

Čas in umetna inteligenca - mednarodna poletna šola v Bolzanu

DANILO ŠUSTER

Od leta 1988 v slikovitem Bolzanu vsako leto poteka mednarodna poletna šola o filozofiji in umetni inteligenci. Organizacijo je zadnja leta prevzel Srednjeevropski inštitut za kulturo iz Bolzana. Šola, ki poteka v začetku julija, je dvodelna - prvi teden je namenjen filozofskim temam, ki so tipično "srednje-evropske", saj združujejo pristope analitične filozofije in fenomenologije (brentanovska tradicija, graška šola, poljska filozofija). Drugi del pa povezuje povsem "tehnična" vprašanja umetne inteligence (UI) s filozofskimi problemi.

Letošnja (1992) šola je bila posvečena Filozofiji jezika, drugi del (13.7. - 17.7.) pa je nosil naslov "Razmišljanje v času" ("Temporal reasoning"). Predavatelji so bili James Allen (Rochester), Johan Van Bentham (Amsterdam), Drew McDermott (Yale) ter Erik Sandewall (Link-ping). Udeleženci iz Italije, Francije, Portugalske, Nemčije, Nizozemske, Škotske, Madžarske, Poljske in Slovenije pa so v glavnem prišli z različnih inštitutov za računalništvo, informatiko in umetno inteligenco, nekaj pa je bilo tudi logikov in filozofov.

Kaj je "Razmišljanje v času"? Tema je zadnje čase v UI-skupnosti deležna izjemne pozornosti. Raziskovalci UI želijo konstruirati sisteme (robote), ki bodo sposobni reševati določene "živiljenjske" probleme. Zato pa morajo biti sposobni razsojati o dogodkih, trajanju, dejanjih v času. Takšno razsojanje, ki ga skušajo raziskovalci prevesti v učinkovite algoritme, pa temelji na neki časovni logiki in logiki dejanj, ki ju filozofi in logiki že dolgo preučujejo.

Običajno so to ne-monotona sklepanja. V grobem gre za proces izpeljave zaključkov, ki jih lahko ovrže pridobitev nove informacije. Ta tip razmišljanja je blizu običajnemu človeškemu rezoniranju, ki temelji na racionalnosti in ne na resnici. Naši zaključki niso vedno resnični, pa vendar niso povsem naključni - opiramo se na naše vedenje o svetu ter na naša izkustva, preden se odločimo, ali je nek zaključek sprejemljiv ali ne. Prav ta racionalna upravičenost zaključkov je tisto, za kar bi raziskovalci umetne inteligence radi našli ustrezen logični model, ga prevedli v algoritem in z njim opremili robote.

Največkrat si izberejo neko idealizirano okolje, nek mini-svet, v katerega postavijo robota, ki mora s svojimi dejanji tako predelati okolje, da bo dosegel postavljeni cilj. Gre recimo za nek svet naključno razporejenih kock, ki jih mora robot razporediti v neko želeno zaporedje. S tem, ko izvede neko dejanje (recimo prebarva rdečo kocko v zeleno), robot spremeni stanje svojega "mini-sveta", s čimer postane uskladiščeno

vedenje, stavki, ki so prej pravilno opisovali svet, napačni. Revizija množice prepričanj o svetu glede na učinke je že eden glavnih (in slavnih) problemov "razmišljanja v času", problem okvirja ("frame problem").

Na kratko - gre za problem razdelitve aspektov sveta na tiste, ki ostanejo nespremenjeni, in tiste, ki se spremenijo, ko robot izvede neko dejanje. Delovanje običajno spremeni samo nekatere lastnosti, recimo, če robot prebarva neko kocko, se spremeni samo obarvanost, vse druge lastnosti ostanejo nespremenjene. Toda robota je treba nekako vnaprej opremiti z znanjem o teh invariantah, z njihovo eksplicitno reprezentacijo. To dosežemo z ne-monotonimi pravili tipa: vsi aspekti sveta ostanejo nespremenjeni, razen tistih, ki so se eksplicitno spremenili. Obstaja več načinov za logični izraz te ideje. Običajno klasični logiki dodamo pravila sklepanja oblike: "iz premise p ob neveljavnosti q , izpelji r ". Zaključek r je običajni sklep, relativiziran glede na informacijo q . Recimo: če je a ptič (p) in če je konsistentno s teorijo, da a ni ne-normalen ptič (q), tedaj a lahko leti (r). Glavno težavo za vse te logične sisteme predstavlja dejstvo, da so neprimerni za sklepanja v domeni, kjer se stvari in dogodki spreminjajo v času. S tem pa smo že pri glavni temi predavanj Allena, Sandwella in McDermotta - kako modelirati proces običajnega ("common-sense") sklepanja in razmišljanja v času.

McDermott je predstavil "Planning problem" - temporalno sklepanje v okviru problema načrtovanja. Robot mora razmisliti o posledicah različnih dejanj, da bi dosegel želeni cilj. Ob danem opisu začetne situacije S in opisu želene situacije I je treba robota opremiti z načrtom, s katerim robot doseže situacijo, ki ustreza opisu I , kadar načrt izvedemo v situaciji, ki ustreza opisu S . McDermott je potem podrobno predstavil izdelavo "načrtovalca" - logični jezik za opis dejanj, predstavitev načrtov kot grafov in izdelavo algoritma (zaporedja korakov in delnih načrtov), s katerim robot dokazljivo doseže postavljeni cilj. To je zelo grob povzetek, saj je McDermott zelo natančno obdelal številne fineise in predstavil, kolikor lahko sodi nekdo, ki se ne ukvarja s problemi umetne inteligence, tudi nekaj novosti (poudarek na algoritmih reaktivnega načrtovanja - od cilja nazaj do začetne situacije, spreminjanje načrtov v času, probabilistično sklepanje v času, ločitev interpreter/načrtovalec ipd.)

James Allen je predstavil model načrtovanja za situacije, ko imamo več med seboj neodvisnih istočasnih dejanj, s katerimi želimo doseči želeni cilj (primer je zabijanje žebelja v zid, kjer gre za dve dejanji - istočasno držimo žebelj in udarjamo po njem s kladivom). Interakcija dejanj proizvede različne učinke, s katerimi se mora robot spopasti. Če temu dodamo še medsebojno interakcijo robotov, dobimo že kar zapleten in "življenjski" mini-svet. Allenov osnovni model je predstavljala železniška postaja z vlaki, ki prihajajo in odhajajo, imajo zamude, se srečujejo z ovirami, vse to pa upravljajo roboti, ki natovarjajo in raztovarjajo vagoni, med seboj komunicirajo, se usklajujejo in izbirajo prave načrte delovanja.

Če pustim ob strani tehnične podrobnosti (njegov časovni sistem temelji na logiki intervalov in ne časovnih točk), potem je za filozofa najbolj zanimiv osnovni okvir Allenovega modela. Predpostavka je, da roboti komunicirajo v naravnem jeziku, tako da mora robot na osnovi informacij, ki mu jih v naravnem jeziku sporoča drug robot, izbrati pravi načrt. Koraki so naslednji: opazuj drugega dejavnika, skušaj prepoznati pravi načrt (kaj mi drugi želi sporočiti), izberi cilj, izdelaj načrt in izvedi dejanje. Allenova teoretska predpostavka je Griceova teorija pomena in teorija jezikovnih dejanj. Svoje robote načrtuje tako, da ugibajo o "intencah" drugih robotov, ki svoja

sporočila prenašajo z "intenco", da bi drugi robot prepoznal njihovo "intenco". Allen se zadnjih 15 let ukvarja s projektom, kako teorijo, po kateri komunikacija ni možna, če ni nekega nivoja sodelovanja med govornici (dejavniki), prevesti v nek komputacijski algoritem.

Predavanja Erica Sandwella so imela naslov: Dinamična sklepanja v umetni inteligenci - sklepanja o času in delovanju. Zaradi sistematičnega pristopa in izdelave celovite teorije so bila predavanja izredno zanimiva tudi za filozofa, saj gre za problem racionalnega delovanja v svetu, ki se spreminja, to pa je problem vsake "intelligence", ne le "umetne". Izhodišče predstavljata dve predpostavki: (i) v dinamičnem sistemu se s časom dogajajo spremembe (ii) inercija ali pravilo ne-monotonosti: stvari se ne spreminjajo, če ni nekega posebnega razloga za to (neko dejanje, vzročna veriga, nenaden nastop presentljivega dogodka...).

Sandwell je predlagal naslednji postopek. Izberi množstvo primerov običajnega (največkrat deduktivnega) sklepanja v neki vsakdanji domeni, identificiraj (predpostavljeno) splošno vedenje o tej domeni, razvij formalni jezik, v katerem je mogoče izraziti to vedenje, pazljivo definiraj primitivne izraze jezika in razvij logični sistem za sklepanja v tej domeni. Sandwellovo izhodišče je bilo dvanajst tipičnih problemov običajnega sklepanja, ki jih do sedaj noben računalniški model ni uspel zadovoljivo rešiti.

Najbolj zanimiva je bila teza, da je prehod od realnosti do logike, ki kodificira naša zdravorazumska sklepanja o tej realnosti, odvisen od pravilne idealizacije, abstrakcije te realnosti. Temu problemu je posvetil precej pozornosti in izdelal, za nekoga, ki se ukvarja s problemi algoritmov in njihove učinkovite implementacije, presentljivo podrobno in prepričljivo ontologijo (dogodki, dejanja, aktivnosti, poteze, aspekti, vzročna in ne-vzročna odvisnost, naključja, čudeži ...), kot ta vmesni korak med realnostjo in logiko. S tem so tudi njegove tehnične rešitve v logiki časovnega sklepanja zaživele v perspektivi, ki je najbolj na široko odpirala vrata za dialog s filozofskimi teorijami akcije.

Johan Van Bentham je pravzaprav edini predstavil tisto, kar filozofi in logiki običajno razumemo kot sklepanje v času - časovno logiko, ki jo je kot sistem za ocenjevanje argumentov, v katerih nastopajo izjave, katerih resničnostna vrednost se spreminja s časom, razvil A. N. Prior. Časovna logika razširja standardno logiko s časovnimi operatorji (G - V bodočnosti bo vedno res, da..., P - V preteklosti je bilo res, da...) in posebnimi aksiomi za te operatorje, recimo (4): Če bo vedno res, da p , tedaj bo vedno res, da je vedno res, da p , ali $Gp \supset GGp$.

Van Bentham je v prvem delu predstavil teorijo korespondence, ki je morda najbolj znano delo tega slavnega logika. Ideja sicer izhaja iz semantike za modalno logiko, Van Bentham pa je v vrsti razprav pokazal, kako jo je mogoče uporabiti za sistematični študij časovne logike, logike kondicionalov, epistemske logike, intencionalne logike, dinamične logike in celo relacijske algebre. Pravzaprav je trdil, da je mogoče s pomočjo korespondence med neko strukturo, ki jo opišemo v predikatni logiki prvega reda in intenzionalnimi aksiomi logike, ki jo preučujemo, pojasniti in semantično modelirati prav vsako ne-klasično logiko!

V semantiki za časovno logiko nastopa struktura T , kjer je T množica trenutkov, pa relacija časovne predhodnosti. Semantična interpretacija izjave Gp temelji na prevodu v standardno logiko, v kateri kvantificiramo čez trenutke časa. Izjava p je resnična v vseh trenutkih časa, ki sledijo izhodiščnemu trenutku. Teorija korespondence nam potem pove, da je teza $Gp \supset GGp$ resnična, če in samo če je relacija tranzitivna.

Sistematična korespondenca, ki jo preučuje Van Bentham, pa nastopa med izjavami, v katerih nastopajo intenzionalni operatorji (v tem primeru časovni), in izjavami o "strukturni", ki pa so v jeziku predikatne logike prvega reda - za katerekoli trenutke x, y, z , če je trenutek x pred trenutkom y in trenutek y pred trenutkom z , tedaj je trenutek x pred trenutkom z . Teorija korespondence se potem ukvarja z dvema problemoma - če je dana teza v jeziku intenzionalne logike, poišči relacijski pogoj prvega reda na ustreznih strukturi in če je dan relacijski pogoj, poišči ustrezní intenzionalni aksiom.

Van Bentham je predstavil različne metode, s katerimi odkrijemo, ali obstaja prevod za teze v časovni logiki ali ne. V drugem delu predavanj pa je razlagal osnovne ideje dinamične logike, ki izhaja iz tega, da se nekatere komponente semantične interpretacije v času spreminjajo. Dandanes skušajo logični sistemi modelirati dinamične aspekte spreminjanja in dopolnjevanja informacij, kognitivnega premikanja skozi vzorce informacijskih stanj. Van Bentham je zanimala predvsem aplikacija v računalniški znanosti, saj je povezava očitna - v dinamični logiki formule razumemo kot napotke za izvedbo semantične evaluacije. Kljub temu, da je tudi to logiko mogoče prevesti v strukture prvega reda, pa je vendarle precej "deviantna" - v njej ni pojma resnice in tudi logična sklepanja imajo drugačno interpretacijo. Recimo - iz A, B sklepaj na C pomeni: procedura C sledi iz procedur A in B; ali pa: procesiranje premis A in B upraviči skok na zaključek C. Dinamična logika je zelo zanimivo področje, ki pa je sorazmerno neznano in precej zahtevno, ne nazadnje zato, ker jo tisti, ki se z njo ukvarjajo, po nepotrebnem, kot pravi Van Bentham, obdajajo z avreolo misterioznosti.

Morda je zanimivo opozoriti na razliko med logično in filozofsko perspektivo glede razmišljanja v času in pristopom raziskovalcev s področja umetne inteligence. Van Bentham je predstavil problem tako, kot to stori logik - zanimajo ga strukture same po sebi in možne posplošitve. Za raziskovalce UI je logika samo prvi korak do algoritmov, arhitekture in konstrukcije sistema (robota). Tak pristop ima dobre in slabe strani. Omejitev na mini-svetove kock, opredeljene z začetnimi aksiomi, in osredotočenje na enostavna dejanja (poberi predmet / spusti predmet) omogočajo preverljivost modela. Glede na to, ali in kako sistem deluje, je mogoče izločiti neuspešne ali slabe pristope. Omogoča tudi, da se raziskovalec osredotoči le na bistvene poteze problema - večina predavateljev je recimo menila, da je propozicionalna logika povsem zadostno orodje za robota, ki se premika v svetu in zadeva na fiksno, končno domeno.

Tisto, kar so za filozofa glavna vprašanja - kaj je vzrok, dejanje, dogodek, intencionalo dejanje, dejavnost... je že na začetku definirano z aksiomi izbranega mini sveta. Stara filozofska uganka, kako identificirati isto dejanje pod različnimi opisi, v tem kontekstu sploh ni problem. Predpostavka je, da če robot lahko izvrši dejanje pod enim opisom, ga lahko izvrši pod katerimkoli opisom. Zdi se mi, da si raziskovalci UI včasih prav s takšnimi poenostavitvami in slepoto za načelno, teoretsko (in s tem neposredno ne-aplikabilno) dimenzijo zapirajo pot do dobrih rešitev.

Skupen imenovalce celotnega ciklusa predavanj je bilo logično modeliranje običajnega zdravorazumskega sklepanja v svetu, ki se spreminja v času, iskanje neke "človeške" logike, zaradi katere so ljudje neprimerljivo uspešnejši v reševanju problemov od najboljših do sedaj narejenih umetnih sistemov. Van Bentham je lepo opozoril na paradoks kompleksnosti - vsi sistemi zdravorazumskega sklepanja so veliko bolj zapleteni, kot pa klasična logika. Bolj ko se skušaj približati na videz enostavnemu, vsakdanjemu človeškemu razmišljanju, bolj zapleteni postajajo logični sistemi. Na koncu smo soočeni z vprašanjem, ali so ljudje, ki očitno brez težav izpeljujejo pravilne

(racionalne, upravičene) zaključke v svetu dinamičnih procesov, zares tako dobri "sklepalci". Po predstavljenih teorijah v vsakdanjem življenju igraje obvladamo sklepanja, katerih teorija (logika) je izjemno zapletena, težko jo je sistematično prikazati in izpeljevati sklepanja po njenih napotkih.

Gre za zanimivo vprašanje, na katerega je možnih več odgovorov. Eden je ta, da ljudje morda niso popolni glede na svojo lastno gramatiko ("vrojeni logični sistem") - tako proizvajajo resnične zaključke, ki jih ni mogoče dokazati, izpeljati po pravilih logičnega sistema, s katerim so opremljeni. Do pravih zaključkov torej ne prihajamo po zapleteni logični poti, ki jo trasirajo sistemi dinamične časovne logike, ampak na nek drug način. Seveda se potem postavi vprašanje - kako?

Bližje mi je odgovor, ki pravi, da je treba bolj natančno opredeliti problem vsakdanjega zdravorazumskega sklepanja. Gre pravzaprav za paket, v katerem moramo ločiti dve komponenti - reprezentacijo informacije (znanja) in samo sklepanje. Sklepanje morda temelji na enostavni in najbolj razumljeni klasični logiki. Večina problemov, ki jih rešujejo raziskovalci umetne inteligence, morda temelji na tem, da ni dobrega teoretskega modela o tem, kako je vedenje o svetu reprezentirano. Kar pa je seveda klasični problem okvirja.

Prepad med sredstvi, ki so teoretiku na voljo (razni sistemi monotonih in nemonotonih logik) in želenim ciljem (uspešno modeliranje običajnega razmišljanja v času) še vedno ostaja, čeprav so predavatelji in drugi udeleženci simpozija (ti so predstavili svoje delo v dveh "delavnicah") pokazali, da gre za področje, kjer je teoretska iniciativa v velikem razmahu. Morda bi položaj lahko opisal s pomočjo šale iz Adamsovega Vodiča: kako doseči, da bi človek poletel? Odgovor pravi, da je to zelo enostavno, vržeš se na tla in zgrešiš. Podobno vprašanje je, kako z logičnimi sredstvi zajeti kompleksnost običajnega razmišljanja v času. (Ali ekvivalentno, če imamo pred očmi, da je cilj vseh raziskovalcev na tem področju uspešna implementacija te logike v algoritmih - kako doseči, da bi se robot uspešno spopadel z reševanjem problemov v času?) Odgovori, ki so jih ponujali predavatelji so, če nekoliko parafraziram: tako, da vzameš klasično logiko in jo skušaš modificirati, pravilno "zgrešiti" z raznimi nemonotonomimi pravili. Seveda pa je problem v tem, ali je s temi sredstvi sploh mogoče zgrešiti.