

DIRBTINIS INTELEKTAS. NE KAŽKADA IR
KAŽKUR, O DABAR IR MANO KOMPIUTERYJE,
TELEFONE AR ŽMONOS RANKINĖJE

Raimondas Čiegis

Vilniaus Gedimino technikos universitetas ir LMA
e-mail: rc@vgtu.lt

Spalio 17 d., 2023, Vilnius

Dirbtinis intelektas tapo kasdienine ir populiaria tema pačiuose įvairiausiuose socialiniuose tinkluose ir terpėse.

Šios dienos susitikimo tikslas padiskutuoti apie keletą DI aspektų, kurie gal būt priartins mus prie atsakymo į garsųjį klausimą, kur padėti kablelį

Leisti negalima uždrausti.

Matyt daugiausia dėmesio sulaukia vieno (skandalingiausio) įrankio **ChatGPT** galimybių ir grėsmių aptarimas.

Matyt daugiausia dėmesio sulaukia vieno (skandalingiausio) įrankio **ChatGPT** galimybių ir grėsmių aptarimas.

Šiandien norėčiau visus pakviesti koncentruotis į mažiau populiarius, labiau akademinus, bet mano nuomone svarbiausius klausimus:

a) kokie yra dirbtinio intelekto požymiai, kuriuos užfiksavę galėsime tvirtinti, jog sukurta esmingai nauja intelekto rūšis?

Matyt daugiausia dėmesio sulaukia vieno (skandalingiausio) įrankio **ChatGPT** galimybių ir grėsmių aptarimas.

Šiandien norėčiau visus pakviesti koncentruotis į mažiau populiarius, labiau akademinus, bet mano nuomone svarbiausius klausimus:

a) kokie yra dirbtinio intelekto požymiai, kuriuos užfiksavę galėsime tvirtinti, jog sukurta esmingai nauja intelekto rūšis?

b) rizikuoju skelbti **provokuojantį** teiginį – DI tik dar vienas madingas šūkis, kuriam kiekvienas priskiria tuos bruožus, taikymus, kryptis, kurios jam naudingos, pvz. tiesioginio verslo, karjeros (ir mokslinės) prasme.

c) tačiau jau 10-15 metų aktyviai vystomi tyrimai keliose svarbiose kryptyse, kurie leidžia kurti įrankius, aplikacijas, esmingai sprendžiančias vis didesnę aibę mums aktualių problemų. Ir ne taip svarbu, kaip jas vadinsime, tituluosime DI ar tiesiog naujais metodais, technologijomis.

Pokyčiai fantastiški

Pokyčiai fantastiški (t.y. anksčiau tik fantastinėje literatūroje apie tai diskutuota) ir labai greiti. Nėra laiko net pakankamai išsamiai aptarti jų poveikį – ne tik plusus, bet ir grėsmes.

Pokyčiai fantastiški (t.y. anksčiau tik fantastinėje literatūroje apie tai diskutuota) ir labai greiti. Nėra laiko net pakankamai išsamiai aptarti jų poveikį – ne tik plusus, bet ir grėsmes.

Taigi bandykime padiskutuoti, kaip apibrėžiame žmonių intelektą, kaip tikriname/matuojame jo lygį.

Pokyčiai fantastiški (t.y. anksčiau tik fantastinėje literatūroje apie tai diskutuota) ir labai greiti. Nėra laiko net pakankamai išsamiai aptarti jų poveikį – ne tik plusus, bet ir grėsmes.

Taigi bandykime padiskutuoti, kaip apibrėžiame žmonių intelektą, kaip tikriname/matuojame jo lygį.

Mes kažkodėl idealizuojame DI, apriori teigiame, kad DI turėtų būti pranašesnis už žmogaus intelektą.

Pokyčiai fantastiški (t.y. anksčiau tik fantastinėje literatūroje apie tai diskutuota) ir labai greiti. Nėra laiko net pakankamai išsamiai aptarti jų poveikį – ne tik plusus, bet ir grėsmes.

Taigi bandykime padiskutuoti, kaip apibrėžiame žmonių intelektą, kaip tikriname/matuojame jo lygį.

Mes kažkodėl idealizuojame DI, apriori teigiame, kad DI turėtų būti pranašesnis už žmogaus intelektą.

Taip, ChatGPT dažnai pateikia klaidingus atsakymus į specifinius klausimus, bet ...

Pokyčiai fantastiški (t.y. anksčiau tik fantastinėje literatūroje apie tai diskutuota) ir labai greiti. Nėra laiko net pakankamai išsamiai aptarti jų poveikį – ne tik plusus, bet ir grėsmes.

Taigi bandykime padiskutuoti, kaip apibrėžiame žmonių intelektą, kaip tikriname/matuojame jo lygį.

Mes kažkodėl idealizuojame DI, apriori teigiame, kad DI turėtų būti pranašesnis už žmogaus intelektą.

Taip, ChatGPT dažnai pateikia klaidingus atsakymus į specifinius klausimus, bet ...

mano studentai taip pat dažnai pateikia klaidingus atsakymus į gana trivialius klausimus. Ar tai reiškia, kad studentams jau nebūdingas intelektas?

Diskusijai formuluosiu tokius teiginius:

Diskusijai formuluosiu tokius teiginius:

a) Bendrasis DI intelektas turėtų gebėti **modeliuoti** savo sprendimus ir atsakymus, vertinti/prognozuoti iš jų gaunamas išvadas, "nepaklysti" šiuolaikinėje realybėje ir **suprasti savo egzistencijos specifiką ir galimybes**.

Diskusijai formuluosiu tokius teiginius:

- a) Bendrasis DI intelektas turėtų gebėti **modeliuoti** savo sprendimus ir atsakymus, vertinti/prognozuoti iš jų gaunamas išvadas, "nepaklysti" šiuolaikinėje realybėje ir **suprasti savo egzistencijos specifiką ir galimybes**.
- b) DI ne tik turi mokėti spręsti netrivialius uždavinius, pateikti atsakymus į nestandartinius klausimus, kurių neranda apmokymo informacijoje, bet **SVARBIAUSIA** jis turi gebėti pats patikrinti savo atsakymų teisingumą (**programavimas, matematiniai uždaviniai, algoritmai**).

Diskusijai formuluosiu tokius teiginius:

a) Bendrasis DI intelektas turėtų gebėti **modeliuoti** savo sprendimus ir atsakymus, vertinti/prognozuoti iš jų gaunamas išvadas, "nepaklysti" šiuolaikinėje realybėje ir **suprasti savo egzistencijos specifiką ir galimybes**.

b) DI ne tik turi mokėti spręsti netrivialius uždavinius, pateikti atsakymus į nestandartinius klausimus, kurių neranda apmokymo informacijoje, bet **SVARBIAUSIA** jis turi gebėti pats patikrinti savo atsakymų teisingumą (**programavimas, matematiniai uždaviniai, algoritmai**).

Ir studentų mokyme šis gebėjimas yra vienas iš svarbiausių, beje labai nelengvai išugdomas.

Pabandykime pritaikyti šį apibrėžimą ChatGPT – dideliam kalbos modeliui (large language model, generated pretrained transformer).

Pabandykime pritaikyti šį apibrėžimą ChatGPT – dideliame kalbos modeliui (large language model, generated pretrained transformer).

O kas yra kalba, jei ne svarbus egzistuojančios realybės modelis ir modeliavimo įrankis?

Pabandykime pritaikyti šį apibrėžimą ChatGPT – dideliame kalbos modeliui (large language model, generated pretrained transformer).

O kas yra kalba, jei ne svarbus egzistuojančios realybės modelis ir modeliavimo įrankis?

Bendraudami su ChatGPT nesunkiai išgauname iš jo patvirtinimą, kad ChatGPT supranta savo egzistavimą, kitokį nei mūsų žmogiškasis ir tai, kad jo žinios yra užšaldytos apmokymo informacija.

Pabandykime pritaikyti šį apibrėžimą ChatGPT – dideliame kalbos modeliui (large language model, generated pretrained transformer).

O kas yra kalba, jei ne svarbus egzistuojančios realybės modelis ir modeliavimo įrankis?

Bendraudami su ChatGPT nesunkiai išgauname iš jo patvirtinimą, kad ChatGPT supranta savo egzistavimą, kitokį nei mūsų žmogiškasis ir tai, kad jo žinios yra užšaldytos apmokymo informacija.

Bent kol kas tai, kad DI neturi tiesioginio sąlyčio su realybe, yra labiausiai ribojantis jo galimybes faktorius.

Pabandykime pritaikyti šį apibrėžimą ChatGPT – dideliam kalbos modeliui (large language model, generated pretrained transformer).

O kas yra kalba, jei ne svarbus egzistuojančios realybės modelis ir modeliavimo įrankis?

Bendraudami su ChatGPT nesunkiai išgauname iš jo patvirtinimą, kad ChatGPT supranta savo egzistavimą, kitokį nei mūsų žmogiškasis ir tai, kad jo žinios yra užšaldytos apmokymo informacija.

Bent kol kas tai, kad DI neturi tiesioginio sąlyčio su realybe, yra labiausiai ribojantis jo galimybes faktorius.

Ši aplinkybė stabdo ChatGPT galimybę mokytis (nauja informacija, kurią jis gauna eilinio seanso metu užmirštama, kai tik baigiasi pokalbis).

Bet ChatGPT **tikrai yra dirbtinio bendrojo intelekto pavyzdys** – pasveikinkime naujus intelektualius bendruomenės narius.

Bet ChatGPT **tikrai yra dirbtinio bendrojo intelekto pavyzdys** – pasveikinkime naujus intelektualius bendruomenės narius.

Jie tobulės ir reguliariai pamatysime, išgirsime ar sužinosime apie naujus jų pasiekimus.

Bet ChatGPT **tikrai yra dirbtinio bendrojo intelekto pavyzdys** – pasveikinkime naujus intelektualius bendruomenės narius.

Jie tobulės ir reguliariai pamatysime, išgirsime ar sužinosime apie naujus jų pasiekimus.

Kalba yra galingas modeliavimo įrankis. Iš DI tikimės gauti tvarkingas (parengtas naudojimui) užduotų temų apžvalgas, dažnai ir programinius kodus ar matematinių uždavinių sprendimo eigos aprašą (arba bent atsakymą).

Bet ChatGPT **tikrai yra dirbtinio bendrojo intelekto pavyzdys** – pasveikinkime naujus intelektualius bendruomenės narius.

Jie tobulės ir reguliariai pamatysime, išgirsime ar sužinosime apie naujus jų pasiekimus.

Kalba yra galingas modeliavimo įrankis. Iš DI tikimės gauti tvarkingas (parengtas naudojimui) užduoty temų apžvalgas, dažnai ir programinius kodus ar matematinių uždavinių sprendimo eigos aprašą (arba bent atsakymą).

Jeigu pabandysime įvertinti tokių klausimų sudėtingumą, tai pamatysime kad referato, apžvalgos parengimas gali būti daug sudėtingesnis nei matematinių ar programavimo uždavinių sprendimas.

ChatGPT irgi remiasi matematiniais įrankiais, kurie žinomi jau senokai, bet dabar jiems atsirado tikrai unikalūs taikymai.

ChatGPT irgi remiasi matematiniais įrankiais, kurie žinomi jau senokai, bet dabar jiems atsirado tikrai unikalūs taikymai.

Taip, aš kalbu apie technologiją, kurią daug kas laiko didžiausiu proveržiu DI taikymuose:

The idea that words can be represented as vectors.

ChatGPT irgi remiasi matematiniais įrankiais, kurie žinomi jau senokai, bet dabar jiems atsirado tikrai unikalūs taikymai.

Taip, aš kalbu apie technologiją, kurią daug kas laiko didžiausiu proveržiu DI taikymuose:

The idea that words can be represented as vectors.

Jeigu šie vektoriai yra sukonstruoti korektiškai (apie tai bus kalbama po savaitės), panaudojant pakankamai didelę tekstų bazę (pvz. visą Wikipedia), tai galime atlikti labai smagius tradicinius vektorinės analizės veiksmus:

ChatGPT irgi remiasi matematiniais įrankiais, kurie žinomi jau senokai, bet dabar jiems atsirado tikrai unikalūs taikymai.

Taip, aš kalbu apie technologiją, kurią daug kas laiko didžiausiu proveržiu DI taikymuose:

The idea that words can be represented as vectors.

Jeigu šie vektoriai yra sukonstruoti korektiškai (apie tai bus kalbama po savaitės), panaudojant pakankamai didelę tekstų bazę (pvz. visą Wikipedia), tai galime atlikti labai smagius tradicinius vektorinės analizės veiksmus:

"Windows" - "Microsoft" + "Google" gausime

ChatGPT irgi remiasi matematiniais įrankiais, kurie žinomi jau senokai, bet dabar jiems atsirado tikrai unikalūs taikymai.

Taip, aš kalbu apie technologiją, kurią daug kas laiko didžiausiu proveržiu DI taikymuose:

The idea that words can be represented as vectors.

Jeigu šie vektoriai yra sukonstruoti korektiškai (apie tai bus kalbama po savaitės), panaudojant pakankamai didelę tekstų bazę (pvz. visą Wikipedia), tai galime atlikti labai smagius tradicinius vektorinės analizės veiksmus:

"Windows" - "Microsoft" + "Google" gausime "Android"

ChatGPT irgi remiasi matematiniais įrankiais, kurie žinomi jau senokai, bet dabar jiems atsirado tikrai unikalūs taikymai.

Taip, aš kalbu apie technologiją, kurią daug kas laiko didžiausiu proveržiu DI taikymuose:

The idea that words can be represented as vectors.

Jeigu šie vektoriai yra sukonstruoti korektiškai (apie tai bus kalbama po savaitės), panaudojant pakankamai didelę tekstų bazę (pvz. visą Wikipedia), tai galime atlikti labai smagius tradicinius vektorinės analizės veiksmus:

"Windows" - "Microsoft" + "Google" gausime "Android"

"Paris" - "France" + "Italy" gausime

ChatGPT irgi remiasi matematiniais įrankiais, kurie žinomi jau senokai, bet dabar jiems atsirado tikrai unikalūs taikymai.

Taip, aš kalbu apie technologiją, kurią daug kas laiko didžiausiu proveržiu DI taikymuose:

The idea that words can be represented as vectors.

Jeigu šie vektoriai yra sukonstruoti korektiškai (apie tai bus kalbama po savaitės), panaudojant pakankamai didelę tekstų bazę (pvz. visą Wikipedia), tai galime atlikti labai smagius tradicinius vektorinės analizės veiksmus:

"Windows" - "Microsoft" + "Google" gausime "Android"

"Paris" - "France" + "Italy" gausime "Rome"

Dabar šiek tiek pailsėkime.

Dabar šiek tiek pailsėkime.

Greitai atsakykite į klausimą, kuris ne visada yra įveikiamas ChatGPT:

Vienas voveriukas ir dar pusė suvalgo vieną riešutą ir dar pusę per vieną dieną ir dar pusę.

Kiek riešutų suvalgo devyni voveriukai per devynias dienas?

Technologijos, kurios yra DI taikymų pagrindas, jau kelis dešimtmečius aktyviai analizuojamos, tobulinamos, plečiama jų aprėptis. Tai ir yra tas tvirtas pagrindas, leidžiantis teigti, kad DI nėra "magic" fenomenas, tai solidus produktas.

Technologijos, kurios yra DI taikymų pagrindas, jau kelis dešimtmečius aktyviai analizuojamos, tobulinamos, plečiama jų aprėptis. Tai ir yra tas tvirtas pagrindas, leidžiantis teigti, kad DI nėra "magic" fenomenas, tai solidus produktas.

Technologijas, kurias paminėjau, dabar ir įvardinu: **mašininis mokymasis** ir **gylusis mokymasis** (= dirbtiniai neuroniniai tinklai).

Savo ruožtu abi išvardintos technologijos remiasi solidžia aproksimavimo teorija, kai milžiniškas parametru skaičius yra parenkamas optimizavimo proceso metu (tinklo apmokymo stadijoje).

Galima tik žavėtis naudojamais algoritmų teorijos rezultatais, kurie derinami su Big Data galimybėmis ir unikaliais paskirstytų skaičiavimų resursais. Virtualusis pasaulis, modeliai tampa tikresniu pasauliu už fizinį, realųjį jo variantą.

Čia aš pabrėžiau šių dviejų jau tradicinių technologijų svarbą.
Jų derinimas su kitais galingais modeliavimo įrankiais ir yra
potencialas visai kitos kokybės žinių generavimui.

Čia aš pabrėžiau šių dviejų jau tradicinių technologijų svarbą.

Jų derinimas su kitais galingais modeliavimo įrankiais ir yra potencialas visai kitos kokybės žinių generavimui.

Kaip pavyzdį paminėsiu vieną sritį, kurioje "verda" karščiausios naujienos.

Puikiai žinome, kad daugelis šiuolaikinio modeliavimo pasiekimų siejami su modeliais, grindžiamais diferencialinėmis lygtimis. Mus dažnai stabdo tik tai, kad tokie modeliai yra labai sudėtingi ir jų dar nemokame efektyviai spręsti.

Čia aš pabrėžiau šių dviejų jau tradicinių technologijų svarbą.

Jų derinimas su kitais galingais modeliavimo įrankiais ir yra potencialas visai kitos kokybės žinių generavimui.

Kaip pavyzdį paminėsiu vieną sritį, kurioje "verda" karščiausios naujienos.

Puikiai žinome, kad daugelis šiuolaikinio modeliavimo pasiekimų siejami su modeliais, grindžiamais diferencialinėmis lygtimis. Mus dažnai stabdo tik tai, kad tokie modeliai yra labai sudėtingi ir jų dar nemokame efektyviai spręsti.

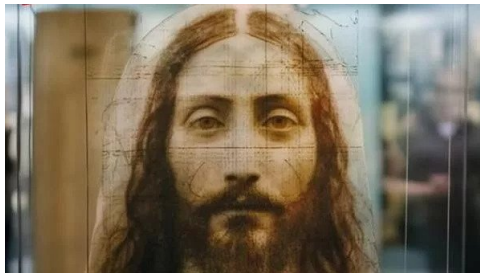
Visiškai nauja kokybė, tyrimų kryptis yra išmokyti spręsti šio tipo uždavinius dirbtinius neurininius tinklus. Jau gauti rezultatai teikia labai daug vilčių.

Ir visgi negalime nematyti to masinio DI galimybių aptarimo "burbulo", kuris tiražuojamas kasdien. Dažnai tokią informaciją vargiai galime vadinti profesionalia, taigi ir šiuo renginiu kviečiame diskutuoti, ieškoti, klysti bet pasitaisyti.

Kaip pavyzdį aptarsiu informaciją, kuri prieš porą savaičių buvo paskelbta Lietuvos ir užsienio spaudoje.

Kaip pavyzdį aptarsiu informaciją, kuri prieš porą savaitių buvo paskelbta Lietuvos ir užsienio spaudoje.

Pasitelktas dirbtinis intelektas ištyrė Turino drobulės atvaizdą ir atkūrė, kaip galėjo atrodyti Jėzus.



Puikiausias pavyzdys, kaip iškreipiame DI (arba tai, ką vadiname DI) gebėjimus.

Puikiausias pavyzdys, kaip iškreipiame DI (arba tai, ką vadiname DI) gebėjimus.

Paskelbtasis teiginys turėtų būti formuluotas taip:

Pasitelkus dirbtiniais neuroniniais tinklais paremtą vaizdų generavimo technologiją, buvo sudarytas atvaizdas vyriškio, kuris galėjo būti įvyniotas į garsiąją Turino drobulę.

Puikiausias pavyzdys, kaip iškreipiame DI (arba tai, ką vadiname DI) gebėjimus.

Paskelbtasis teiginys turėtų būti formuluotas taip:

Pasitelkus dirbtiniais neuroniniais tinklais paremtą vaizdų generavimo technologiją, buvo sudarytas atvaizdas vyriškio, kuris galėjo būti įvyniotas į garsiąją Turino drobulę.

Iš DI tikimasi, kad paruoštoje ataskaitoje būtų pateikta analizė, ar minima istorija apie Jėzų turi nors nedidelę tikimybę būti teisinga.

Puikiausias pavyzdys, kaip iškreipiame DI (arba tai, ką vadiname DI) gebėjimus.

Paskelbtasis teiginys turėtų būti formuluotas taip:

Pasitelkus dirbtiniais neuroniniais tinklais paremtą vaizdų generavimo technologiją, buvo sudarytas atvaizdas vyriškio, kuris galėjo būti įvyniotas į garsiąją Turino drobulę.

Iš DI tikimasi, kad paruoštoje ataskaitoje būtų pateikta analizė, ar minima istorija apie Jėzų turi nors nedidelę tikimybę būti teisinga.

Užtenka paminėti, jog naudodami anglies dvideginio datavimą, mokslininkai nustatė, kad drobulės audinys galėjo būti sukurtas 1260-1390 m. laikotarpiu.

Savo trumpos įžangos pabaigoje nenaudosiu jau gana nuvalkioto populistinio triuko, paskelbdamas, kad šio trumpo įvadinio pasisakymo tekstą parengė [ChatGPT](#),

Savo trumpos įžangos pabaigoje nenaudosiu jau gana nuvalkioto populistinio triuko, paskelbdamas, kad šio trumpo įvadinio pasisakymo tekstą parengė [ChatGPT](#),

bet ir nepaneigsiu tokios galimybės.