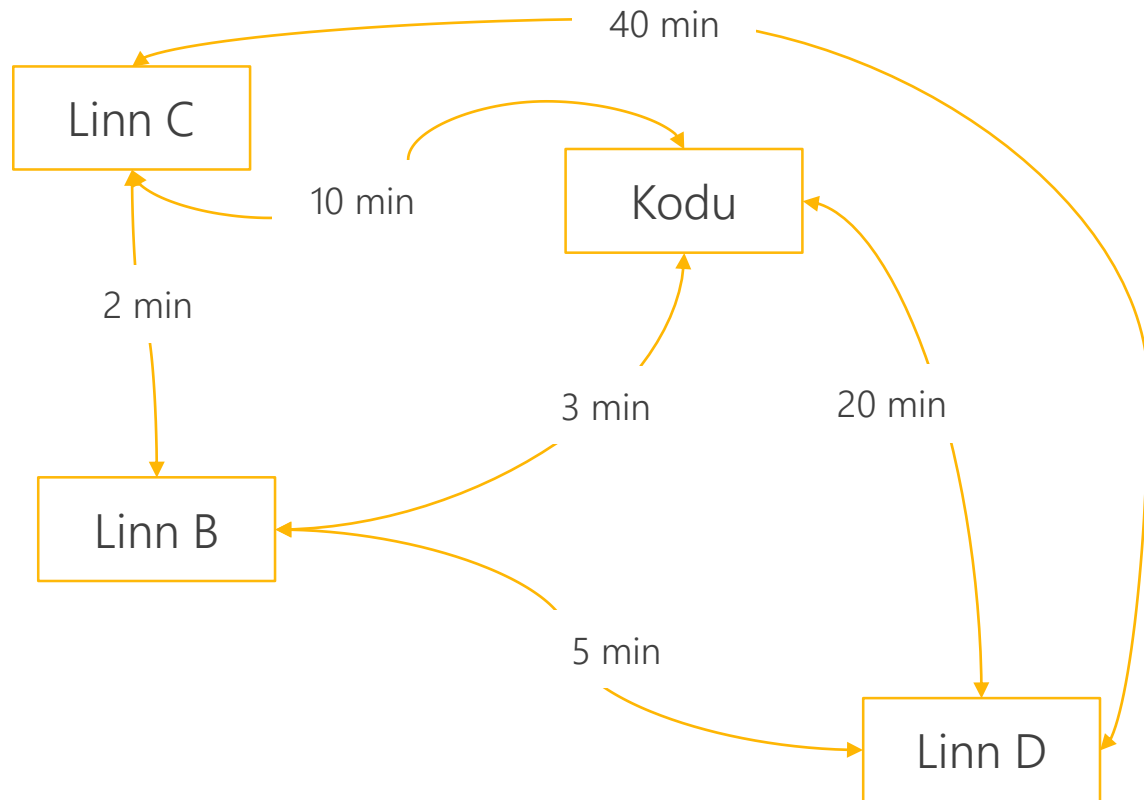


1. Keskkond: Liigu ühest linnast teise.
2. Tegevused: Kodu, B, C
3. Preemia: Ajakulu ühest linnast teise liikumiseks.

**Eesmärk:** Õppida selgeks, kuidas käia läbi kõik linnad ja jõuda koju minimaalse ajakuluga.



# Reisiva müügimehe ülesanne – Naiivne lahendus



Naiivne strateegia: mine alati ajaliselt lähimasse punkti, mida pole veel külastanud.

Kodu -> Linn B = 3 min

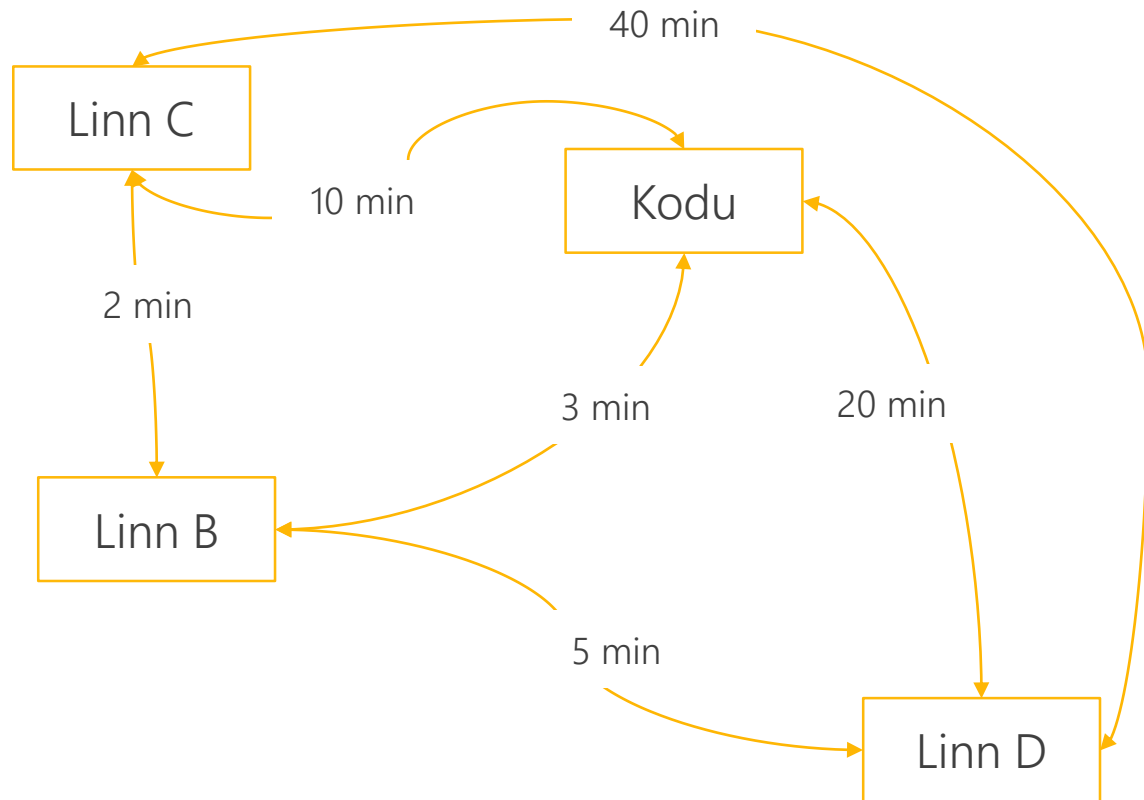
Linn B -> Linn C = 2 min

Linn C -> Linn D = 40 min

Linn D -> Kodu = 20 min

**Kokku = 65 min**

# Reisiva müügimehe ülesanne – Mõtle samm ette



Mõtle samm ette: kaalun ka järgmise sihtkoha kaugust.

Kodu -> Linn C = 10 min

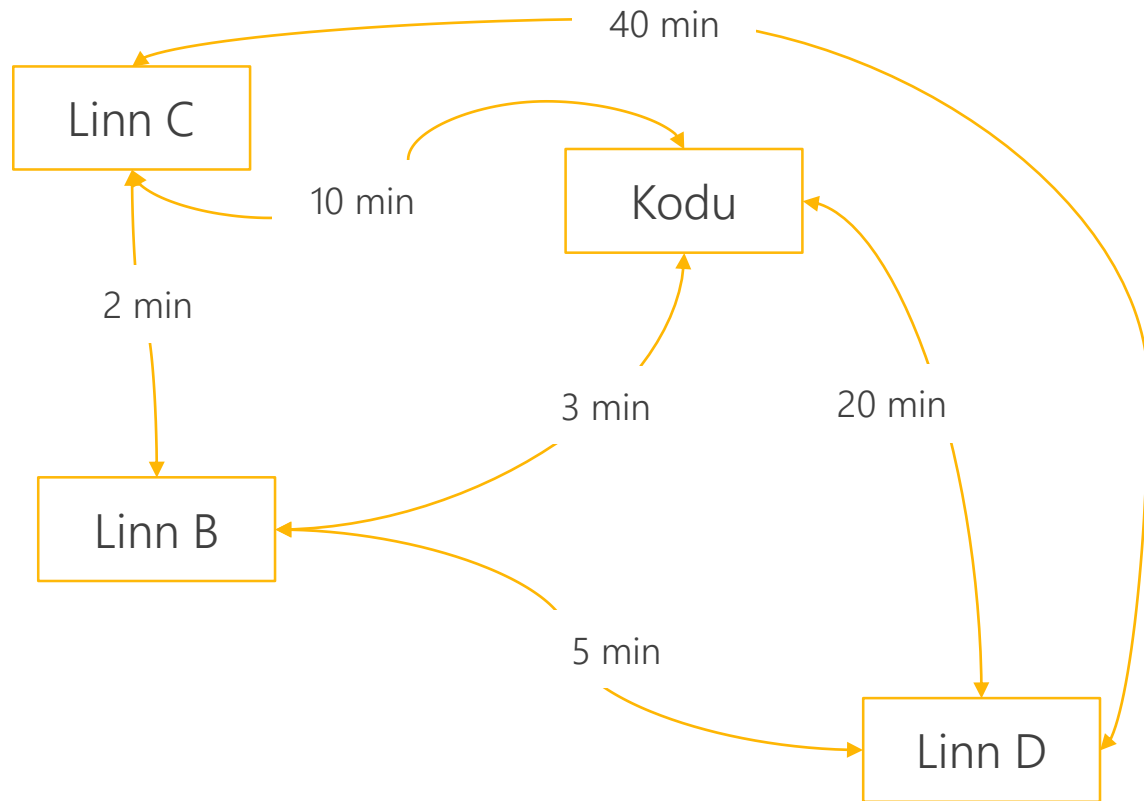
Linn C -> Linn B = 2 min

Linn B -> Linn D = 5 min

Linn D -> Kodu = 20 min

**Kokku = 37 min**

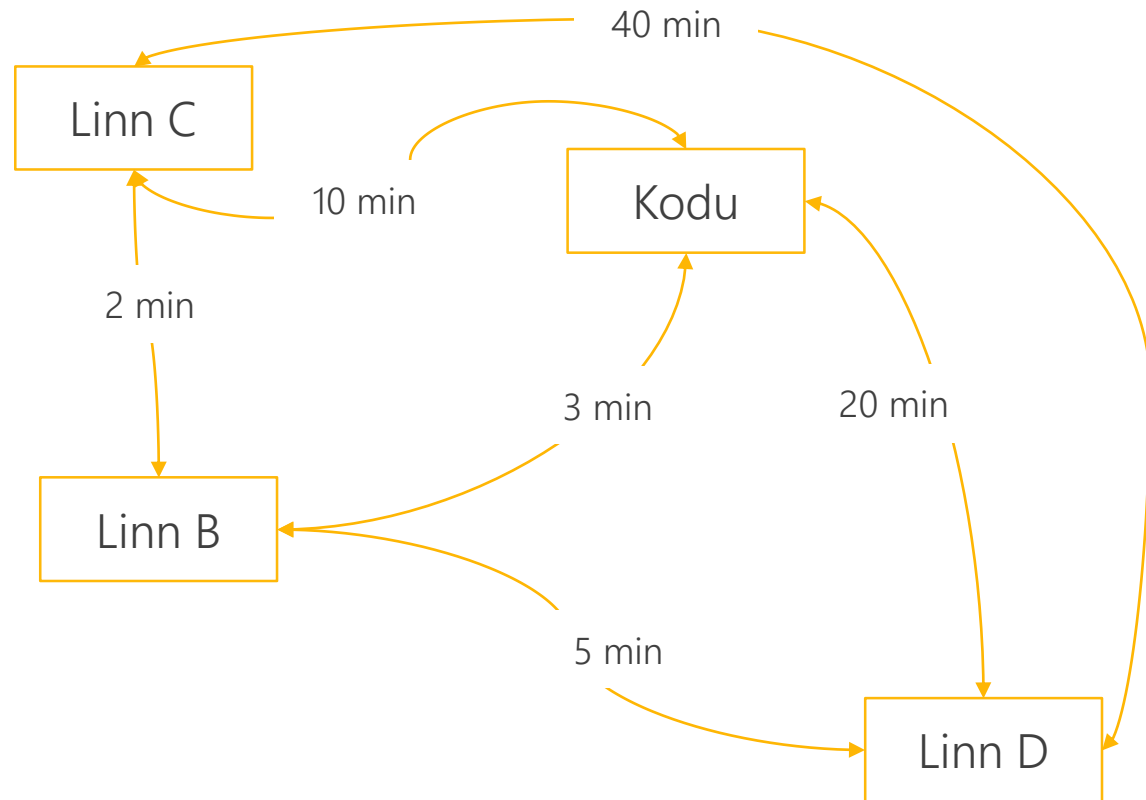
# Reisiva müügimehe ülesanne – Samm tagasi



Samm tagasi: äkki on mõistlik vahepeal minna samasse linna, kust juba tulnud.

Kodu -> Linn B = 3 min  
Linn B -> Linn C = 2 min  
Linn C -> Linn B = 2 min  
Linn B -> Linn D = 5 min  
Linn D -> Kodu = 20 min  
**Kokku = 32 min**

# Reisiva müügimehe ülesanne – Lahendus



Optimaalne lahendus:

Kodu -> Linn B = 3 min

Linn B <-> Linn C = 2 + 2 min

Linn B <-> Linn D = 5 + 5 min

Linn B -> Kodu = 3 min

**Kokku = 20 min**

# Mida stiimulõpe õpib!?

|   |  |
|---|--|
| <b>Hetkeseis:</b><br>Asukoht: Kodu<br>Ajalugu:                                  | Lahendamise käigus treeniti masinõppemudel, mis võtab sisendiks hetkeseisu ja prognoosib kõigi tegevuste „pikaajalise“ väärtuse.                 |
| <b>Tegevuse väärtus:</b><br>Kodu: 120<br>Linn B: 19<br>Linn C: 24<br>Linn D: 32 | Näiteks algseisust (Kodu) prognoosib mudel Linna B mineku pikaajaliseks väärtuseks 19 minutit ( <i>teame, et tegelik väärtus on 20 minutit</i> ) |

# Mida stiimulõpe õpib!?

|   |   |   |   |
|---|---|---|---|
| <b>Hetkeseis:</b><br>Asukoht: Kodu<br>Ajalugu:                                  | <b>Hetkeseis:</b><br>Asukoht: Linn B<br>Ajalugu: Kodu                           | Järgmise otsuse tegemiseks anname ette uue hetkeseisu: Asukoht on Linn B ja varasemalt oleme külastanud Kodu. |   |
| <b>Tegevuse väärtus:</b><br>Kodu: 120<br>Linn B: 19<br>Linn C: 24<br>Linn D: 32 | <b>Tegevuse väärtus:</b><br>Kodu: 24<br>Linn B: 117<br>Linn C: 16<br>Linn D: 17 | <b>Tegevuse väärtus:</b><br>Kodu: 27  | Sellest seisust prognoosib mudel, et kõige väiksem ajakulu tekib, kui liigume linna C (16 minutit). |

# „Toore jõuga“ optimeerimine vs Stiimulõpe

Traditsioonilise „toore jõu“ meetodi korral **proovib** arvuti läbi kõik variandid ning ütleb, mis on parim. 4 linna puhul on see lihtne, kuid variantide arvu suurenemisel muutub see kiiresti võimatuks!

**Kodu>B>C>D>... = 120 min**  
**Kodu>A>C>B>... = 90 min**  
**Kodu>B>C>B>... = 35 min**  
**Kodu>D>C>A>... = 62 min**  
...

„Toore jõu“! korral ei õpi me midagi! Iga kord tuleb optimeerimine uuesti teostada.

Stiimulõppe puhul **õpib** arvuti prognoosima tegevuse väärtust igast **hetkeseisust**.

**Hetkeseis:**  
Asukoht: Kodu  
Ajalugu:

**Tegevuse väärtus:**  
**Kodu: 120**  
**Linn B: 19**  
**Linn C: 24**  
**Linn D: 32**

Treenitud mudelilt saab küsida prognoosi väga kiiresti ükskõik mis hetkeseisu kohta.

# Stiimulõppe rakendusvaldkonnad

- Finants - Väärtpaberiinvesteering ja –kauplemine.
- Robotika - Autonoomsed masinad.
- Turundus - Turunduse automatiseerimine.
- Kaubandus - Hoone energiakasutuse optimeerimine.
- Tootmine - Tehase töö optimeerimine.
- Logistika – masinate liikumise optimeerimine.
- Personalijuhtimine - Töögraafikute optimeerimine.
- Vestlusrobot - Chatbot, mis arvestab hetkeseisu ja lausete pikaajalist mõju.
- Meditsiin – Parim võimalik raviplaan patsiendi hetkeseisu arvestades.
- Inseneriteadus ja disain – uute disainide loomine, mis võimaldavad saavutada maksimaalse tulemuse.
- **Aga tegelikult ... Igas otsuseprotsessis!!!**



# Mida on saavutatud: Rubiku kuubiku lahendamine



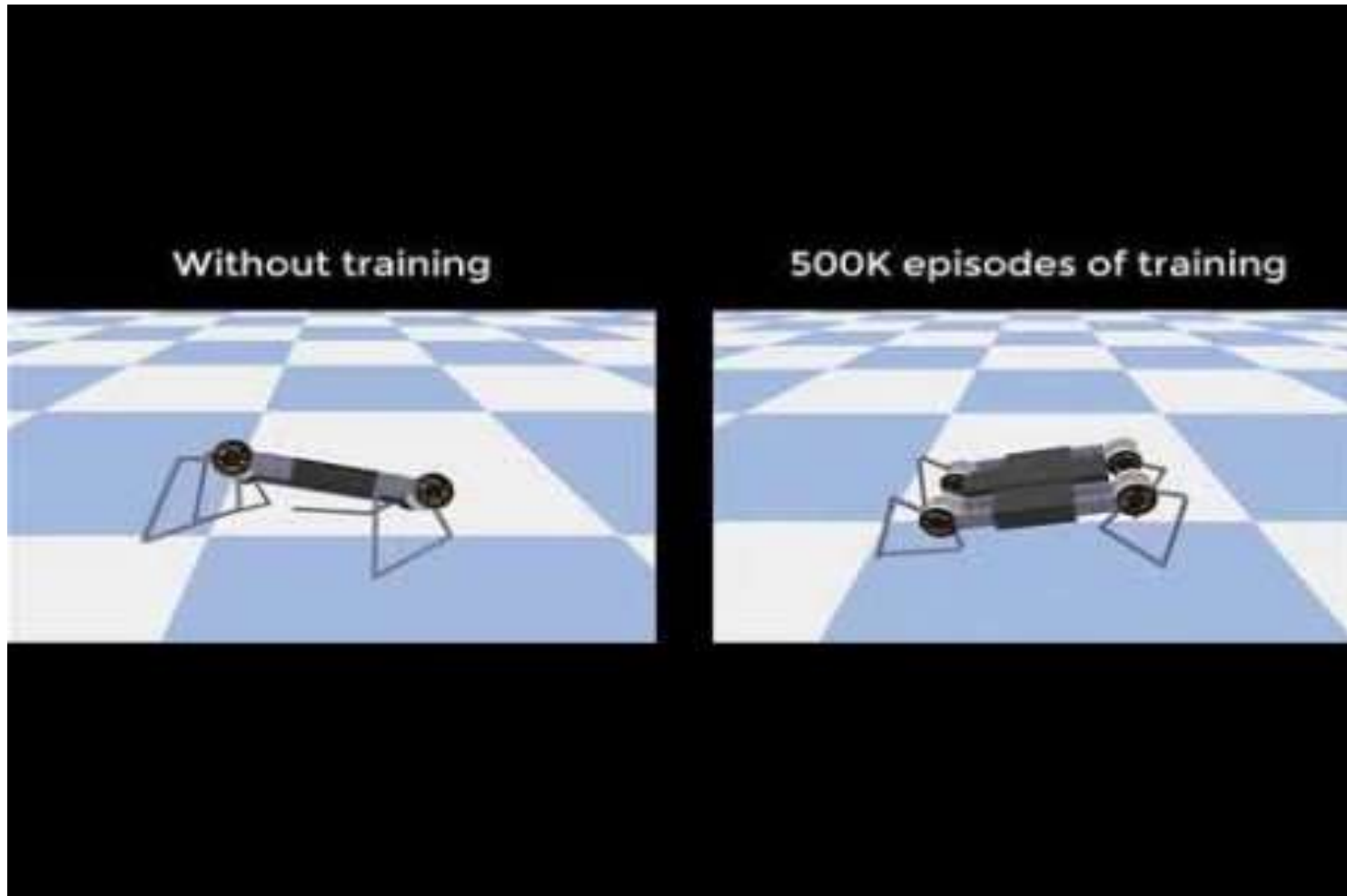
2019. aastal näitas OpenAI, et stiimulõppega treenitud agent on suuteline lahendama Rubiku kuubiku.

# Mida on saavutatud: Mitme agendi koostöö



Näide, kuidas mitu digitaalset agendi suudavad õppida tegema koostööd.

# Mida on saavutatud: Treenime roboti kõndima

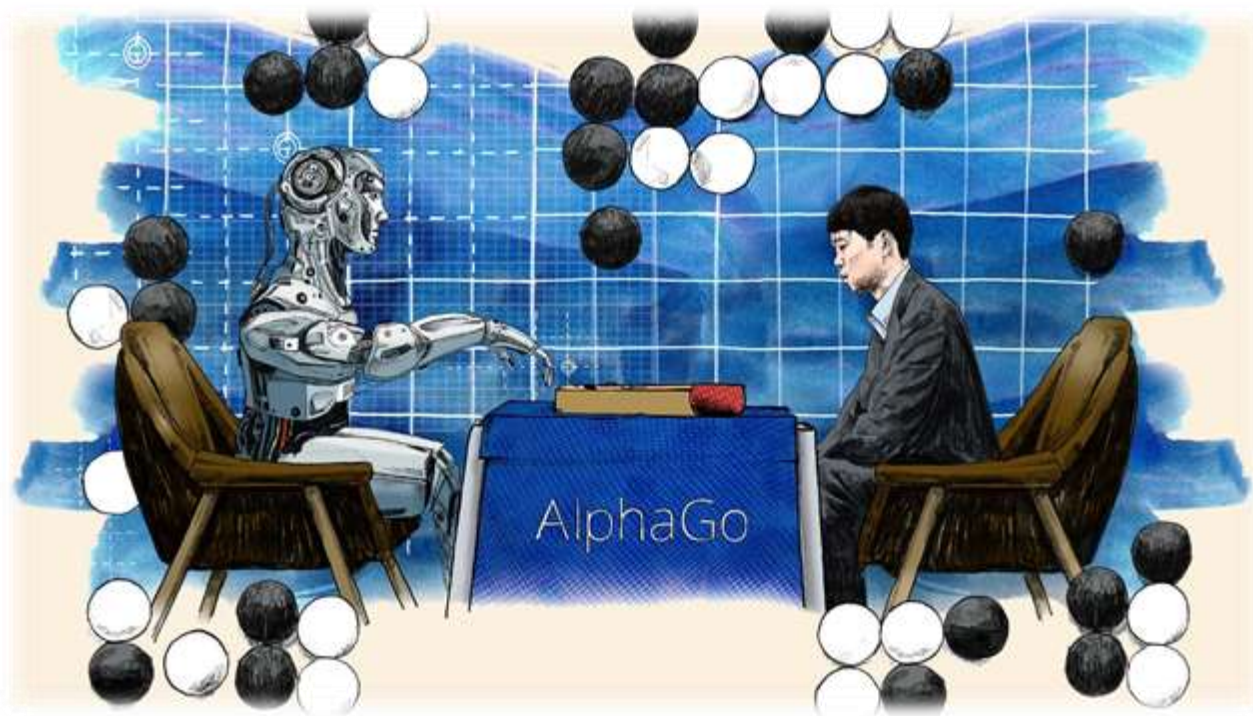


**Minitaur:** Neljajalgne kõndiv robot.

**Tegevusruum:** Saame igal ajahetkel määratleda 8 mootori nurka.

**Eesmärk:** Liikuda võimalikult kiiresti edasi.

# Mida on Stiimulõppega saavutatud



**AlphaGo:** Google DeepMind'i poolt arendatud agent, mis 2016. aastal võitis maailmameistrid Lee Sedol'i 4:1.

**AlphaGo Zero:** 2017. aasta agent, mis õppis, mängides iseenda vastu ja võitis algset agentti 100:0

Koolitus „Andmeteadus on Popp“

# Autonoomsed otsustusesüsteemid

November 2021

Kristjan Eljand



Euroopa Liit  
Euroopa Sotsiaalfond



Eesti  
tuleviku heaks