

Ocjenjivanje

Lavoslav Čaklović

Revidirano 25. veljače 2025.

<https://web.math.hr/~caklovic/eseji/ocjenivanje.pdf>

Sadržaj

1	Uvod	2	7.9	Preferencijalna (ne)zavisnost atributa	30
2	Vrednovanje	2	7.9.1	Preferencijalna nezavisnost atributa	31
3	Što je produkt obrazovanja?	4	7.9.2	Ilustracija MAVT.	33
4	Ciljevi vrednovanja u obrazovanju	5	8	Subjektivno mjerenje	35
4.1	Standardizacija vrednovanja znanja	5	8.1	Ordinalna mjera	35
4.2	Ocjena, što je to?	6	8.2	Graf preferencije	36
4.3	Svrha ocjenjivanja	7	8.3	Mjerenje razlike preferencije	37
4.3.1	Održavanje standarda ili razvijanje sposobnosti?	8	9	Snadi se kako znaš	38
5	Hijerarhija odluke i potciljevi	8	9.1	Rangiranje profila	38
5.1	Mišljenje orijentirano na proces	9	9.2	Dualnost u ocjenjivanju	39
5.1.1	Proces kao hijerarhija.	11	9.3	Agregacija ispitnih bodova	41
5.2	U potrazi za ciljevima	13	9.4	Subjektivna i egzaktna procjena	43
6	Motivacija za učenjem	13	9.5	Zaključna ocjena kao pripadnost klasi	44
6.1	Pojam motivacije u literaturi	13	9.6	Važnost poznavanja skale	45
6.2	Postignuća kao ciljevi	14	9.6.1	Primjer ocjenske skale	46
6.3	Spoznaja kao cilj	15	9.7	Post-hock analiza rezultata pismene provjere znanja	47
6.4	Nastup kao cilj	15	10	Eksterna evaluacija	49
6.5	Slabosti AGT teorije	16	10.1	Procedura ocjenjivanja	50
6.6	Primarnost interesa	17	10.2	Principi, snaga i ograničenost ljudskog prosuđivanja	51
7	Agregacija ocjena	19	10.3	Određivanje granica među ocjenama	51
7.1	Subjektivna procjena	20	10.3.1	Neke nove/stare sugestije	52
7.2	O skalama	20	10.4	Državna matura u Hrvatskoj	52
7.3	Centralna mjera skale.	21	10.4.1	Iz pravilnika o polaganju državne mature, NN 87/08	52
7.4	Funkcija vrijednosti	23	10.4.2	Iz javnih glasila	53
7.5	Utjecaj mjerenja kod agregacije ocjena	24	10.4.3	Eksterna evaluacija nužnost ili zabluda	53
7.6	Kvalitativna i kvantitativna smislenost agregacije	25	11	Dodatak	53
7.7	Agregacija ordinalnih vrijednosti	25	11.1	Metoda potencijala	53
7.7.1	Bordino pravilo	26	11.1.1	Konzistentnost toka.	54
7.8	Većinska procjena. Medijan.	26	11.2	Agregacija grafova preferencije	54
7.8.1	Neke 'paradoksalne' situacije kod agregiranja medijanom	27	11.3	Potencijal tablice odlučivanja	55
7.8.2	Bordin većinski indeks	28	11.3.1	Dualnost u ocjenjivanju	56
7.8.3	Usporedba medijana i aritmetičke sredine	29	11.4	Ocjenjivanje učenika. Primjer	57
			11.4.1	Kompozitni tok tablice	57
			11.4.2	Samoprocjena	58
			11.5	Ordinalno rangiranje s metodom potencijala	58

11.6	Agregacija unutar hijerarhije	59	11.10.1	Dualnost ciljeva i obrazov- nih sadržaja	63
11.6.1	Zavisnost ocjena	59			
11.7	Samodualna hijerarhija	59	12	Umjesto zaključka	65
11.8	Revizija državnog proračuna	60	12.1	Agregacija	65
11.9	Inteligentni mobilni roboti	61	12.2	Metoda potencijala	65
11.10	Hijerarhija obrazovnih ciljeva	62	12.3	Kako vidim budućnost ocjenjivanja	66

Svi računi provedeni su uz pomoć softvera razvijenog od strane autora, a može se koristiti u akademske svrhe. Za upute i download posjetite web sučelje <https://web.math.pmf.unizg.hr/caklovic/software.html>

1 Uvod

Ocjena, u našem obrazovnom sustavu, predstavlja odraz učenikovog stupnja ovladanosti znanja iz nekog predmeta podučavanja i/ili vještina. Takva ocjena može se izraziti brojem na nekoj skali, obično je to skala 1–5 (1–6) ili vrijednošću na temperaturnoj ljestvici (1–100). Neki obrazovni sustavi, ruski na primjer, koriste uređene kategorijske vrijednosti A–F gdje je A najviša, a F najniža ocjena. Običaj je, barem u našem obrazovnom sustavu, ocjenjivanje na temperaturnoj ljestvici nazivati bodovanjem. Završna ocjena je ocjena nastala nekom procedurom agregiranja predmetnih ocjena, obično na polugodištu ili na kraju godine.

Iz gore rečenog proizlazi da termin 'ocjena' nije jednoznačan, tj. značenje tog termina je kontekstualno ovisno. Ako se smatra da je ocjena 'mjera' ovladavanja određenog gradiva, onda se tu implicitno podrazumijeva da ona mjeri učenika nekim 'metrom'. Taj 'metar' bi trebao biti *nepristran* u smislu da bi svako (ponovljeno) mjerenje učenikovog znanja trebalo dati istu ocjenu. Prevedeno na svakodnevni obrazovni riječnik:

Svaki nastavnik koji ispituje (istog) učenika na temelju istog testa (ili druge provjere znanja) trebao bi mu dati istu ocjenu.

Takav metar u obrazovnom sustavu ne postoji iz jednostavnog razloga što je ocjena (1) subjektivna i (2) ne postoji zacrtana procedura donošenja ocjene koja bi osiguravala nepristranost. Ako mi ne vjerujete čitajte nacionalni kurikulum.

U ovom tekstu djelomično ćemo se baviti konstrukcijom ocjenskog metra (skale) s posebnim naglaskom na *metodu potencijala* zasnovanu od autora, a opća teorija mjerenja nadilazi okvire ovog časopisa¹. Nećemo se baviti smisljenošću i opravdanošću ocjenjivanja u obrazovnom sustavu; uzimimo kao gotovu činjenicu da se ocjene u školama donose i da će tako zasigurno biti još neko vrijeme. Pozabavit ćemo se agregacijom već postojećih ocjena koja je dovoljno zanimljiva sama po sebi, tim više što nastavnici to rade svakodnevno, svaki na svoj način.

2 Vrednovanje

Vrednovanje kao pojam javlja se u *užem* i u *širem* smislu. U užem, tehničkom smislu, vrednovanje je *funkcija vrijednosti* na skupu S sa zadanom relacijom slabe preferencije \succsim koja odražava kvalitativni odnos među njegovim elementima. Ako je (S, \succsim) preferencijalna struktura onda *funkcija*

¹Znatiljnjog čitatelja upućujemo na autorovu knjigu *Teorija vrednovanja s naglaskom na metodu potencijala* i to na onaj dio koji se bavi konstrukcijom izmjerive funkcije vrijednosti.

vrijednosti $V : S \rightarrow \mathbb{R}$ zadovoljava

$$(\forall x, y \in S) x \succcurlyeq y \iff V(x) \geq V(y).$$

Drugim riječima, vrednovanje je kvantifikacija kvalitativnog odnosa među objektima. Vrijednost funkcije V može biti lingvistička ili numerička. Pozicioniranje na uređenoj lingvističkoj ljestvici uobičajeno je nazivati *rangiranjem*, a na numeričkoj ljestvici *bodovanjem*. Na primjeru temperature (kao kvalitete) ljestvica može biti: *vrelo, vruće, toplo, mlako, hladno, ledeno*. Takva ljestvica je uređena.

U kontekstu obrazovanja vrednovanju se daje širi smisao, a to je proces koji uključuje praćenje napredovanja učenika kao i razvoj instrumenata i protokola za prikupljanje informacija korisnih za roditelje, obrazovne ustanove i kreatora obrazovanja. Što se učenika tiče, vrednovanje se svodi na testiranje, ocjenjivanje i davanje sažetih procjena kao povratnih informacija o njegovom napretku.

Blisko značenje termina vrednovanje nose i termini: mjerenje, skaliranje, ocjenjivanje, rangiranje . . . i svaka od tih procedura u obrazovnoj praksi je stvar dogovora ili predstavlja običajno nasljeđe. Ono što dodatno komplicira vrednovanje u obrazovanju je kondenzacija ili agregacija dobivenih vrijednosti uvriježena kao zaključna ocjena, bilo da se radi u predmetnoj nastavi ili općem uspjehu.

Enciklopedija <https://plato.stanford.edu/entries/measurement-science/> Stanfordskog univerziteta ne nudi neku opću definiciju mjerenja ili što je to što bi trebalo omogućavati mjerenje. Postoji ustaljena praksa mjerenja unutar pojedinih područja, postoji filozofija mjerenja započeta Aristotelom i Euklidom koja raspravlja o veličinama i intenzitetu kvalitete i postoji matematičko zasnivanje mjerenja započeto Hölderovim teoremom (1911) i dovršeno osamdesetih godina prošlog stoljeća (Krantz et al., 1971; Roberts, 1979).

U nedostaku mjerne jedinice i strožih procedura mjerenja u obrazovanju se koristi testiranje. U najširem smislu test je bilo kakav sustavni uzorak ponašanja u kontroliranim uvjetima. To može biti rješavanje zadataka u matematici, izrada reljefa nekog geografskog područja, pisanje eseja, argumentirana rasprava ili čitanje teksta. Mišljenje o učenikovom napredovanju bazirano je na takvim uzorcima iz jednostavnog razloga jer nije moguće, a nije ni potrebno, da se procjenjuje sve ono što čini učenikov potencijal. Današnje obrazovanje priprema učenike za obavljanje društveno korisnih poslova i testiranje je tome prilagođeno.

Drugi, učestaliji oblik provjere znanja, u hrvatskom obrazovnom sustavu je usmena provjera znanja u kojoj nastavnik ili više njih, postavlja pitanja i vrednuje učenikove odgovore. Takva provjera znanja je stresna za učenika, ovisna o nastavniku, njegovom raspoloženju i namjeri, ona je dio kulturnog nasljeđa, posljedica autoritarnog režima i oblika nastave koja ne omogućava neke druge oblike vrednovanja kao što je projektna nastava na primjer. Takva provjera znanja završava ocjenom.

Ocjenjivanje je proces pridjeljivanja broja učenikovim "usvojenim znanjima i vještinama". Taj broj može biti postotak korektnih odgovora na seriju pitanja ili rezultat agregacije podataka dobivenih praćenjem tijekom određenog perioda. Ocjenjivanje učenika, brojčanom ocjenom od 1–5 na primjer, nakon što su mu postavljena pitanja na satu, nije vrednovanje u užem smislu jer nije donešeno na temelju neke preferencijalne strukture i uspoređivanju s drugim učenicima. Takvo ocjenjivanje na prvi pogled ima sličnosti s ekstenzivnim mjerenjem (Čaklović, 2014, str. 139), samo što u ovom slučaju ne postoji jedinična mjera nego se radi o zadovoljavanju zadanih normi. Nastavnik u svojoj glavi posjeduje imaginarne referentne učenike (standardi) za svaku ocjenu, a učenika kojeg ocjenjuje uspoređuje s tim standardima (v. odjeljak 9.5 pod nazivom [Zaključna ocjena kao pripadnost klasi](#).)

Procjena učenikovog napretka je sažetak dobiven iz više izvora, ukaz na učenikove "kvalitete" i "slabosti" kao i davanje uputa za daljnji napredak. Takva informacija otvorena je učeniku, razrednoj zajednici, školi i roditeljima. Može biti pismena i dodatno biti graduirana na nekoj skali. Ne bih ulazio u detalje što i kako se radi u praksi niti što je dobro, a što nije dobro jer je to nemoguće

odvojiti od načina izvođenja nastave i zacrtanog kurikula. Cilj ovog teksta je proces vrednovanja usvojenog gradiva staviti u kontekst u čijoj pozadini leži neka matematička struktura.

3 Što je produkt obrazovanja?

Obrazovne ustanove se često uspoređuju s poduzećima čiji proizvodni proces nije pojmljiv ili ga nije moguće kontrolirati. Ustanove sličnog tipa su bolnice, banke... *Efikasnost* takvih ustanova mjeri se količinom *outputa* podijeljenog s količinom *inputa*. Uspoređivanje na temelju tako definiranog pojma *efikasnosti* još je i razumljivo ako postoji samo jedna vrsta *inputa* i samo jedna vrsta *outputa* kao na primjer u slučaju bolnica ako je *output* broj pacijenata, a *input* broj osoblja. Ako kao *input* promatramo broj liječnika i broj sestara, a kao *output* broj vanjskih i unutarnjih pacijenata onda je pojam *efikasnosti* nejasan ako nije poznata vrijednost liječnika u terminima sestara ili obratno.

Jedan od načina, danas često korišten i uvažavan je *analiza omeđivanja podataka*². Osim što mjeri efikasnost ustanova s više *inputa* i više *outputa*, ta metoda pronalazi i primjere efikasnog ponašanja ustanova u koje bi se neefikasne ustanove trebale ugledati³. Ono što je u toj analizi jedino bitno je da ustanove koje se uspoređuju, djeluju u istom ekonomskom okruženju (država, regija).

Ako govorimo o obrazovnim jedinicama (školama, fakultetima, univerzitetima) onda nije posve jasno što bi trebao biti *input*, a što *output*. Proizvod koji je rezultat obrazovnog procesa je svjedodžba s imenom i prezimenom učenika ili, analitički rečeno, svi oni parametri koji definiraju takvog učenika. Ulazni parametri mogli bi biti nastavnici i pomoćno osoblje, njihove plaće, materijalni troškovi... Netko će možda reći da je izlazni produkt znanje i vještine koje učenici posjeduju. Čak i da jest tako, teško je, ako ne i nemoguće, mjeriti efikasnost na gore definiran način jer znanje i vještine koje učenik na izlasku iz škole posjeduje nisu samo posljedica školovanja, već i posljedice utjecaja neposrednog i šireg učenikovog okruženja, kao i posljedice samo-obrazovanja. Iz toga proizlazi neposredan zaključak da su svi koncepti koji 'mjere' znanje i vještine učenika i donose zaključke o kvaliteti školovanja ograničeni, jednostrani i neučinkoviti.

Kako uopće organizirati nastavu i obrazovne procese ako nije jasno što je *output* niti kako ga mjeriti? Umjesto orijentiranosti na *output* prihvatljivije je govoriti o svrsi i ciljevima obrazovanja. Poteškoća u ostvarivanju tih ciljeva je taj što nisu uvijek u skladu s interesima učenika i ostalih sudionika obrazovnog procesa. Ako je cilj obrazovanja postizanje znanja, a učenik je *orijentiran na nastup*, v. odjeljak 6 i ima podršku roditelja i šireg okruženja za takvo ponašanje, onda će biti vršen pritisak na nastavnika da mu se (učeniku) povećaju ocjene bez obzira na znanje. Ocjene postaju neselektivne i gube smisao vlastitog postojanja.

Čini se da glavni problem leži u ljudskom poimanju i shvaćanju procesa na najapstraktnijoj razini jer još uvijek prevladava mehanicistički pogled na svijet i zbivanja u formi *uzrok* \mapsto *posljedica*. U kontekstu obrazovanja sudionici procesa su: učenici, nastavnici, škola, roditelji i interesne skupine, uključujući i državu. Prava djece u obrazovanju regulirana su zakonskim aktima, ali je odgojna komponenta škole potpuno zanamarena. Pravila ponašanja sudionika u obrazovanju nisu jasno artikulirana niti od strane škole niti od strane roditelja, a roditelji ne doživljavaju školu kao nešto u čijem bi modeliranju mogli aktivno sudjelovati. Njihova komunikacija sa školom uglavnom se svodi na roditeljske sastanke. Fizičko i psihičko nasilje, demotivacija i neuspjeh učenika u školi su posljedice poremećaja zdravog funkcioniranja i društva i pojedinca i dokaz da nisu stvoreni uvjeti u kojima cvijeta motivacija i zadovoljstvo spoznaje.

Ako je problem dijagnosticiran to ne znači da postoji pilula koja ga rješava. Ako nije promijenjen način života onda će nakon nekog vremena trebati još jedna pilula, pa još jedna... U posljednje vrijeme novinari javnih glasila u Hrvatskoj prozivaju Ministarstvo znanosti i obrazovanja (MZOŠ) što ne poduzima ništa povodom 'loših' rezultata hrvatskih petnaestogodišnjaka na PISA testovima. MZOŠ nije funkcionalan jer nema viziju obrazovanja u Hrvatskoj, a i bolje je da ništa ne poduzima

²eng. *Data Envelopment Analysis*, skraćeno DEA.

³eng. benchmarking

jer PISA rezultati nisu problem, samo su naznaka da naši učenici ne koriste svoje znanje na način kako bi to možda htjela svjetska ekonomska elita.

Kakve to mjere i promjene obrazovnog sustava MZOŠ može pokrenuti. U prvom redu to bi trebale biti mjere koje se mogu pretočiti u oblikovanje neposrednog nastavnog procesa. Ako nema izvršitelja nastave sposobnih za takvu provedbu reforma je neizvediva, a dugoročno gledano, tada se ruši sadašnji sustav koji funkcioniše na svima poznat način. Osim ciljeva obrazovanja, potrebno je definirati i proceduru izvršavanja promjena i ugraditi kontrolu procedure, ali ne u formi vanjske kontrole, nego na način da obrazovni proces postane jedan autopoietički sustav.

4 Ciljevi vrednovanja u obrazovanju

Što je cilj vrednovanja u obrazovanju? Ocijeniti trenutačno znanje, vještine, postignuća, urođene vještine, neke buduće sposobnosti, da služi u neke pedagoške svrhe ili kombinacija svega toga. Naivno gledajući, primarni cilj vrednovanja je doći do argumentiranih i opravdavajućih zaključaka o učenikovim sposobnostima, interesima i postignućima. Ciljevi vrednovanja proizlaze i u uskoj su vezi s ciljevima obrazovanja općenito, a njihova sekundarna uloga je unapređivanje obrazovanja.

Trenutačni stavovi o vrednovanju u obrazovnoj literaturi su pod utjecajem psihometrijskih koncepata i teorija korištenih u analizama psihometrijskih testova, kao IQ test na primjer. Te teorije zasnovane su na vjerovanju da:

- unutar svakog pojedinca postoji nešto što se može zvati 'sposobnošću' ili 'postignućem' u izvršavanju postavljenih zadataka,
- taj entitet je mjerljiv na objektivnoj, jednodimenzionalnoj skali,
- cilj testa je sakupiti podatke na temelju kojih ga je moguće procijeniti.

French (1989) smatra da cilj javnih ispita (u Engleskoj i Walesu) nije objektivno mjerenje nečega o kandidatu i da ne postoji neki entitet unutar kandidata koji može biti mjereno na taj način. Prema Frenchu, cilj takvog ispita je iznijeti prosudbu donešenu od strane ispitivačke komisije. Bodovi i ocjene ne predstavljaju mjeru entiteta već ispitivačevu prosudbu kvalitete ispitanikove 'predstave', kako je oni vide, u ispitanikovom testu ili nekoj drugoj formi pokazivanja sposobnosti. Dalje, French iznosi da

... ta distinkcija nije nevažan i ezoterijski akademski stav; ona ima značajne implikacije za formu mnogih numeričkih procedura korištenih kod manipulacije u postupku vrednovanja. Te procedure nisu tu da procjenjuju kandidatove sposobnosti u prisutnosti 'mjernih grešaka' niti imaju za cilj procjenu ispitivačevih sudova. Njihova je uloga suptilnija. One pomažu ispitivačima da bi formirali pravedne i konzistentne procjene.

4.1 Standardizacija vrednovanja znanja

Na koji način psihometrijski principi utječu na evaluaciju i razvoj ispitivanja znanja u školi kao što je državna matura na primjer? Bez obzira jesu li rezultati ispita prosudbe ili mjerenja, one trebaju biti *pouzdanе, valjane, standardizirane i oslobođene pristranosti*, što god to značilo u ovom trenutku⁴.

Što se standardizacije procesa vrednovanja u obrazovanju tiče tu ima nekoliko pristupa. U literaturi se najčešće spominju dvije forme standardizacije: *normativno* orijentirana i *kriterijski* orijentirana. U praksi se gotovo uvijek radi o mješavini jedne i druge.

⁴O tim pojmovima je dovoljno rečeno u obrazovnoj literaturi. Pitanje je u kojoj mjeri ti zahtjevi nisu zadovoljeni i zašto ih se prešutno zanemaruje.

Normativni test. Normativno orijentirani test (eng. *norm-referencing*) daje rang ispitanika u odnosu na rezultate testa na nekoj predefiniranoj grupi ispitanika koji su bili ispitivani u istim uvjetima (normativna grupa). Očigledan nedostatak normativno orijentiranog testa je taj da ne može mjeriti napredak ispitanika. Prednost takvog testa je što i studenti i profesori znaju što se očekuje od testa i kako će test biti ocjenjivan. Uspjeh na takvom testu izražava se najčešće u *percentilima*.

Tipični primjer takvog testa je test inteligencije. Izborom normativne grupe već je intuitivno jasno što se tim testom želi mjeriti. Mišljenje autora je da škole ne bi trebale koristiti takve testove kao mjeru znanja jer zaobilaze zacrtani kurikulum i ne ukazuju na to što bi studenti trebali znati. Njihova je glavna namjena da razlikuju ispitanike i rang postignut na takvom testu ne govori o tome što student zna a što ne zna.

Neke institucije za normativnu grupu uzimaju grupu ispitanika⁵. Defekt takvog pristupa je taj da ocjena dobivena na ispitu danas i nakon godinu dana ne odražava jednaku kvalitetu ispitanika.

Kriterijski test. Kriterijski orijentirani testovi (eng. *criterion-referencing*) namijenjeni su da provjere u kojoj je mjeri osoba naučila specifično područje ili ovladala zacrtanim vještinama. Najčešće se sastoji od pitanja s više ponuđenih odgovora ili dozvoljava kratke rečenice kao odgovor. Pojam 'kriterij' ovdje ne treba miješati s prolaznim pragom kojim se zahtijeva od studenta da sakupi minimalan broj brojeva kako bi zadovoljio na testu, već se misli na vještine koje se provjeravaju. Mnogi državni testovi u raznim zemljama koriste tu vrstu testova.

Čak i ako testiraju bliska znanja i vještine, test koji je dizajniran za provjeru sposobnosti koristit će drugačija pitanja nego test koji je namijenjen diferencijaciji ispitanika. Razlog tome je što neka pitanja bolje odražavaju aktualna postignuća ispitanika, a neka druga pitanja su bolja za diferencijaciju između 'dobrih' i 'loših'. Mnoga pitanja čine i jedno i drugo. Kriterijski orijentirani test će koristiti pitanja na koja će korektno odgovoriti studenti koji poznaju određenu materiju, a normativno orijentirani test će koristiti pitanja na koja korektno odgovaraju 'najbolji', a nekorektno 'najlošiji' studenti.

Zanimljiva je sljedeća misao (Linn, 1993):

Općenito se priznaje da termini 'kriterijsko' i 'normativno', primijenjeni na instrumente vrednovanja, vode u zabludu jer se interpretacija daje ocjenama, a ne samim instrumentima.

Još neki tipovi referenciranja. U posljednje vrijeme sve više se govori o *limen referencing* u kojem se profil kandidata uspoređuje s tipičnim zamišljenim predstavnikom klase ispitanika za svaku ocjenu. Taj tip referenciranja postaje zanimljiviji utoliko što je porasla svijest o važnosti kurikula i sličnih standarda u obrazovanju. Važnost takvog referenciranja u ocjenjivanju dolazi do izražaja u fazi ocjenjivanja tzv. *graničnih testova*, v. odjeljak 10.3.

Slabo kriterijsko referenciranje je oslabljeno kriterijsko referenciranje koje nastoji zadržati kvalitetu i nivo izvođenja ispita. Pri tome se nastoji zadržati težina samog ispita ali se ne zahtijeva potvrda specifičnih znanja i vještina.

Construct referencing je način dodjeljivanja ocjena koji je fleksibilniji od kriterijskog, a uvažava i projektne zadatke i subjektivne dojmove o njima.

Niti jedan od navedenih tipova nije referentni sustav koji daje značenje ocjeni, to su metode koje pomažu u određivanju granica među ocjenama.

4.2 Ocjena, što je to?

Ljudi su skloni vjerovati da je ocjena kvantitativna mjera sposobnosti ispitanika. Sama po sebi ocjena ne govori ama baš ništa o ispitaniku. Ako ispitanika i njegovu ocjenu stavimo u neki kontekst

⁵eng. *cohort referencing, cohort—drug, ortak*

(školu, profesora, fakultet, sustav...) tada o njemu stičemo neki dojam. Zapravo, mijenjamo sliku o sebi jer sebe stavljamo u kontekst iz kojeg promatramo osobu kroz prizmu njegove ocjene.

Ocjena dobivena na ispitu/testu, pretpostavljam, zamišljena je kao mjera koja odražava stupanj ovladanosti zacrtanih obrazovnih ciljeva od strane kandidata. Način ocjenjivanja dio je kulturnog nasljeđa institucije, sustava, države. Ocjena kao pedagoški pokazatelj i motivacija u procesu savladavanja gradiva (interno ocjenjivanje) dio je obrazovnog sustava i nema drugih korisnika osim učenika, njegovih staratelja i nastavnika.

Ocjena na državnom ispitu/maturi (vanjsko ocjenjivanje) i opći uspjeh u srednjoj školi, kao informacija koja prati studenta i markira njegov budući život, koristi se u mnogim situacijama i korisnici te ocjene manipuliraju s njom na način kako to njima odgovara. Tako na primjer, fakulteti mogu i ne moraju koristiti uspjeh u srednjoj školi i/ili maturi prilikom upisa na fakultet. Ovdje se neka pitanja i sugestije otvaraju sami po sebi:

- Odražava li državna matura dinamiku promjena u društvu?
- Razne institucije imaju potrebu za različitim profilima zaposlenika. Nije li bolje da država osigura intrumente evaluacije koje će institucije unajmljivati?
- Takvi instrumenti evaluacije mogu se lakše internacionalizirati i modularno nadograđivati.

4.3 Svrha ocjenjivanja

Čemu služi ocjenjivanje? U *dokimologiji* i uputama za nastavnike najčešće se na to pitanje odgovara da je ocjena:

- rezultat vrednovanja učenikovih postignuća,
- način komuniciranja sa studentom, njegovim roditeljima i budućim poslodavcima o učenikovim sposobnostima u predstavljanju njihovog znanja,
- pokretač učenikove motivacije za daljnje napredovanje u učenju,
- polazište u organizaciji nastave i nadogradnji znanja.

Odgovori su suhoparni, vrlo općeniti, podrazumijevaju da čitatelj razumije što je to *znanje, vrednovanje, motivacija, nadogradnja znanja* i ne govore o tome kako se ocjenjivanje uklapa u ciljeve obrazovanja (koji uglavnom nisu ili su rijetko kad navedeni). U prvom odgovoru naslučuje se da je ocjena krajnji ishod vrednovanja, a procedura vrednovanja uključuje ocjenjivanje kao svoju bitnu komponentu u gotovo svim obrazovnim razinama. Je li to znači da su vrednovanje i ocjenjivanje sinonimi? Jedna od bitnih komponenti motivacije je *reciprocitet* u smislu potrebe za suradnjom u grupi kako bi se završio neki zadatak. Je li organizacija nastave takva da to dozvoljava? Je li ocjenjivanje zaista u službi pokretanja motivacije? O ocjeni kao kazni se već dugo raspravlja ali nema smjernica kako izbjeći kažnjavanje učenika davanjem loših ocjena. Što se komunikacije tiče ona zahtijeva razgovor i suradnju, ne samo između učitelja i učenika i njegovih roditelja, nego i među samim učenicima u procesu ovladavanja znanjem. Ocjena kao broj često predstavlja kraj komunikacije, a ne njen početak.

Stručni tekstovi o ocjenjivanju uglavnom daju upute o tome kako ocjenjivanje učiniti *transparentnim, učinkovitim, manje stresnim* (i za učitelje i za učenike), kako organizirati pitanja za različite forme testova... Nisam siguran da obrazovna psihologija može išta doprinijeti u formi recepata i uputa nastavnicima ako im se ne dozvoli da eksperimentiraju i kreiraju svoje vlastite ispitne metode. Umjesto da se raspravlja o tome kako usavršiti ocjenjivanje, bilo bi korisnije raspravljati o provjeri stečenih znanja i sposobnosti kao funkciji u organizaciji nastave i nastavnog sadržaja. To je dinamički proces u smislu da nedovoljno dobra razina usvojenosti od strane učenika ima kao posljedicu učenikovu nesposobnost daljnjeg sticanja znanja.

Koji god aspekt ocjenjivanja, od gore spomenutih, razmatramo on je orijentiran na rezultate nekog djelovanja. Ocjena je posljedica rezultata djelovanja i zanemaruje djelovanje kao proces. Proces učenja ima vrijednost sam po sebi. Ako rezultat tog procesa ukazuje učeniku da ne treba nastaviti u smjeru u kojem je krenuo to je za njega jednako vrijedno kao i potvrda da je krenuo u dobrom smjeru. U svakom učenju naizmjenice se izmjenjuju povratne informacije jedne i druge vrste, a škola kao institucija je slijepa na tu kvalitetu nastavnog procesa. Na dijete se vrši psihološki pritisak za postizanjem rezultata od strane nastavnika, roditelja i od budućih poslodavaca. Rezultat! Rezultat!... Rezultat! Preopterećenost rezultatima vodi u krah, a usredotočenost na djelovanje generira suštinsku unutarnju vrijednost. Orijentacija na rezultat ukazuje na cilj koji (još) nije ostvaren, orijentacija na proces učenja je bivanje u sadašnjem trenutku.

4.3.1 Održavanje standarda ili razvijanje sposobnosti?

Greatorex (2003) zaključuje da sva nastojanja u praksi i teoriji, pod nazivnikom 'pravednijeg' ocjenjivanja, imaju u pozadini očuvanje standarda radije nego interpretaciju postignuća (rezultata).

Možda bi naglasak trebao biti na poboljšanju kvalitete svih sudionika u obrazovnom sustavu. U tom smislu bilo bi poželjno osmisliti i kreirati moderno i kvalitetno vrednovanje koje će pokazivati ispitaničeve sposobnosti i znanje radije nego ih referencirati prema standardima. To opet kreira nove probleme jer je diskutabilno što znači termin 'sposobnosti'. S druge strane, promjene u obrazovnom sustavu zahtijevaju dodatni napor i za učenike i za nastavnike i čine ga teško razumljivim, ako ni zbog čega onda zbog pomanjkanja kontinuiteta, na primjer. U promjene treba ulaziti s jasnim zahtjevima od zainteresiranih (autoriteta) što je to obrazovni standard, a od profesionalaca ocjenjivača da razmotre valjane i stabilne metode vrednovanja i ocjenjivanja. Ono što se pri tome najčešće zaboravlja je tko su sve korisnici ocjena i kakvi su njihovi interesi. Jedna od zabluda, ne samo u hrvatskom obrazovnom horizontu, je da javni ispiti (matura) mogu mjeriti i kvalitetu škola i nastavničkog kadra u njima. Čak se ta 'komponenta' mature ističe kao jedna od važnijih. U svrhu određivanja kvalitete škola treba razvijati drugačije mjerne instrumente.

5 Hijerarhija odluke i potciljevi

Pretpostavimo da želite kupiti digitalnu kameru (KupiKameru). Nakon dugotrajnih razgovora s prijateljima i pretraživanjem web portala ostale su vam dvije opcije (alternative):

- A) Kamera A je slabijih mogućnosti (6 osnovnih funkcija), ali je jednostavna za rukovanje.
- B) Kamera B je naprednija (6 osnovnih i 6 dodatnih funkcija) ali je kompliciranija za rukovanje.

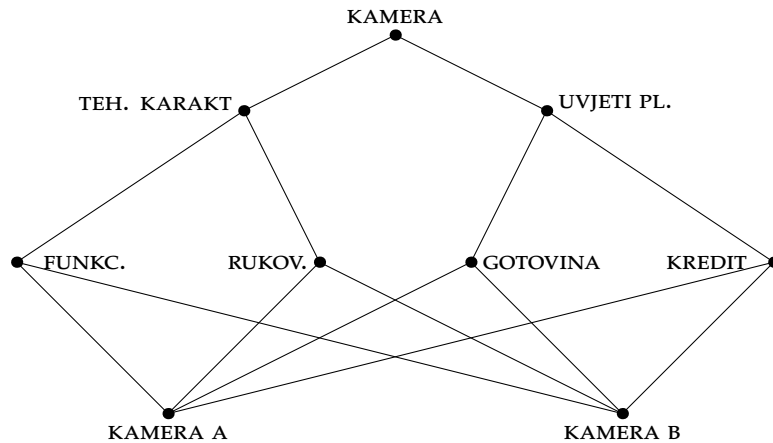
Vi se dvoumite jer nemate iskustva s kamerama i ne možete procijeniti što vam je važnije: funkcionalnost (Funkc.) ili lakoća rukovanja (Rukov.). Problem se može još i dodatno zakomplicirati ako je jedna kamera nešto jeftinija, ali drugu možete kupiti na mjesečne rate. U tom slučaju dobili smo još dvije dodatne informacije: gotovinsko plaćanje (Gotovina) i otplata na kredit (Kredit).

Svi kriteriji nisu istog tipa. Funkc i Rukov su tehničke karakteristike kamere (Teh. Karakt), a Gotovina i Kredit su uvjeti plaćanja (Uvjeti pl.). Proces odluke moguće je hijerarhijski organizirati na način da sve elemente odlučivanja: kamere, tehničke karakteristike i uvjete plaćanja rasporedimo po razinama:

- 0. razina: (Cilj) Kupnja kamere
- 1. razina: (Tipovi atributa) Teh. karakt i Uvjeti pl.
- 2. razina: (Atributi) Funkc, Rukov, Gotovina, Kredit
- 3. razina: (Alternative) Kamera A i Kamera B

Tipovi atributa su potciljevi glavnog cilja: Kupnja kamere. Kažemo da je Kupnja kamere *korijen*, a Teh. karakt i Uvjeti pl. su njegovi *listovi*. Isto tako su Funkc i Rukov *listovi* od Teh. karakt, a Gotovina i Kredit su *listovi* od Uvjeti pl.. Kamera A i Kamera B su *listovi* od svakog atributa: Funkc, Rukov, Gotovina i Kredit. Na slici 1 dana je hijerarhijska struktura odluke

Slika 1: Hijerarhijska struktura odluke za kupnju kamere.

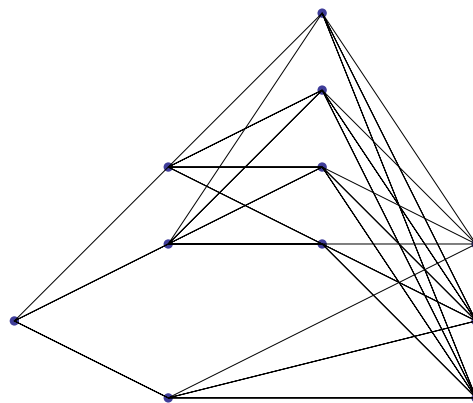


navedenog primjera za kupnju kamere.

U hijerarhijskom modelu odluke svaki element odlučivanja je u jednoj i samo jednoj razini. Najviša razina je ona s najmanjim indeksom (Cilj), a najniža razina je ona s najvećim indeksom (Alternative). Svaki element odlučivanja, osim alternativa, je korijen za jednog ili više elemenata odlučivanja iz iste (niže) razine. Svaki element odlučivanja (osim Cilja) mora biti list bar jednog korijena. Za korektnu definiciju hijerarhijske odluke v. (Čaklović, 2014, str. 454).

Na slici 2 dan je primjer hijerarhijske odluke s jednim ciljem (sasvim lijevo) koja se širi slijeva udesno, a listovi su uređeni od nižeg prema višem. Svaki čvor hijerarhije, osim onih u posljednjoj

Slika 2: Primjer hijerarhijske struktura s jednim ciljem. Orijehtacija hijerarhije je s lijeva na desno. Prvi čvor druge razine ima djecu u posljednjoj (četvrtoj) razini.



razini, ima svoju djecu u nekoj od nižih razina. Čvor (2, 1) (prvi čvor u drugoj razini) ima listove u posljednjoj (četvrtoj) razini zaobilazeći treću razinu.

5.1 Mišljenje orijentirano na proces

Taylor et al. (1998) definiraju mišljenje orijentirano na proces kao mišljenje koje uključuje korak-po-korak razrađenu proceduru akcija koje dovode do ostvarenja cilja. Ako kormilo određuje pravac kretanja broda (cilj), onda su vesla proces — to je simpatična metafora. Ako je cilj odgovor na pitanje: "Što?" onda je proces odgovor na pitanje: "Kako?".

Prvotna istraživanja sugeriraju da je procesno orijentirana mentalna stimulacija efikasnija od mentalne stimulacije orijentirane na posljedice u smislu efikasnijeg ostvarivanja ciljeva, samoregulacije i hvatanja u koštac s ostvarenjem odluke. Na primjer, studentima koji su se pripremali za ispit bilo je sugerirano da zamišljaju kako se pripremaju za višu ocjenu (sjede za stolom, listaju knjigu, prolaze gradivo). Takvi studenti su više vremena posvetili učenju i postigli su bolje rezultate od studenata kojima je sugerirano da zamišljaju kako su već postigli dobre ocjene (osjećaju se zadovoljno zbog toga, puni su samopouzdanja) (Taylor et al., 1998).

Novija istraživanja nude nešto drugačije zaključke. Thompson et al. (2009) proveli su nekoliko studija i svaka od njih je rezultirala jednakim zaključkom. Prikazujemo jednu od njih:

Grupi studenata (iz istog kampusa) bila su ponuđena 2 stana na izbor s jednakom najamninom. Stan A je nešto manji (50 m²) ali bliži kampusu (1 km), dok je stan B veći (90 m²), ali udaljeniji od kampusa (9 km).

Ispitanicima u grupi **orijentiranoj na proces** bilo je sugerirano sljedeće:

Usredotočite se na to kako će život u stanu A (B) utjecati na vaše svakodnevne rutine i navike. Zamislite sebe u tom stanu i usredotočite se na korake u obavljanju vaših dnevnih rutina i kako se pri tome osjećate.

Ispitanicima u grupi **orijentiranoj na posljedice** bilo je sugerirano:

Usredotočite se na prednosti življenja u stanu A (B) i što time dobivate. Zamislite sebe kao stanara i fokusirajte se na dobrobiti koje vam donosi bivanje u tom stanu.

Nakon što su ispitanici iznijeli svoje mišljenje za svaki stan, zamoljeni su da ocijene složenost (težinu) mentalne simulacije na skali lako/vrlo teško, kao i stupanj uvjerenja u ispravnost svoje odluku na skali (neuvjerljivo/vrlo uvjerljivo). Na kraju su bili upitani o složenosti i težini fokusiranja na veličinu stana i udaljenost od kampusa na skali (uopće ne/vrlo teško).

Sudionici grupe usmjerene na proces svjedočili su o povećanoj težini mentalne stimulacije, jednako kao i o povećanoj težini fokusiranja na attribute stana. Ispitanici u toj grupi izabirali su u većini slučajeva manji stan (91%), a ispitanici ciljno orijentirane grupe izabirali su veći stan (70%). Sudionici grupe usmjerene na proces također su svjedočili o povećanoj složenosti (težini) odlučivanja zbog povećane koncentracije na attribute stana. Drugim riječima, grupi orijentiranoj na proces bilo je daleko teže donijeti odluku.

U dvije različite studije (učenje i izbor stana), ponuđeni su suprotni zaključci što se tiče **efikasnosti mentalne usmjerenosti na proces**. Kako to objasniti? Kao prvo, te dvije studije se bitno razlikuju. U prvoj studiji nije bila ponuđena alternativa i nije bilo suprotstavljenih atributa (površina vs. udaljenost) što može ukazivati na povećanu složenost procesa odlučivanja u drugoj studiji. U ovom trenutku ništa više i bolje se ne može zaključiti jer **nije ponuđeno dublje značenje orijentacije na proces** osim korak-po-korak upute.

Pogledajmo malo detaljnije kako razni autori shvaćaju *proces* ili *orijentaciju na proces* i kako je taj pojam vezan uz *ciljeve*:

- Možete misliti na proces kao na naviku koju treba uvježbati. Napr. ako ste odlučili smršaviti možda bi bilo korisno da se više krećete i pazite na prehranu. U fazi uvježbavanja navike vi sebi postavljate za cilj da postignete nešto što će vam na dugu stazu donijeti zadovoljavajuće rezultate.
- Kada nastojite postići neki cilj mogu vam se u pozadini vrtiti ovakve misli: "Još nisam dovoljno dobar, ali kad postignem cilj bit ću uspješan." Neki tvrde da postavljanje ciljeva reducira našu sreću. Drugi će reći da nas postavljanje ciljeva odvlači od sadašnjosti.
- Za ciljno orijentirane ljude rast i napredak se dešava do mjere koju im dopuštaju njihovi ciljevi. Ljudi orijentirani na procese postižu ciljeve jer ih proces tamo vodi u svakom slučaju. Za njih rast nije uzgredan, on je nužan, a ciljevi su na kraju procesa.

"Uspjeh je put, a ne destinacija." — kaže narodna izreka. To je zato, jer je uspjeh zagarantiran u procesu. Najveća mjera osobnog uspjeha nije ono što je čovjek postigao, već ono što je u tom procesu postao.

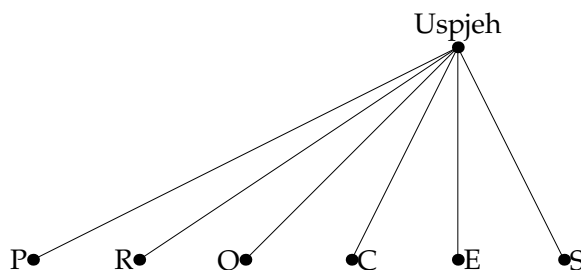
U nekim od tih 'definicija' proces je navika ili nešto što treba uvježbati. Drugim riječima, proces je postao automatizam i ugrađen u strukturu odluke. On više nije ni faktor niti varijabla nego dio uvjerenja. Ako proces promatramo u procesu uvježbavanja, to što uvježbavamo je cilj kojem težimo. **Procesu je potreban cilj** bez obzira kako na njega gledamo.

"Uspjeh je put, a ne destinacija" je lijepa rečenica koja nas tjera da razmislimo o njenom značenju i o našim uvjerenjima, ali u sebi nema nikakvu naznaku da bi se išta moglo 'izmjeriti'. Osnova bilo kakvog mjerenja je uspoređivanje (relacija), a cilj kao pojam to u sebi nosi. Tvrdimo da je **ciljno mišljenje osnovna misaona struktura** koju možemo prepoznati i u procesu i u svakom strukturiranom mišljenju.

5.1.1 Proces kao hijerarhija.

Pogledajmo kao motivacijski primjer studiju od Taylor et al. (1998). Jednoj grupi ispitanika u toj studiji je sugerirano da se usredotoče na proces, tj. na to kako sjede za stolom, prelistavaju udžbenik i ponavljaju gradivo za ispit. Te upute su konkretne, razumljive i sastavni su dio uspješnog savladavanja gradiva i mogu se shvatiti kao potciljevi uspjeha, v. hijerarhiju na slici 3. Na slici su

Slika 3: Proces kao hijerarhija.



dani generički nazivi potciljeva: *P, R, O, C, E, S*. Shvaćeni kao elementi odlučivanja u hijerarhiji, ti se elementi mogu vrednovati i rangirati po važnosti pa je moguće staviti naglasak na neke od tih procesa kao važnije u postizanju uspjeha. U ovakvom shvaćanju procesa jasno je zašto su studenti orijentirani na proces imali bolje rezultate na ispitu od studenata koji su bili usmjereni samo na cilj. Studenti usmjereni na cilj sami kreiraju neki misaoni proces kako postići uspjeh ili djeluju po ustaljenim navikama koje ne traže svjesnu angažiranost, ali ne vode nužno ka efikasnijem ostvarenju cilja.

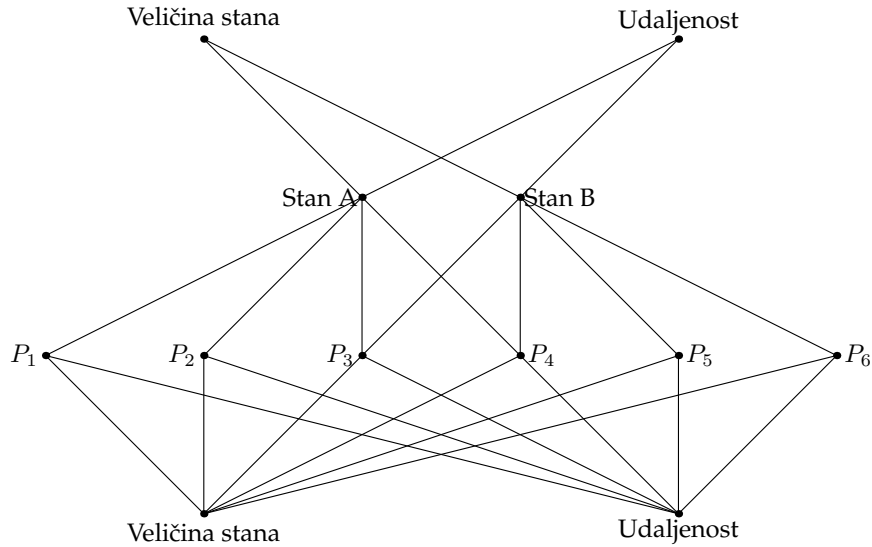
Mnogi procesi, naročito u proizvodnji, su strogo hijerarhijski i neki elementi procesa se ne mogu nastaviti ako prethodni nisu obavljeni. Takvi, algoritamski procesi, nisu predmet teorije odlučivanja.

Slikoviti primjer za proces je nabava hrane. Sigurno ste bar jednom u životu bili na tržnici s namjerom da kupite hranu za goste koje ste pozvali na sutrašnji rođendan, ali imate samo grubu sliku što biste im pripremili. Ako ste vrsna kuharica, dovoljna je jedna šetnja tržnicom i vi ćete imati mentalnu sliku svih jela koja biste mogli ponuditi. Voće i povrće na tržnici su samo sastojci, a vrsna kuharica u njima vidi gotova jela jer poznaje proces pripreme hrane. Ona kupuje sastojke za svoju hranu i ti sastojci su ništa drugo nego potciljevi-djelovi-procesa. Ako niste vrsna kuharica vaši gosti će se morati zadovoljiti voćem i miješanom salatam. Sami zaključite kome je lakše donijeti odluku: vrsnoj kuharici ili nekome tko zamišlja gotova jela.

Model odluke u studiji sa stanovima (Thompson et al., 2009) prikazan je na slici 4 (vidi str. 10 za opis studije). Svaki od stanova generira neki proces u glavi ispitanika. Za razliku od prethodne studije, ovdje nije sugeriran neki precizniji proces, već je ostavljeno ispitaniku da ga sam kreira.

Pretpostavimo da je stan A generirao, u glavi ispitanika, proces: (P_1, P_2, P_3, P_4) , gdje je proces identificiran s potciljevima (listovima). Na isti način je stan B je generirao proces: (P_3, P_4, P_5, P_6) . Elementi

Slika 4: Model mentalne simulacije orijentirane na proces kod izbora stana.



odluke P_3, P_4 su zajednički za oba procesa, što ne mora nužno biti slučaj, ali u ovom primjeru je tako. Procesu mogu biti disjunktni, tj. nemaju zajedničkih potciljeva ili mogu biti identični. I jedno i drugo su krajnosti koje su moguće. Poteškoća u kreiranju procesa proizlazi iz kompetitivnih informacija (ciljeva): *Udaljenost* i *Veličina stana*. Što se udaljenosti tiče povoljniji je stan A, a što se veličine tiče povoljniji je stan B.

Fokusiranost ispitanika na jedan od stanova u drugom retku hijerarhije aktivira njemu odgovarajuće procese u trećem retku hijerarhije, a svaki od elementarnih procesa P_i uključuje i *Udaljenost* i *Veličinu stana*. Na primjer, ako ispitanik posjeduje bicikl i zamišlja kako ga unosi u stan (proces), onda mu je možda važnija *Veličina stana* od *Udaljenosti* jer bicikl negdje treba smjestiti u stan. Ako ispitanik ne posjeduje nikakvo prevozno sredstvo možda mu je važnija *Udaljenost* kao kriterij. Drugačije rečeno, to znači da među elementima odluke postoji povratna veza što dodatno komplicira model.

Ovakav hijerarhijski model sugerira da je generiranje mentalnih modela orijentiranih na procese, u slučaju dvije alternative i dvije kompetitivne karakteristike alternativa, kompliciranije i teže nego generiranje modela orijentiranih na rezultat (prve dvije razine u hijerarhiji), što i jest zaključak u Thompson et al..

O takvim hijerarhijama s povratnim vezama, kao što je ova na slici 4, bit će još riječi u odjeljku 11.7 pod nazivom **Samodualna hijerarhija**. Napomenimo da u gornjoj hijerarhiji, prva razina ima više od jednog elementa (dva) i nisu poznate njihove težine. To znači da vrednovanje hijerarhije, kao što je opisano u odjeljku 11.6, nije moguće bez tih težina. Ono što razlikuje samodualnu hijerarhiju od obične je to što prva i posljednja razina u samodualnoj hijerarhiji imaju iste elemente. Ako se preciziraju neke polazne težine elemenata u prvoj razini, mogu se izračunati težine elemenata u posljednjoj razini, što revidira polazne težine. Ponovimo li proces vrednovanja hijerarhije s revidiranim težinama, dobit ćemo novu reviziju težina tih istih elemenata. Ponavljanjem tog postupka težine se stabiliziraju i granična vrijednost težina ne ovisi o početnom izboru (v. odjeljak 11.7).

Neke od 'definicija' procesa na stranici 10 spominju čovjekov rast i napredak. Imajući u vidu samodualnu strukturu, promjene težina (odnosa) među elementima odlučivanja je moguće pratiti iz jedne revizije u drugu (u ciklusima), a te promjene slobodno možemo interpretirati kao napredak, učenje ili sazrijevanje.

5.2 U potrazi za ciljevima

Kako shvatiti ovaj naslov? Ako se radi o pokretanju nekog novog posla i ako imam mogućnosti (sposobnost + materijalna osnova), koji je doseg mog poslovanja? To je svakodnevno pitanje svakog poduzetnika i poslovna literatura je puna savjeta što činiti. Odgovor je svuda isti: postavi si ciljeve i procese kako ih postići. Ako i postavimo neke moguće ciljeve, tada je logično pitanje koje od njih je lakše⁶ ostvariti pomoću zadanih mogućnosti. U svakom slučaju radi se o povratnoj vezi između ciljeva i mogućnosti koji treba modelirati i evaluirati. Današnja poslovna psihologija nije u stanju niti modelirati niti riješiti tako kompleksan zadatak.

Locke and Latham (1990) razvili su *Goal setting theory* s osnovnom namjerom da kompanije razviju metode za povećanje motivacije svojih zaposlenika u svrhu postizavanja zadanih ciljeva, te kako i na koji način te ciljeve formulirati. Bez obzira na jasnoću i razumljivost njihovih uputa, **osnovni problem danas u svijetu i dalje ostaje nemotiviranost**. Jedan od glavnih nedostataka postavljenih ciljeva je taj što generiraju mentalnu krutost zaposlenika i onemogućavaju njihovu autonomiju, a to je opet prepreka u prihvaćanju i interiozaciji tih ciljeva kao generatora kreativnosti i unutrašnjeg zadovoljstva.

Pitanje je može li se naučiti nešto iz poslovne psihologije i primijeniti to u obrazovanju? Može da je to i istina, ali su, s jedne strane, problemi obrazovanja toliko kompleksni, a s druge strane ne postoji globalna institucionalna podrška za razumijevanjem i rješavanjem tih problema. To je posljedica orijentiranosti najodgovornijih subjekata i vladajućih struktura na ciljeve kao što su profit i rast BDP-a⁷, kao i na skretanje toka novca poreznih obveznika u privatne škole i nevladine organizacije (slučaj SAD).

Bez obzira radi li se o obrazovnom ili proizvodnom kontekstu orijentacija na ciljeve može ubrzati proces ali je ljudska priroda takva da lako mijenja ciljeve što je posljedica kontekstualnih ograničenja koje ciljevi stvaraju. S druge strane, ciljevi koje postavlja donositelj odluke samo su mala proširenja njegovog sadašnjeg spoznajnog trenutka. Ako mu netko drugi postavi cilj on neće biti u stanju niti približno odabrati načine njegovog mogućeg ostvarenja.

Proces je zahtjevan i po svojoj prirodi složeniji od organizacije ciljeva, ne samo zato što ga treba usvojiti već i stoga što se rezultati procesa u principu razlikuju od zacrtanih ciljeva. Možda bi trebalo razmišljati o procesu transformacije ciljeva tokom procesa iz jednostavnog razloga što u procesu ljudi rastu i sazrijevaju.

6 Motivacija za učenjem

Ako netko nije voljan penjati se uz ljestve sam, uzaludno ga je gurati.

Andrew Carnegie, biznismen i filantrop

6.1 Pojam motivacije u literaturi

Motivacija se definira kao poriv u pozadini nečijeg djelovanja i ponašanja. Ona obuhvaća i naš pristanak i spremnost za pokretanje akcije. Može biti privremena, diktirana trenutačnom situacijom ili može biti trajna osobna kvaliteta. Nekoliko je teorija motivacije koje su manje više komplementarne jedna drugoj i svaka od njih nudi objašnjenja veze između studentove motivacije i njegovih postignuća.

Motivacija je sama po sebi neodređen i nejasan pojam i u principu naznačuje samo da postoji povezanost između aktivnosti i njene svrhe. Razne teorije samo detaljnije pojašnjavaju vidove te povezanosti i nude motivacijske obrasce u obrazovnom procesu s ciljem svrsishodnijeg usvajanja

⁶ili koji od ciljeva su: realniji, donose veći profit, kratkoročno ili dugoročno ostvarljiviji.

⁷Bruto domaći proizvod, makroekonomski pokazatelj napretka.

gradiva i povećanja učenikovih kognitivnih i drugih sposobnosti. Poteškoće u razumijevanju psihološke literature, vezane uz obrazovanje i motivaciju, leže u nejasnoći termina koje teoretičari koriste. Nejasnoća proizlazi iz činjenice da ne postoji dogovor među autorima o značenju određenog pojma, a još je manje jasno u kakvom odnosu ti koncepti stoje jedni u odnosu na druge. U tekstu koji slijedi nastojali smo izdvojiti samo ona istraživanja i koncepte koji čine okosnicu današnje obrazovne psihologije i neka najnovija istraživanja koja iz temelja propituju ova prva.

Murphy and Alexander (2000) proveli su iscrpno istraživanje terminologije vezano uz motivaciju u kojem iznose razlike i sličnosti među istim terminima i konceptima kod raznih autora. Tako na primjer, od 51 studije uključene u njihovo istraživanje, samo 4 (8%) je definiralo eksplicitno taj termin. Implicitna definicija je navedena u 17 (33%) studija, ostatak studija pretpostavlja da je termin 'motivacija' jasan sam po sebi. Od dvadesetak različitih termina oni su izdvojili početnu listu od devet najučestalijih termina i za svaku studiju su komentirali je li navedeni termin eksplicitno, implicitno ili nije definiran i navedena je definicija ako je eksplicitna. Ti termini su (na engleskom): *achievement, affect, attribution, self-competence, self-efficacy, goals, engagement, motivation i self-regulation*.

U razlozima za pokretanje akcije leži njen motivacijski aspekt. Ti razlozi mogu biti u samoj akciji⁸, kao što je zadovoljstvo i osjećaj ispunjenosti ili nagrada kao posljedica akcije, u formi priznanja ili novca. Prvi razlozi su unutrašnji porivi, a drugi su vanjski porivi za akciju. Unutrašnju motivaciju nalazimo u hobiju, čitanju knjiga, sportu (ako volimo izazove) ili učenju ako smo željni novih znanja. Vanjsku motivaciju nalazimo u tjelovježbi čiji je cilj dobra kondicija ili mršavljenje, sticanje vještina zbog novog i bolje plaćenog posla ili poslušnost radi dobivanja nagrade i priznanja od roditelja, šefa ili vlastitih kolega. U vanjsku motivaciju također se svrstava i izbjegavanje obaveza, napr. nedolazak u školu ako bi to rezultiralo lošom ocjenom na testu. Svaka akcija s prefiksom "moram" motivirana je izvana. Korektno ponašanje u saobraćaju, na primjer, je dio našeg ponašanja koje može biti motivirano strahom od kazne (moram), a isto tako odražava i naš karakter i stavove prema drugim ljudima općenito (korektnost).

Unutrašnja motivacija dovodi se u vezu s nesvjesnim i automatski pokrenutim akcijama koje su dio kulture ponašanja, dok su vanjski motivi svjesni i dostupni analizi i argumentaciji (McClelland et al., 1989). Unutarnja motivacija olakšava studentovo samoizražavanje, uključuje i dopušta divergentno mišljenje i omogućava studentu da slijedi vlastiti interes bez brige što će reći njegovi učitelji. Ona dodatno potiče težnju ka samostalnošću i kontrolu nad vlastitim životom.

6.2 Postignuća kao ciljevi

Jedna od raširenijih teorija motivacije je *teorija ciljeva-postignuća* (eng. *Achievement Goal Theory*) koja ciljeve postavlja kao glavne pokretače i nositelje motivacije. *Ciljevi* organiziraju, održavaju i daju smisao našem ponašanju. *Ciljevi* su kognitivna reprezentacija razloga i svrhe naših akcija u različitim područjima: u sportu, poslovanju, učenju, održavanju zdravlja... Konretizacija ciljeva u formi koja se može okarakterizirati kao 'mogućim' ili 'ostvarivim' je poželjna jer je tada, u principu, moguće ustanoviti stupanj postizanja cilja ili mjeru njegovog ostvarenja. Takve ciljeve je lakše prihvatiti, moguće ih je nadograđivati i koristiti u vođenju učenika kroz obrazovni proces i evaluaciji njihovog napretka. Dva su glavna aspekta (dimenzije) takvih *ciljeva-postignuća*: *kompetentnost* i afektivan odnos prema cilju-postignuću ili *valencija*. Valencija može biti pozitivna i negativna, a u literaturi se još koriste termini: *prihvatanje (dostupnost)* ili *izbjegavanje cilja*.

Kompetencija se odnosi na stupanj zadovoljavanja *cilja-postignuća* koji se najčešće formuliraju kao standardi. Standarde usvaja osoba referirajući se na sebe samu, može se uspoređivati s drugima ili može usvojiti standarde koje nameće zadatak (cilj) sam po sebi. Sva tri oblika usvajanja standarda su prisutna u obrazovanju. Učenik koji mjeri svoj vlastiti napredak mjereći svoj uspjeh postignutim bodovima na testu koristi *samo-referencirajuće* standarde. Nastavnik koji ocjenjuje učenike prema Gaussovoj ili nekoj drugoj raspodjeli temelji standarde na uspoređivanju učenika – *norma-*

⁸figurativno rečeno.

tivni pristup. Apsolutni standardi proizlaze iz prirode problema (zadatka) koji se mora savladati kao pretpostavka za daljnje napredovanje, napr. savladano/nije savladano.

U raspravama o obrazovanju, koncept mišljenja i ponašanja učenika, je uglavnom ciljnog karaktera. Iako je krajnji cilj učenja postizanje određene razine znanja i vještina, karakterne osobine učenika i njegova okolina formiraju i *proxi* ciljeve koji mogu biti prilagođeni trenutku, interesu ili organizacijskim mogućnostima. Takve *proxi* ciljeve možemo zamišljati kao bliska zamjena za stvarne ciljeve ili su njima hijerarhijski podređeni.

Ranija istraživanja (Dweck, 1986) zagovarala su dvo-faktorski model ciljne orijentacije koji su prisutni kod učenika. Jedan je faktor težnja za usvajanjem novih znanja i vještina (spoznaja), a drugi je težnja za dobivanjem potvrde od svojih vršnjaka i učitelja koja rezultira ocjenom (nastup). Učenici drugog tipa, osim što traže potvrdu, ujedno i izbjegavaju negativne prosudbe o svom znanju i zalaganju (kasniji istraživači dodaju to kao treći faktor).

U svakodnevnoj školskoj praksi učenik mijenja svoju orijentaciju ovisno o svom dobu, životnoj situaciji i predmetu podučavanja. Učenikov interes za sticanjem znanja nije interes potrošača za potrošnom robom kako to tradicionalisti⁹ razumiju. Naprotiv, "učenik je 'sirovi materijal' obrazovanja kao kipara umjetnika i gravni produkt obrazovne transformacije, a pri tome je još i aktivni sudionik obrazovnog procesa" (Lengnick-Hall and Sanders, 1997).

6.3 Spoznaja kao cilj

Orijentacija na znanje i vještine, kao model, pomaže u kreiranju okruženja poticajnog na učenje i u tom smjeru američki National Research Council (2000) sugerira sljedeće: (1) Učitelji trebaju biti svjesni kakvo znanje i vještine učenici donose dolaskom u školu i (2) fokus usmjeriti na sadržaj, tj. na to što se uči, obrazložiti zašto se to uči i poticati usavršavanje vještina. Dokaz da su učenici savladali gradivo je snalaženje u novim situacijama i problemima. U tom smislu je formativna evaluacija učenikovih postignuća nezaobilazna i ne bi trebala biti u obliku ocjene (zaključak autora).

Pojam *orijentacija na znanje* nije posve jasan u psihološkoj literaturi jer ne razlikuje: motivaciju za znanjem u epistemološkom smislu ili kao posljedicu čisto praktičnog interesa učenika. Možda je bolji termin *Students' Approaches to Learning* (Biggs, 1987) koji razlikuje: (1) *površni pristup* motiviran strahom od neuspjeha, (2) *duboki pristup* motiviran porivom za razumijevanjem i (3) *strateški pristup* motiviran težnjom za uspjehom. Biggs je razvio i odgovarajući upitnik koji se vrlo uspješno koristi u analizama motivacije za učenjem kod studenata. *Spoznaja* kao cilj, u literaturi se nalazi i pod nazivom *vještina*. Drugim riječima, ti se pojmovi isprepliću, nadopunjuju i često je nejasno što se pod tim misli.

6.4 Nastup kao cilj

Nastup (eng. *performance*) razni autori prevode na hrvatski kao *izvedba, demonstracija, učinak* i u daljnjem tekstu koristit ćemo sva tri termina. Unutar *orijentacije na nastup*¹⁰ Elliot (1999) razlikuje *performance approach* ciljeve i *performance avoidance* ciljeve. *Performance approach* ciljevi odnose se na orijentaciju učenika za pokazivanjem visokih sposobnosti, a *performance avoidance* ciljevi odnose se na orijentaciju učenika prema izbjegavanju pokazivanja niskih sposobnosti. Učenici prvog tipa su motivirani da budu bolji od drugih, a učenici drugog tipa izbjegavaju neuspjeh i pokazivanje nekompetencije.

Wilson (2009) je provela studiju među studentima prve godine medicine i potvrdila Eliotov dvo faktorski model strukture¹¹ orijentacije na nastup. Štoviše, studenti prvog tipa (*performance approach*) crpe svoju orijentaciju iz samopoštovanja, željom da budu bolji od drugih ili iz percepcije medicine (struke) kao sredstva za poboljšanje njihovog društvenog položaja.

⁹Primjedba se odnosi na interpretaciju marketinških i poslovnih modela u obrazovanju koja zagovaraju tezu: "student = potrošač, obrazovanje = usluga".

¹⁰Neki ih nazivaju ego-ciljevima ili ciljevima sposobnosti.

¹¹U tom istraživanju analizirali su se razlozi za upisivanje studija medicine.

Orijentacija na *vještine* i *nastup* su tradicionalno shvaćeni kao suprotstavljeni ciljevi. Empirijska istraživanja daju drugačiju sliku; oni mogu biti negativno korelirani, nekorelirani ili pozitivno korelirani Pintrich (2000a). Modeli motivacije danas jasno razlikuju konstrukte kao: *ciljevi*, *uspješnost*, *interes* i *uvjerenje*, koji kao moderirajuće varijable služe u interpretaciji posljedica kao što su: *izbor*, *upornost* i *ponašanje*. Njihova međusobna interakcija u dinamici postignuća nema samo teorijsku vrijednost, već i praktičnu pedagošku, jer je **nedovoljno konstatirati da su studenti "motivirani" ili "nemotivirani"** u terminima nekog sadržaja, već na koji način bazični konstrukti moderiraju učenje i uspjeh. U kvalitativno drugačijem pristupu, ti bazični konstrukti nisu nezavisne varijable i konstituenti su latentnog faktora kojeg bi okarakterizirali kao motivacijskim.

6.5 Slabosti AGT teorije

Achievement Goal Theory (AGT) mogli bismo prevesti kao *Teorija postignuća* ili *Teorija postignuća kao ciljeva*. Takvi ciljevi su dohvatljivi i određuju kontekst djelovanja i ponašanja učenika u savladavanju gradiva. Raniji dvofaktorski model, a i kasniji trofaktorski model učenikove orijentacije, tvrde da orijentacija na vještine od strane učenika vodi ka boljim postignućima. Međutim, eksperimenti provedeni u posljednjih desetak godina nude zaključak da to i nije baš tako. Takvi ciljevi mogu imati pozitivan efekt, negativan efekt ili nemaju nikakav efekt na postignuća. Štoviše, postavlja se pitanje što ako su istodobno prisutne obje orijentacije, *orijentacija na vještine* i *orijentacija na nastup* i kakva je interakcija među njima, ako interakcija opće postoji.

Robustelli (2008) zastupa hipotezu o interakciji obje orijentacije, koju su već ranije pretpostavili Pintrich (2000c) i Barron and Harackiewicz (2001). Štoviše, ona postavlja *hipotezu usklađenosti* koja tvrdi da veća usklađenost učenikovih ciljeva s ciljevima nastavnika, pogoduje većem uspjehu na standardiziranim testovima i završnim ispitima.

U hijerarhijskom linearnom modelu kojeg razmatra Robustelli, indeksi i , j predstavljaju indeks studenta i indeks razreda, a varijable su:

- P_j – izvedba (eng. *performance*) kao cilj nastavnika
- M_j – vještina (eng. *mastery*) kao cilj nastavnika
- p_{ij} – izvedba kao cilj učenika
- m_{ij} – vještina kao cilj učenika
- y_{ij} – učenikov rezultat na testu
- α_j – sredina studentskih rezultata za j -ti razred.

Hijerarhijski linearni model ima dvije razine: razinu učenika (1. razina) i razina nastavnika (2. razina). U učeničkoj razini, p_{ij} predstavlja vrijednost boda i -tog učenika u j -tom razredu dobivenom na izvedbi. Nastavnička razina sadrži varijable P_j i M_j za $j \in \{1, \dots, 32\}$, tj. za 32 razreda koja su obuhvaćena testiranjem s ukupno 400 učenika. Na učeničkoj razini osnovni model izgleda

$$y_{ij} = \alpha_j + \pi_j p_{ij} + \mu_j m_{ij} + \varepsilon_{ij}$$

s *random efektima* nastavnikovih vještina i izvedbi

$$\pi_j = \pi_0 + \pi_P P_j + \pi_M M_j + \varepsilon_\pi,$$

$$\mu_j = \mu_0 + \mu_P P_j + \mu_M M_j + \varepsilon_\mu.$$

Vrijednosti y_{ij} su bodovi koje je student dobio na testu iz: standardiziranog testa engleskog jezika kojeg polažu svi učenici¹², a vrijednosti varijabli P i M dobivene su testiranjem nastavnika.

Analiza rezultata pokazuje da orijentacija na *vještina-kao-cilj* ima značajnu pozitivnu asocijaciju s učenikovim postignućima, a orijentacija na *nastup-kao-cilj* ima negativnu asocijaciju s učenikovim

¹²Zapravo se radilo o tri testa i analiza je provedena za svaki test.

postignućima. Međutim, nakon što je u analizu uvedena kontrolna varijabla (kovarijata) nazvana *prethodna postignuća* (učenika), gornje asocijacije potpuno nestaju. Drugim riječima, uključivanje *prethodnih postignuća* negira utjecaj ciljeva na učenikova postignuća. Nameće se zaključak da, ako se ciljevi unose u model onda su oni posljedica prethodnih postignuća, prije nego li uzrok budućih postignuća. Rezultati istraživanja time propituju smisao dosadašnjih 2-faktorskih i 3-faktorskih motivacijskih modela u kojima koncept ciljeva, kao uzroka učenikovih uspjeha, čini se da nije posve opravdan. U kontekstu navedenog eksperimenta pokazuje se upravo suprotno, ciljevi su posljedica prethodnog uspjeha. To ne mora značiti da su ciljevi nepotrebni, već da mogu biti korisni u službi održavanja motivacije i interesa za predmet podučavanja.

Prethodne zaključke treba uzeti s rezervom jer su se u eksperimentu pojavile izvjesne anomalije. Jedna od njih je neproporcionalno mnogo studenata koji su bili orijentirani na vještine i nema dovoljno argumenata za podršku interakcijskom ciljnom modelu niti aditivnom ciljnom modelu. Sam autor iznosi još neke ograde na proceduru sakupljanja podataka koji je varirao od škole do škole. Što se nastavnika tiče, njih 29 od 32 je bilo uključeno u hijerarhijski linearni model jer nisu svi bili testirani zbog malog broja učenika u njihovom razredu. Osim toga, broj razreda također nije reprezentativan za ozbiljniju studiju. Još jedan nedostatak je broj škola koje su bile uključene u analizu, samo dvije, tako da se zaključci ne mogu interpretirati kao generalni. Naposljetku, učenici u drugačijem školskom okruženju mogu usvojiti iste ciljeve ali s različitim efektima na učenikova postignuća. Ima još jedan element, koji se proteže kroz mnoga slična istraživanja, a taj je da sama postignuća nisu i ne moraju biti samo posljedica školskog sustava (op. autora).

Hipoteza usklađenosti učiteljevih i učenikovih ciljeva nije potvrđena. Ima nekoliko razloga za to. Jedan je taj što testovi kojima su bili podvrgnuti nastavnici pokazuju slabu pouzdanost, a drugi je taj što nema jasnog pokazatelja je li studenti korektno percipiraju učiteljevu ciljnu orijentaciju. Samoiskazi (eng. *self-report*), korišteni u ovoj studiji, su često iskazani u socijalno prihvatljivom stilu od strane ispitanika, tj. osoba često odgovara na način da bude podobna socijalnom okruženju u kojem se nalazi, što stvara poteškoću u pripisivanju rezultata stvarnoj ciljnoj orijentaciji osobe ili njenoj želji da izgleda podobna.

Ova analiza postavlja mnoga pitanja i nudi sugestije za daljnja istraživanja. Osim povećanja uzorka i odabira pouzdanijih testova prirodno se nameće i zahtjev za longitudinalnom studijom koja bi mogla odgovoriti na pitanja jesu li nastavnikove ciljne orijentacije stabilne, a učenikove možda i nisu. Znači li to da su učenici slabo inficirani učiteljevim ciljevima i više inficirani vlastitim učincima? Sve skupa nas vraća na sam početak i propituje jesu li ciljevi tu sami po sebi, nameće li ih netko drugi ili su rezultat prošlih postignuća?

Pintrich (2000b) već ranije pretpostavlja da su ciljevi, kao kognitivne reprezentacije strukture znanja, osjetljivi i na kontekst i na osobne karakteristike. Kao i svaka druga struktura znanja, ciljevi mogu biti aktivirani apriori od strane pojedinca u trenutku kad im kontekst ponudi potrebnu informaciju za njihovo aktiviranje. Termin *ciljna orijentacija* se koristi kako bi naznačili da je aktiviran sustav vjerovanja o vlastitoj kompetenciji, uspjehu, sposobnostima i greškama u nekoj situaciji, ali to ne znači da tom pojedincu nije dostupan i neki drugi sustav vjerovanja. Pojedinaac može izražavati različite strukture znanja i vjerovanja u različitim situacijama. Na primjer, korektno i znanstveno objašnjenje topline i temperature na satu fizike može biti u suprotnosti s naivnim poimanjem istog u svakodnevnom životu. Takva osjetljivost ciljeva otvara nova pitanja o tome "kako" i "kada" se ciljevi aktiviraju, a otvara i novu mogućnost mjerenja utjecaja ciljeva putem samodualne hijerarhije (odjeljak ??), kao i novih načina kako se ciljevi kognitivno reprezentiraju.

6.6 Primarnost interesa

Ovaj odjeljak direktno se nastavlja na odjeljak 5.2 na str. 13 u kojem se govori o ciljevima u poslovnom okruženju. Posljedica postavljanja ciljeva u poslovnom okruženju je povećanje mentalne krutosti i smanjivanje autonomije zaposlenika. Činjenica je da se nastavnici susreću sa sve učestalijom demotiviranošću za učenjem kod učenika i da se rasprave o tome najviše bave posljedicama, a to su metode kako 'izliječiti' nemotiviranost.

Nastavnik je taj koji treba biti u stanju prepoznati je li učenik motiviran ili ne¹³, jer odsustvo motivacije za bilo što predstavlja dublji poremećaj kojeg osoba sama, bez pomoći okoline i bliskih osoba teško može prebroditi. Kako prepoznati motivaciju ili gubitak motivacije?

- Radoznalost. Djeca su po prirodi radoznala i radoznalost treba njegovati jer potiče i osobnu motivaciju učenika i pozitivno djeluje na njegovu okolinu.
- Samostalnost, kompetentnost, ustrajnost doprinose i osobnoj vrijednosti učenika.
- Strah od neuspjeha i strah općenito je socijalno prihvatljiv, ali može predstavljati poteškoću u učenju. Osoba u strahu orijentirana je na budućnost i kaznu kao posljedicu sadašnjeg djelovanja. Strah je dokaz nedostatka poticaja i, dugoročno gledano, umanjuje šanse za uspjehom.
- Utjecaj porodice je presudan na motivaciju djeteta. Za razliku od poticajnih roditelja postoje i toksični roditelji koji kontaminiraju djetetov osjećaj sebstva i ignoriraju njegove emocionalne potrebe.

Što može učiniti nastavnik ako ima u razredu nemotivirane učenike? Ništa, ako i sam nije motiviran za svoj posao i ako nije dobro obučen i fokusiran na potrebe djece. Dobar učitelj je inspirativan, zanimljiv, poticajan i otvoren za komunikaciju, a metode podučavanja djetetu trebaju ponuditi alate i znanje koje može odmah ili u budućnosti prilagoditi svojim potrebama. Školsko ozračje treba djetetu ponuditi sigurnost i zaštitu. Više i detaljnije govori se u preglednom članku Williams and Williams (2011), a D'Souza and Maheshwari (2010) zaključuju da se motivacija potiče neprestanim motivacijskim iskustvima na dnevnoj bazi.

A što ako motivacijska iskustva u školi nedostaju i ako nastavnik nije spreman uhvatiti se u koštac s izazovima svog posla? Takav nastavnik pokazuje nezainteresiranost za vlastiti posao i kao takav nije u stanju pobuditi interes za učenjem kod djece. Robustelli (2008) navodi na nedostatak istraživanja u kojima se propituje prihvaćanje ciljeva nastavnika od strane učenika.

Djeca su danas dodatno demotivirana jer ne vide direktnu korist od naučenog u svakodnevnom životu, a za to su primarno krivi kurikuli koji nude nezanimljive sadržaje. Ocjena kao motivacijski faktor (cilj) je dio priče koja pogrešno shvaća obrazovni proces. *Orijentacija na nastup (performance)*, kako smo nazvali takvu vrstu učenikovog ponašanja, usmjerena je prema izvanjskim nametnutim obrascima ponašanja, a ne istinskim vrijednostima. Sve rasprave o tome kako ocjena potiče motiviranost učenika su na krivim temeljima je ne uvažavaju činjenicu da **sve više današnje djece nije zainteresirano za obrazovanje**. Ona ne doživljavaju obrazovanje kao dio životnog procesa jer od njega ne vide koristi.

Interes kao dinamična varijabla u poticanju uspješnog učenja sve je više predmet istraživanja, ali u okruženju (modelu) u kojem svi elementi, bazični konstrukti i ponašanje čine mrežu s međusobnim utjecajem i u kojoj se nastoje prepoznati uzorci koji utječu na spoznavanje i ponašanje. Istraživanja tog tipa su još u povojima, a ciljevi u takvim modelima, kao kognitivne reprezentacije strukture znanja, su kontekstualno osjetljivi i ovisni o unutrašnjim osobnim faktorima. Na primjer, student u kompetitivnom okruženju može biti orijentiran na nastup, a u individualnom e-učenju može biti orijentiran na vještinu pri usvajanju znanja.

Eckblad (1981) sugerira da *interes*, kao afektivno stanje, igra centralnu ulogu u odnosu između znanja i sustava vrijednosti pojedinca i njegove okoline. U ranijoj fazi proučavanja interes je razmatran kao okidač podržavan od okoline (drugi pojedinci, zadaci...), a u kasnijoj fazi se shvaća kao samoregulirajući proces u kojem pojedinac započinje angažman i traži odgovore na pitanja iz čiste radoznalosti. Interes nije nikad sasvim izvanjski niti je čista unutarnja potreba subjekta; on je odraz onoga što pojedinac doprinosi zadatku, onoga što podržava okolina i načina na koji je pojedinac sposoban djelovati u svom okruženju

Edelson and Joseph (2004) opisuju *Interest-Driven Learning (IDL)* okruženje za dizajniranje aktivnosti koje koriste moć interesa u motivaciji za učenjem. Buđenje interesa i njegovo širenje s

¹³ako to već nisu učinili roditelji.

jednog sadržaja (fokusa) na drugi je osnovna tehnika poticanja za učenjem jer polazi od već postojećeg interesa kojeg djeca intrinzično posjeduju ili ga trenutačno probude. Okidač za buđenje interesa može biti priča ili pokus čiji su rezultati u suprotnosti s njihovim sadašnjim iskustvom.

IDL okruženje razlikuje interes koji proizlazi iz sadržaja aktivnosti i od kontekstualnih motivatora koji proizlaze iz okolnosti u kojima se aktivnost dešava. Za uspješno učenje potreban je balans jednog i drugog. Na primjer, ako je student zainteresiran za takmičenje u raspravljanju može puno energije potrošiti na dodvoravanje sucima što ga odvlači od sadržaja rasprave.

Poteškoće u kreiranju IDL okruženja za učenje leže u činjenici da intenzitet učenikovog interesa varira u vremenu kao i u činjenici da fokus njegovog interesa ne pokriva cjelokupno predviđeno gradivo. Edelson and Joseph razvijaju strategije za savladavanje tih poteškoća koje počivaju na: *zadovoljstvu, sadržaju interesa, potvrdi identiteta, životnim ciljevima i zadovoljavanju znatiželje*. Osim toga, ljudi abstraktno vrednuju aktivnosti i rezultate koji nemaju samo direktan utjecaj na njih. Na primjer, briga za okolinu može biti dio njihovih duhovnih uvjerenja o povezanosti svega što nas okružuje, a ne samo odgovornost za buduća pokoljenja. Ta briga može potaknuti interes i aktivnosti za mnoge nastavne i nenastavne sadržaje u njegovom životu koje doprinose očuvanju prirode.

Kako iskoristiti spomenute izvore interesa u dizajniranju IDL okruženja? Životni ciljevi su neprestani izvori interesa. Djeca će pokazivati interes za čitanjem i pisanjem jer time proširuju svoju autonomiju. Pojedince svih uzrasta će zanimati funkcioniranje društvenog sustava u kojem žive jer im razumijevanje dinamike udruživanja i hijerarhije u školi i na radnom mjestu pomaže u stvaranju osjećaja pripadnosti i vlastitog identiteta.

Znatiželja je posljedica raskoraka između učenikovog razumijevanja stvari i trenutačnih observacija. Kreator IDL okruženja treba taj raskorak učiniti dovoljno velikim ali na način da nova situacija ostane u dohvat učenikovih kognitivnih sposobnosti. Tipično buđenje znatiželje je takvo da učitelj pokaže neki eksperiment i vodi ih na putu do objašnjenja te pojave. Druga strategija je da učenici artikuliraju svoje sadašnje koncepcije o zadanoj temi kao uvod u širu raspravu o toj nastavnoj jedinici. Pri tome učenici identificiraju nekonzistentnosti u svom sadašnjem znanju što pobuđuje znatiželju i želju za povezivanjem starog znanja i novog sadržaja. Edelson and Joseph opisuju jednu takvu nastavnu jedinicu o klimi u kojoj su učenici bili zamoljeni da bojama izraze raspored temperature zraka u svijetu (kontinentu, državi) u određenom mjesecu. To je izazvalo živi raspravu o poimanju topline i temperaturnim procesima.

Održavanje i poticanje interesa postiže se kontekstualnim motivatorima. Na primjer, interes za razumijevanjem kosog hitca može se potaknuti pokretanjem takmičenja u konstrukciji katapulte. Ovdje su prisutna čak dva kontekstualna motivatora: takmičenje kao socijalni motivator i izazov do konačnog razumijevanja trajektorije gibanja.

U kreiranju IDL okruženja investicija učenika u rješavanje problema može biti značajni održavatelj motivacije, ali je tu potrebno stručno vođenje od strane kreatora procesa kako ne bi došlo do gubitka interesa. Pri tome učenik mora imati osjećaj da je produktivan, a nova investicija treba biti nadovezana na staru. Osobno sam bio zadivljen rješenjima koja sam dobivao od studenata kad sam im dao zadatak da sami kreiraju grafički prikaz statističkih podataka, s napomenom da pokušaju kreirati dinamički prikaz koji pokazuje dodatne podatke (pop up) ovisno o položaju miša, ali da ne smatraju neuspjehom ako to ne uspiju.

U već spomenutom članku Edelson and Joseph detaljno opisuju dizajn interesno poticanog kurikula kojeg su nazvali *Passion Curriculum* čija realizacija počiva na IDL okruženju. Takvo okruženje proširuje i rasterećuje tradicionalni pogled na motivaciju i učenje kao instrumente za postizanje postavljenih obrazovnih ciljeva.

7 Agregacija ocjena

Pod agregacijom ocjena podrazumijeva se određivanje ocjene učeniku na temelju njegovog profila i profila svih učenika u njegovom okruženju/razredu. U takvu proceduru spada i zaključna ocjena na kraju polugodišta ili na kraju školske godine ili agregacija ocjena od nekoliko evaluatora.

Pod profilom se smatra skup ocjena učenika/ispitanika na temelju raznih kriterija ili nastavnih predmeta. U tablici 1 dana su tri profila od tri učenika A, B, C, a I, II, III mogu biti predmeti, uzastopni testovi ili nešto treće.

Tablica 1: Skup učeničkih profila u formi tablice.

	Predmet		
	I	II	III
A	63	59	66
B	66	57	69
C	69	56	68

Najčešće korištena metoda agregiranja je *aritmetička sredina* bodova ili, što je ekvivalentno, *zbroj bodova*. Ako gledamo rang-listu učenika iz tablice 1 dobivenu tom metodom onda je to: $A \geq B \geq C$ ¹⁴.

Mnogi nastavnici zagovaraju medijan kao agregiranu ocjenu bez dublje argumentacije. Ovdje ćemo prodiskutirati mane jedne i druge procedure.

One metode koje zadovoljavaju aksiom nezavisnosti o irelevantnoj alternativi, koriste samo učenikov profil, dok će neke druge metode koristiti i profile ostalih učenika.

U ovom odjeljku razmatrati ćemo nekoliko načina kako dobiti agregiranu (prosječnu) ocjenu učenika na temelju više kriterija, uključujući i metodu potencijala razvijenu od strane autora ovog teksta.

7.1 Subjektivna procjena

Najjednostavnija moguća situacija ilustrirana je primjerom u kojem nastavnik subjektivno uspoređuje učenike po nekom kriteriju, koji može biti i "opći dojam", na način da za dva učenika odredi koji je *bolji* ili ih proglašava *jednako dobrima* ako je neodlučan.

Primjer. Ivica je bolji od Marice, Marica je bolja od Jožeka, Petar i Jožek su jednako dobri, a Petar je bolji od Ivice. Ostali parovi nisu uspoređeni. Bez obzira kako definirali relaciju 'bolji' ovo je dovoljan broj informacija da se ti učenici rangiraju. Primijetimo da relacija nije tranzitivna jer sadrži ciklus:

$$\text{Ivica} > \text{Marica} > \text{Jožek} \sim \text{Petar} > \text{Ivica}.$$

U takvoj situaciji, kad postoji ciklus, većina metoda zakazuje i ne zna kako procesirati ovakve ulazne podatke. Metoda potencijala nudi odgovor:

$$\text{Ivica (0.42)} > \text{Marica (0.25)} > \text{Petar (0.18)} > \text{Jožek (0.15)},$$

gdje brojevi u zagradi predstavljaju dobivenu težinu.

Situacije gornjeg tipa su veoma česte, ali ih izbjegavamo upravo zbog nedostatka tranzitivnosti koji se nameće kao uvjet-bez-kojeg-se-ne-može. Broj učenika u razredu sigurno je veći od 4, ali isti princip je i dalje primjenjiv. Jedini uvjet na relaciju preferencije je da njen graf bude povezan.

7.2 O skalama

Čitateljima je, pretpostavljam, već poznat pojam skale, ako ne drugačije onda kao skup rezultata mjerenja. Ovdje ćemo detaljno pojasniti taj pojam i vidjeti po čemu se one razlikuju. Najčešće korištene skale su: *nominalna*, *ordinalna*, *omjerna* i *intervalna*. Takvu je tipologiju¹⁵ skala uveo S. S. Stevens (1946), a karakterizirane su dozvoljenom grupom transformacija: identitet, monotona,

¹⁴Relaciju \geq čitaj: 'bolji ili jednako dobar'

¹⁵iako postoje i drugačije tipologije

sličnost i afina transformacija. U nazivu skale već je prisutna informacija o prirodi izmjerene vrijednosti, a prava pozadina takve tipologije leži u stupnju jedinstvenosti numeričke reprezentacije mjerenog entiteta.

Nominalna ili kvalitativna skala je skala čije vrijednosti su oznake (nazivi), i nema drugih dozvoljenih transformacija osim identitete. Primjer takve skale je sama Stivensova tipologija. Ona služi, kako i samo njeno ime govori za imenovanje objekata.

Ordinalna skala je invarijantna na rastuće transformacije: $w = \phi(u)$, gdje je $\phi : \mathbb{R} \mapsto \mathbb{R}$ strogo rastuća funkcija. Iako su ordinalne skale sveprisutne u socijalnim znanostima, izbjegavaju se u fizici i tehnici jer se smatra da predstavljaju vrlo slabu formu mjerenja. Odgovori na anketna pitanja koja rangiraju stupanj zadovoljstva nekom uslugom su na ordinalnoj skali.

Drugi tip skale je *omjerna* skala. To je skala koja je invarijantna na množenje pozitivnim brojem: $w = \alpha \cdot u$, $\alpha > 0$. Ako je izmjerena vrijednost dobivena uspoređivanjem s mjernom jedinicom onda je ona izražena na omjernoj skali (masa, naboj, duljina štapa). Omjerna skala ima definirano ishodište. Tipičan primjer omjerne skale je izražavanje brzine gibanja nekog tijela ili odgovori na anketna pitanja tip: "Koliko vremena dnevno provodite u šetnji?" (A) 0.5–1 sat, (B) 1–2 sata, (C) 2–3 sata. Zaključci provedeni na omjernoj skali putem statističkih analiza su korektni jer su invarijantni na množenje pozitivnim brojem.

Intervalna skala dopušta pozitivne afine transformacije ili, izraženo formulom: ako je u varijabla koju mjerimo onda je varijabla

$$w = \alpha \cdot u + \beta, \quad \alpha > 0, \quad \beta \in \mathbb{R}$$

jednako korisna (interpretabilna) u našem modelu kao i u . Omjer vrijednosti na intervalnoj skali je besmislen, ali je zato kvocijent duljina intervala između dviju vrijednosti smislen i poprima vrijednosti na omjernoj skali. Kako interpretirati podatke na intervalnoj skali? Na primjer, ako je temperatura zraka 15° na nekoj skali to ne znači da je vani toplo ili hladno jer toplinu doživljavamo u odnosu na temperaturu tijela.

Primjer intervalne skale je temperatura izražena u $^\circ C$ i F (Fahrenheit). Konverzija jedne skale u drugu je dana formulom $F = 2 \cdot C + 30$. Temperaturna skala nosi u sebi još jedno svojstvo, a to je kvalitativna smislenost razlike temperature. Drugi ilustrativan primjer je *korisnost* u teoriji očekivane korisnosti. Korisnost je također invarijantna na pozitivne transformacije, ali razlika korisnosti nema substancijalnog smisla jer ta razlika nije interpretabilna u modelu. Uz intervalnu skalu veže se i *log-intervalna* skala koja je okarakterizirana dozvoljenim transformacijama oblika $x \mapsto sx^r$, $s, r \in \mathbb{R}$. Logaritam te transformacije je afina transformacija.

Osim spomenutih skala postoje još neke specijalne skale koje se koriste u kartografiji, a pojedini autori predlažu i drugačije tipologije mjernih skala u koje nećemo ulaziti.

Zašto je važno na kakvoj skali vršimo mjerenja? S izmjerenim vrijednostima manipuliramo i na temelju tih manipulacija donosimo određene zaključke. Zaključak ne bi smio varirati ako vršimo dozvoljene transformacije skale.

7.3 Centralna mjera skale.

Centralna mjera skale predstavlja vrlo rudimentarnu agregaciju u kojoj se izmjereni podaci reprezentiraju jednim jedinim brojem. Primjer centralne mjere je: *aritmetička sredina* za intervalnu skalu i *median* za ordinalnu skalu.

Ako je $x = (x_i)$, $i = 1, \dots, n$ vektor izmjerenih vrijednosti onda se aritmetička sredina $\mu(x)$ računa po formuli

$$\mu(x) = \frac{1}{n} \sum_{i=1}^n x_i$$

i predstavlja točku minimuma kvadratičnog odstupanja (devijance), tj.

$$\mu(x) = \arg \min_x \sum_{i=1}^n (x - x_i)^2.$$

Medijan je 2-kvantil izmjerenih podataka i predstavlja točku minimuma apsolutnog odstupanja¹⁶, tj.

$$\text{med}(x) = \arg \min_x \sum_{i=1}^n |x - x_i|.$$

Koju centralnu mjeru uzimamo za naše podatke ovisi o skali. Jedan od zahtjeva na centralnu mjeru je invarijantnost na transformaciju koja definira skalu, a drugi zahtjev treba odražavati usklađenost centralne mjere s algebarskom strukturom koju skala (objekti) posjeduje. U skalu također treba ugraditi i zahtjev invarijantnosti centralne mjere na labeliranje objekata (nazivi) što se formulira kao invarijantnost centralne mjere na permutiranje izmjerenih vrijednosti.

Aritmetička sredina. Centralna mjera $f(x)$ podataka na intervalnoj skali treba zadovoljavati sljedeće zahtjeve:

1. $f(x + \beta) = f(x) + \beta$ – invarijantnost na dodavanje konstante,
2. $f(\alpha x) = \alpha f(x)$ – invarijantnost na množenje brojem,
3. $f(x) = f(x_\sigma)$ za svaku permutaciju σ – invarijantnost na permutiranje objekata,
4. $f(x + y) = f(x) + f(y)$ – invarijantnost na algebarsku operaciju zbrajanja.

Zahtjevi i) i ii) su ekvivalentni zahtjevu invarijantnosti na dozvoljene transformacije skale. Oznaka $x + \beta$ za sumu vektora i konstante znači dodavanje konstante svakoj komponenti vektora, $x + y$ je zbrajanje vektora po komponentama, a x_σ je vektor nastao iz x permutiranjem njegovih komponenti tj. $(x_\sigma)_i := x_{\sigma(i)}$, $\forall i = 1, \dots, n$.

Lemma 7.1. *Centralna mjera podataka na intervalnoj skali je aritmetička sredina $\mu(x)$.*

Dokaz. Dokaz ćemo provesti u 2 koraka.

1. $\sum_i x_i = 0 \implies f(x) = 0$. Zaista,

$$\begin{aligned} f(x) &\stackrel{\text{iv)}}{=} \sum_{i=1}^n f(0, \dots, 0, x_i, 0, \dots, 0) \\ &\stackrel{\text{ii)}}{=} \sum_{i=1}^n x_i f(0, \dots, 0, 1, 0, \dots, 0) \\ &\stackrel{\text{iii)}}{=} \left(\sum_{i=1}^n x_i \right) f(1, 0, \dots, 0) \\ &= 0. \end{aligned}$$

2. $f(x, 0, \dots, 0) = \frac{x}{n}$. Zaista,

$$\begin{aligned} f(x, 0, \dots, 0) &\stackrel{\text{i)}}{=} f\left(\frac{(n-1)x}{n}, -\frac{x}{n}, \dots, -\frac{x}{n}\right) + \frac{x}{n} \\ &= \frac{x}{n} \quad (\text{zbog 1. koraka.}) \end{aligned}$$

¹⁶Gledano na taj način, točka minimuma apsolutnog odstupanja je jednoznačna ako je broj podataka neparan ili srednji interval sortiranih podataka ako je broj podataka paran. U drugom slučaju neki uzimaju za medijan lijevi rub unutarnjeg intervala, desni rub ili sredinu intervala.

Sada je

$$f(x) = \left(\sum_{i=1}^n x_i \right) f(1, 0, \dots, 0) = \frac{1}{n} \sum_{i=1}^n x_i = \mu(x). \quad \square$$

Medijan.

Lemma 7.2. *Centralna mjera podataka na ordinalnoj skali je medijan.*

Dokaz. Neka je $\phi(\cdot)$ strogo rastuća transformacija. Označimo $m = \text{med}(x)$. Tvrdimo $\phi(m) = \text{med}(\phi(x))$. Zbog monotonosti transformacije ϕ :

$$x_i \leq m \leq x_k \iff \phi(x_i) \leq \phi(m) \leq \phi(x_k)$$

Ako je $[x_i, x_k]$ srednji interval sortiranih podataka, onda je $[\phi(x_i), \phi(x_k)]$ srednji interval transformiranih sortiranih podataka, odakle slijedi $\text{med}(\phi(x)) = \phi(m)$. \square

Geometrijska sredina. Jednostavnosti radi promatramo situaciju u kojoj imamo samo dva mjerenja: x i y . Sljedeća svojstva očekujemo od centralne mjere f na omjernoj skali:

1. $f(\alpha x, \alpha y) = \alpha f(x, y)$, $\alpha > 0$ – invarijantnost na promjenu mjerne jedinice,
2. $f(1/x, 1/y) = 1/f(x, y)$ – idempotentnost invertiranja,

Lemma 7.3. *Centralna mjera podataka na omjernoj skali je njihova geometrijska sredina.*

Dokaz. Dokaz ćemo provesti za $n = 2$, gdje je n broj podataka, a u općem slučaju dokaz ide indukcijom po broju podataka. Prvo ćemo dokazati da je $f(1, x) = \sqrt{x}$. Zaista,

$$f(1, x) \stackrel{\text{ii)}}{=} \frac{1}{f(1, \frac{1}{x})} = \frac{x}{x \cdot f(1, \frac{1}{x})} \stackrel{\text{i)}}{=} \frac{x}{f(x, 1)} \implies f(1, x) = \sqrt{x}.$$

Sada je

$$f(x, y) \stackrel{\text{i)}}{=} x f\left(1, \frac{y}{x}\right) = x \sqrt{\frac{y}{x}} = \sqrt{xy}. \quad \square$$

7.4 Funkcija vrijednosti

U analizi tablice učeničkih profila manipuliramo s ulaznim podacima (brojevima) na način koji može biti manje ili više smislen, ovisno o tome održavaju li ti brojevi kvalitativne ili kvantitativne odnose među entitetima. Kvalitativne relacije su na primjer:

- i) ovaj je put *duži* nego onaj, ili
- ii) *više mi sviđa* slika a nego slika b .

Bez obzira kako nazivali taj kvalitativan odnos među objektima (duljina, sviđanje) mi ćemo ga nazvati generičkim nazivom *preferencija* u oznaci \succsim . Prirodno je zahtijevati da relacija preferencije bude tranzitivna, tj.

$$a \succsim b \ \& \ b \succsim c \implies a \succsim c.$$

Jednaku preferiranost među objektima nazivamo *ekvivalencijom* i označavamo

$$a \sim b \iff a \succsim b \ \& \ b \succsim a.$$

Kvantitativni modeli izražavaju kvalitativni odnos nekim brojem na način da se poštuje preferencija. Funkciju V koja to čini nazivamo *funkcijom vrijednosti* i ona zadovoljava

$$(\forall a, b) a \succcurlyeq b \iff V(a) \geq V(b).$$

Na primjeru puteva, funkcija vrijednosti je *duljina puta mjerena u kilometrima (ili satima)*, na primjeru slika, obično nismo u stanju slike postaviti na neku skalu. Sofisticiraniji modeli u tom slučaju pridružuju svakom paru objekata *intenzitet preferencije* i na temelju takvih podataka konstruiraju skalu na kojoj se izražavaju vrijednosti objekata (v. str. 53 pod nazivom [Metoda potencijala](#)).

7.5 Utjecaj mjerenja kod agregacije ocjena

Kao metodu agregacije ocjena uzmimo *sumu*. Pretpostavimo da je ocjenjivač bodovao testove tri ispitanika A, B, C kao u tablici 2 (najveći mogući broj bodova po predmetu je 100). Kvantitativni

Tablica 2: Bodovanje ispitanika.

	Predmet			zbroj
	I	II	III	
A	63	59	66	188
B	66	57	69	192
C	69	56	68	193

odnos između 66 i 63 pokazuje da je kandidat B postigao bolji uspjeh od kandidata A na testu ako se gleda samo predmet I. U posljednjem stupcu tabele dana je suma bodova svakog kandidata koja predstavlja *opći uspjeh* kandidata za sva tri predmeta kao sumu bodova.

Pretpostavimo da je neki drugi ispitivač također ocijenio iste testove i rezultat njegovog ocjenjivanja je skup profila predstavljen u tablici 3. Numeričke vrijednosti pridijeljene uspjehu ispitanika

Tablica 3: Drugo moguće bodovanje ispitanika.

	Predmet			zbroj
	I	II	III	
A	59	65	46	170
B	60	61	48	169
C	61	58	47	166

za svaki predmet predstavljaju isti kvalitativni poredak za oba ocjenjivača. Drugi ispitivač je bio nešto blaži u ocjenjivanju za drugi predmet i stroži za preostala dva. Suma bodova međutim daje inverzni poredak za opći uspjeh: C-B-A za prvog ispitivača i A-B-C za drugog ispitivača iako je kvalitativni odnos (poredak) među učenicima, kod oba ispitivača, za svaki ispit isti. Aritmetička sredina ne uvažava 'kvalitetu' kod agregacije i ako netko želi uvažiti kvalitativni odnos među učenicima u agregiranoj rang-listi, onda suma bodova nije dobra metoda.

Ova jednostavna analiza upućuje na općenitiji pristup agregaciji u kojem se daje lista zahtjeva (aksioma) koje bi metoda trebala zadovoljavati. Pri tome je moguće pretjerati sa zahtjevima, tj. da lista zahtjeva bude prestroga i da ne postoji metoda koja ih zadovoljava. Teorija odlučivanja prošlog stoljeća puna je takvih teorema nemogućnosti koji dokazuju nepostojanje metode za razne sustave aksioma. Posebno je važna i tražena takva situacija kad neki sustav aksioma jedinstveno određuje metodu. Jedan od primjera je *teorija očekivane korisnosti* koju ekonomska literatura i dan danas obilato koristi. Od teorema ne egzistencije spomenuo bih samo Arowljev teorem u teoriji izbora¹⁷

¹⁷Arrow, Kenneth J. (1951), *Social Choice and Individual Values*, John Wiley and Sons, New York.

koji je 'razočarao' zagovornike zapadne demokracije jet tvrdi da ne postoji metoda koja zadovoljava 'na oko jednostavne demokratske zahtjeve'.

7.6 Kvalitativna i kvantitativna smislenost agregacije

Definicija 7.4. Dvije skale su *ordinalno ekvivalentne* ako je jedna skala kompozicija druge sa strogo rastućom funkcijom i svaku od njih nazivamo *ordinalnom skalom*. Strogo rastuća transformacija naziva se *dopuštena transformacija* ordinalne skale.

Teorem 7.5. *Nejednakost između aritmetičkih sredina nije invarijantna na strogo rastuće transformacije. Drugim riječima, ako je V ordinalna funkcija onda*

$$\sum_{i=1}^n \lambda_i V(a_i) \geq \sum_{i=1}^n \mu_i V(a_i)$$

ne povlači da za svaku strogo rastuću funkciju h vrijedi

$$\sum_{i=1}^n \lambda_i h(V(a_i)) \geq \sum_{i=1}^n \mu_i h(V(a_i)),$$

gdje je $\lambda_i, \mu_i \geq 0$ i $\sum_{i=1}^n \mu_i = 1, \sum_{i=1}^n \lambda_i = 1$.

Dokaz je kratak i ostavljamo ga čitatelju.

Teorem tvrdi da aritmetička sredina kao kvantitativni podatak nije smisljena na ordinalnoj skali. U literaturi se koristi termin da ona nije *kvantitativno smisljena*.

Ako broj bodova za svaki predmet, u tablicama 2 i 3 shvatimo kao odraz *poretka* kandidata na ispitu, onda je računanje prosjeka očigledno *kvantitativno besmisljena* operacija jer su oba ocjenjivača dala isti poredak kandidata za svaki ispit (v. tablicu 4), a agregirani rangovi su različiti.

Tablica 4: Kvalitativni odnos učenika po predmetima. Rangovi.

	Predmet		
	I	II	III
A	1	3	1
B	2	2	3
C	3	1	2

Kako bismo interpretirali tu tvrdnju u kontekstu donošenja ocjena? Ako bismo dozvolili nastavnicima da svaki ima svoju skalu ocjena, na primjer jedan koristi skalu 1-2-3-4-5-6, a drugi koristi skalu 10-12-15-17-21-28, onda bi se moglo desiti da aritmetička sredina, kao metoda agregacije, u oba slučaja ne daje isti poredak kandidata. Jedan od načina rješavanja te nelagode je natjerati sve nastavnike da koriste istu skalu i da prođu ozbiljan trening u donošenju ocjena. Primjer takvog okruženja je ocjenjivanje figura na umjetničkom klizanju na primjer, a glavni razlog za takva nastojanja u teoriji izbora je onemogućavanje manipulacije¹⁸.

7.7 Agregacija ordinalnih vrijednosti

Ocjene u hrvatskom obrazovnom sustavu poprimaju vrijednosti u uređenom skupu kategorija, a to što su kategorije označene brojevima u skupu $\{1, 2, 3, 4, 5\}$ nosi dodatne poteškoće. Stoga je opravdano pitanje kako agregirati ocjene ako aritmetička sredina nije kvantitativno smisljena. Ovdje ćemo analizirati neke od metoda agregacije ordinalnih podataka.

¹⁸Pojam *manipulacija* je višeznačan i kontekstualno ovisan. U svakom konkretnom slučaju (metodi) treba eksplicitno reći o kakvoj manipulaciji se radi.

7.7.1 Bordino pravilo

Tablica 4 predstavlja rang-listu kandidata po predmetima s tom razlikom što manji broj predstavlja lošiju poziciju na rang listi. Takva tablica predstavlja najsiroviju informaciju o kandidatima. Pitanje je možemo li ju iskoristiti kao takvu i donijeti negu agregiranu rang-listu na temelju tih podataka, a dodatno je pitanje koliko je ta rang-lista zadovoljavajuća.

Jedan način agregacije ordinalnih lista dao je Jean-Charles de Borda još davne 1770. godine. On je kritizirao način izbora u Francusku akademiju i predložio metodu, danas poznatu kao Bordino pravilo ili Bordin indeks.

Pravilo je jednostavno. Ako je u pitanju n kandidata onda se svakoj rang-listi prvo pridruži stupac numeričkih vrijednosti u skupu $\{1, 2, \dots, n\}$ gdje se prvom mjestu na listi daje n bodova, drugom $n - 1$ i tako dalje kao što je urađeno u tablici 4. Tada se

svakom kandidatu pridijeli zbroj njegovih bodova iz svake liste.

Originalno Bordino pravilo, umjesto tablice oblika 4, koristi tablicu u kojoj su svi bodovi umanjeni za jedan. Jednostavno je pokazati da je krajnji rezultat, tj. rang-lista, u oba slučaja isti.

Tablica 5: Bordin indeks. Rang-lista učenika iz tablice 4.

	Borda
A	4
B	7
C	6

Na dodatno pitanje koliko je rezultat zadovoljavajući (prihvatljiv) teško je reći bez dodatnih kriterija što ovdje znači 'prihvatljivost'. Navodimo neke od tih kriterija koji se najčešće spominju u literaturi

Nezavisnost o irelevantnoj alternativni. U najkraćim crtama taj kriterij karakterizira agregativne procese u kojima dodavanje nove opcije (alternative) u račun ne mijenja poredak starih opcija.

Bordin indeks **ne zadovoljava nezavisnost o irelevantnoj alternativni**.

Pravilo većine. Kandidat na izborima koji ima više od 50% glasova je pobjednik.

Bordin indeks može **narušiti** to pravilo.

Monotonost. Ako izborne rezultate promijenimo na način da pobjednika još više favoriziramo on će i nakon toga biti pobjednik.

Bordin indeks **zadovoljava** monotonost.

Condorcetov kriterij. Ako postoji kandidat koji nadglasava svakog pojedinog protivnika (po broju listića) tada je on pobjednik izbora (tzv. Condorcetov pobjednik).

Bordin indeks može **narušiti** to Condorcetov kriterij.

Bordino pravilo se može jednoznačno okarakterizirati grupom kriterija koje nećemo ovdje navoditi. Štoviše, koliko je autoru ovog teksta poznato, postoje bar dvije grupe kriterija koje jednoznačno određuju Bordino pravilo.

7.8 Većinska procjena. Medijan.

Balinski i Laraki¹⁹ suzili su ocjenjivačku skalu na opće prihvaćenu ordinalnu verbalnu skalu: *odličan (5), vrlo dobar (4), dobar (3), prihvatljiv (2), slab (1), nedovoljan (0)*, nadajući se da će time ublažiti

¹⁹u svojoj knjizi *Election by Majority Judgement: Experimental Evidence*, Technical report. Laboratoire d'Econometrie 1, rue Descartes F-75005, Paris (2007)

moćnost manipulacije. Za operator agregacije predlažu medijan profila jer su ocjene donesene na ordinalnoj skali. Ako je broj ocjena paran, onda se kao medijan uzima **manja vrijednost** od dvije srednje vrijednosti.

Jednostavnosti radi, u primjerima koji slijede umjesto verbalnih ocjena pisat ćemo brojeve, imajući u vidu da se radi o ordinalnoj skali. Primjeri su iz teorije izbora ali se lako mogu 'prevesti' u obrazovno okruženje.

7.8.1 Neke 'paradoksalne' situacije kod agregiranja medijanom

Evo nekoliko situacija u kojima rezultati agregacije pomoću medijana odudaraju od 'očekivanog rezultata', s napomenom da je teško definirati očekivani rezultat jer to može biti posljedica ustaljene prakse ili naše zablude. Iz tog razloga i termin 'paradoks' treba uzeti s rezervom.

Paradoks 1. *Dodavanje glasačkog listića koji svim kandidatima daje najlošiju ocjenu može poremetiti redoslijed kandidata.*

Promatrajmo dva kandidata A i B , čije su ocjene:

	ocjene	medijan
A	1, 2, 4, 4, 6	4
B	2, 3, 3, 6, 6	3

Dakle, A je pobjednik. Ako dodamo još jedan glasački listić s nedovoljnom ocjenom za oba glasača, tada su nove ocjene

	ocjene	medijan
A	0, 1, 2, 4, 4, 6	2
B	0, 2, 3, 3, 6, 6	3

odakle slijedi da je B sada pobjednik jer agregacija uzima donju granicu srednjih vrijednosti.

Paradoks 2. *Dokidanje* Pretpostavimo da su dvoje prijatelja Romeo i Julija prvoj tablici iz gornjeg paradoksa naknadno dodali svoje ocjene: $A \leftarrow 5$, $B \leftarrow 6$ (Romeo) i $A \leftarrow 6$, $B \leftarrow 5$ (Julija). Nova tablica ocjena je

	ocjene	medijan
A	1, 2, 4, 4, 5, 6, 6	4
B	2, 3, 3, 5, 6, 6, 6	5

Romeo i Julija su očekivali²⁰ da će se njihove ocjene 'poništiti', ali nisu, B je postao pobjednikom.

Paradoks 3. *Povećana podrška nekom kandidatu može ga od pobjednika pretvoriti u gubitnika.*

Pretpostavimo da su Romeo i Julija, dvoje novih glasača, dali kandidatu A iz prvog primjera ocjenu 6, a kandidatu B ocjenu 5. Nova tablica ocjena je:

	ocjene	medijan
A	1, 2, 4, 4, 6, 6, 6	4
B	2, 3, 3, 5, 5, 6, 6	5

²⁰Ovo očekivanje je tipično za ljude koji su naviknuti usrednjavati. Usrednjavanjem na intervalnoj skali je moguće 'poništiti' takve ocjene.

Očigledno je A postao gubitnik.

To su samo neke od 'paradoksalnih' situacija. Ima ih još, a mogu se naći u spomenutoj knjizi od Balinskog i Larakija ili u autorovoj knjizi *Teorija odlučivanja s naglaskom na metodu potencijala*.

7.8.2 Bordin većinski indeks

Ordinalna agregacija pomoću medijana prirodno se uklapa u ordinalni kontekst ocjenjivanja, neosjetljiva je na manipulacije, ali ponekad nudi neočekivane i paradoksalne rezultate. S druge strane, aritmetička sredina izgleda da pati od izrazite mogućnosti manipulacije, s manje paradoksalnih mogućih ishoda, pa se prirodno javila ideja o ujedinjavanju obiju metoda. Jedan mogući hibrid nude Zahid i Swart²¹. Oni su eksperimentirali s lingvističkom skalom {odličan, vrlo dobar, dobar, prihvatljiv, slab, nedovoljan} i svoj indeks nazvali su *Bordin većinski indeks*²².

Neka je $\{g_1, \dots, g_k\}$ skup ocjena sa svojstvom $g_1 < g_2 < \dots < g_k$ i a jedan od kandidata na izborima. Neka je n_i broj glasača koji su kandidatu a dali ocjenu g_i , $i = 1, \dots, k$. Definirajmo *Bordin većinski indeks* $C(a)$ kandidata a formulom

$$C(a) := n_1 \cdot 0 + n_2 \cdot 1 + \dots + n_k \cdot (k - 1). \quad (\text{BVI})$$

Funkcija C definirana je na skupu profila s vrijednostima u \mathbb{N}^m , gdje je m ukupan broj kandidata. Napomenimo da vrijednost $C(a)$ ne ovisi o vrijednostima g_1, \dots, g_k , pa skala $\{g_1, \dots, g_k\}$ može biti i ordinalna lingvistička skala.

Za razliku od Bordine metode, ovdje pojedini glasači ne daju rang-listu kandidata, nego ih vrednuju na zadanoj ocjenskoj ljestvici. Bordina metoda svakom glasaču nudi jedino mogućnost da kandidate rangira na nekoj individualnoj skali u svom mentalnom prostoru. Bordin većinski indeks nudi nešto drugačiji pristup, on od svakog glasača zahtijeva da kandidata ocijeni na unaprijed zadanoj ljestvici ocjena koja je (trebala bi biti) opće prihvaćena. Takav pristup zahtijeva jaku socijalnu povezanost glasača i zdrav razum u implementaciji zadane ocjenske skale. Mnogi teoretičari izbornih metoda smatraju da je to smjela pretpostavka. Izborna slika danas je nešto gdje je prisutan sukob interesa, prisutno je taktizirano glasovanje i postoji pritisak na glasače (mediji, stranačka pristranost) da koriste **tuđe procjene**. Takvo ponašanje je blisko, po svojoj formi, kupovanju glasova samo se tako ne zove. U obrazovnom kontekstu glasač je nastavnik koji učeniku (kandidatu) daje ocjenu. Socijalna povezanost unutar jednog obrazovnog sustava je izrazitija jer nastavnici surađuju i imali su iste učitelje.

Agregacija kod računanja Bordina većinskog indeksa samo formulom podsjeća na Bordinu agregaciju. Kod Bordine metode, bodovi koje dobiva svaki kandidat na individualnoj listi ovise o broju kandidata, što ovdje nije slučaj. Na prvi pogled, većinski indeks je više utilitaristički nego ordinalan po svojoj prirodi. Ordinalnost je prisutna jedino u izboru ocjenske skale.

Većinski indeks C ima sva dobra svojstva kao i utilitaristički indeks S i manje je osjetljiv na manipulacije jer je skala sužena s intervala na svega 6 mogućih ocjena. Sve primjedbe i paradoksi koje uzrokuje agregacija pomoću medijana sada ne stoje što trivijalno slijedi iz definicije većinskog indeksa. Štoviše, većinski indeks čuva konzistentnost prioriteta i pobjednika o čemu govori sljedeći teorem.

Teorem 7.6. *Većinski indeks čuva konzistentnost pobjednika, tj. ako neki kandidat pobjeđuje u svakom okrugu onda pobjeđuje i u cijelom elektoratu.*

Većinski indeks čuva konzistentnost prioriteta, tj. ako za dva kandidata a, b vrijedi $a \succcurlyeq b$ u svakom okrugu, onda je $a \succcurlyeq b$ u cijelom elektoratu.

Dokaz. Neka je IUII particija izbornog tijela i $\succcurlyeq_I, \succcurlyeq_{II}$ relacije socijalne preferencije za oba dijela respektivno. Pretpostavimo da je

$$a \succcurlyeq_I b \succcurlyeq_I c \quad \text{i} \quad a \succcurlyeq_{II} b \succcurlyeq_{II} c.$$

²¹u svojoj knjizi *The Borda Majority Count*. Department of Philosophy, Tilburg University, the Netherlands (2010)

²²eng. *Borda majority count*. Gotovo identičan tzv. *Range Voting* samo ima manju skalu ocjena.

Zbog $C = C_I + C_{II}$ slijedi

$$a \succcurlyeq b \succcurlyeq c,$$

što dokazuje i konzistentnost prioriteta i konzistentnost pobjednika. \square

Primjer. Ovo je artifičijelni primjer koji pokazuje da većinski indeks može dati drugačije rangiranje od aritmetičke sredine. Tablica 2 može se dobiti na sljedeći način.

Pretpostavimo da je učenik A na predmetu I dobio niz ocjena a_1 , na predmetu II je dobio niz ocjena a_2 i na predmetu III je dobio niz ocjena a_3 :

$$a_1 = (2, 5, 4, 6, 3)$$

$$a_2 = (4, 2, 7, 5, 2)$$

$$a_3 = (1, 4, 5, 3, 6)$$

Suma ocjena u nizovima je 63, 59, 66 respektivno što predstavlja prvi redak tablice. Ocjene za učenika B neka su:

$$b_1 = (0, 5, 7, 5, 3)$$

$$b_2 = (0, 4, 7, 7, 0)$$

$$b_3 = (3, 4, 7, 8, 1)$$

i za učenika C

$$c_1 = (0, 1, 6, 6, 5)$$

$$c_2 = (0, 1, 3, 5, 5)$$

$$c_3 = (1, 2, 2, 8, 5).$$

Ako niz a_1 zamijenimo s nizom $a'_1 = (0, 3, 4, 5, 5)$ koji ima jednaku sumu ocjena, većinski indeks daje drugačiji poredak kandidata. To je zato jer većinski indeks uvažava broj ukupan ocjena koje je učenik dobio iz svih predmeta. Za većinski indeks važna je povijest dobivanja ocjena, a ne samo njihov zbroj po predmetu kao što to radi aritmetička sredina.

7.8.3 Usporedba medijana i aritmetičke sredine

U tablici 6 učenici A, B, C ocijenjeni su na skali 0-1-2-3-4 (veća ocjena je bolja) iz pet predmeta. U predzadnjem retku je računat prosjek, a u zadnjem medijan. Primijetimo da je učenik A najbolji

Tablica 6: text

	A	B	C
	4	3	1
	4	3	2
	2	0	3
	2	3	4
	1	0	2
BVI	18	12	28
ar. sred.	2.6	1.8	2.4
medijan	2	3	2

po aritmetičkoj sredini, a učenik B je najbolji po medijanu. Učenik C je, ako ga uspoređujemo s drugima prema broju uspješnijih ocjena²³, bolji i od A i od B (posljednja 3 retka tablice ocjena).

²³Condorcetova metoda većine

Bordin većinski index (BVI) daje vrijednosti 18, 12, 28 što stavlja učenika C na prvo mjesto kao i kod metode većine. Na *Stanford Encyclopedia of Philosophy*²⁴ dan je izbor literature koja zastupa medijan, s jedne strane, i izbor literature koja je protiv medijana kao metode izbora, s druge strane. Želim ovdje istaknuti da nema nekog konsenzusa da je medijan bolji ili lošiji od aritmetičke sredine. Zapravo te dvije metode agregiraju različite podatke, medijan agregira položajne liste, a aritmetička sredina agregira originalne vrijednosti.

Što se aritmetičke sredine tiče, ona je podložna manipulaciji, dozvoljava kompenzaciju u ocjenjivanju i stoga obrazovni sustav unosi prolazni prag u ocjenjivanje. Posljedica toga je da nastavnici izjegavaju ocjene neposredno ispod praga. Navedimo još jedan primjer koji govori u prilog zdravog razuma u donošenju grupne rang-liste u slučaju kad postoji *preferencijalna zavisnost atributa*.

7.9 Preferencijalna (ne)zavisnost atributa

Agregiranje ocjena je posebno teško modelirati jer tu dolazi do izražaja efekt interakcije među atributima (kriterijima) koji narušava njihovu aditivnost. Evo primjera za ilustraciju.

Primjer 7.7. Četiri učenika pristupila su ispitima iz fizike, matematike i ekonomije i rezultati tih ispita dani su u sljedećoj tablici učeničkih profila:

	<i>F</i>	<i>M</i>	<i>E</i>
<i>a</i>	18	12	6
<i>b</i>	18	7	11
<i>c</i>	5	17	8
<i>d</i>	5	12	13

Neka je \succsim neka relacija preferencije na studentima. Pitanje je ima li ta relacija aditivnu reprezentaciju, tj. može li suma atributnih funkcija vrijednosti oblika

$$v_F + v_M + v_E$$

predstavljati njenu funkciju vrijednosti. **Jedan od uvjeta** za takvo nešto je *preferencijalna nezavisnost* sributa koju je najbolje objasniti ilustracijom na našem primjeru.

Iz tablice vidimo da studenti *a, b* imaju jednak broj bodova iz fizike, a prema bodovima iz matematike i ekonomije *a* je bolji od *b*. Isto razmišljanje vrijedi i za studente *c, d*. Tada su matematika i ekonomija, kao atributi, *preferencijalno nezavisni* ako $a \succsim b \iff c \succsim d$. Očito je da aritmetička sredina odražava preferencijalnu nezavisnost atributa po konstrukciji.

Osim preferencijelne nezavisnosti postoji još nekoliko uvjeta poznatih kao uvjeti rješivosti za aditivnu reprezentaciju funkcije vrijednosti ali ih nećemo objašnjavati iz jednostavnog razloga jer su previše tehnički.

Pogledajmo sada kako razmišlja nastavnik koji boduje na ljestvici 0–20 ali s pragom prolaznosti 10. Tablica sugerira da studente *a* i *b* rangiramo ispred sudenata *c* i *d*. Student *c* je podbacio u dva predmeta i razumno ga je staviti na dno liste. Student *a* ima veći zbroj ocjena iznad praga od studenta *b* pa nastavnik smatra da bi studenta *a* trebalo rangirati ispred *b*. Slijedeći istu nit razmišljanja student *d* je bolji od studenta *c* jer ima ocjene iz matematike i ekonomije iznad praga, dok student *c* ima samo jednu takvu ocjenu. Stoga je razumno studenta *d* rangirati ispred *c*. Rezultat ovakvog razmišljanja je rangiranje

$$a \succsim b \succsim d \succsim c$$

za koje atributi *M* i *E* nisu preferencijalno nezavisni i koje se ne može dobiti računanjem srednjeg općeg uspjeha (bez obzira na težine predmeta).

Istaknimo još jednom da aritmetička sredina posjeduje mogućnost kompenzacije, tj. da student može podbaciti u jednoj komponenti i nadoknaditi taj neuspjeh uspjehom u drugim komponentama. Ako ispitivač želi kažnjavati nerad, onda će uvesti prolazni prag u ocjenjivanje koji generira nelinearnu skalu i takvo je razmišljanje bilo ugrađeno u rješavanje gornjeg primjera.

²⁴<https://plato.stanford.edu/entries/voting-methods/#VotiGrad>

7.9.1 Preferencijalna nezavisnost atributa.

Pretpostavimo, jednostavnosti radi, da su u igri dva atributa, tj. da je svaka alternativa opisana s uređenim parom $(x, y) \in X \times Y$. Reći ćemo da je atribut X *preferencijalno nezavisan* od atributa Y ako za svake dvije vrijednosti $x, x' \in X$

$$\begin{aligned} & (x, y_0) \succcurlyeq (x', y_0) \text{ za neko } y_0 \in Y \\ \implies & (x, y) \succcurlyeq (x', y) \text{ za svako } y \in Y. \end{aligned}$$

Isto tako, atribut Y je *preferencijalno nezavisan* od atributa X ako za svake dvije vrijednosti $y, y' \in Y$

$$\begin{aligned} & (x_0, y) \succcurlyeq (x_0, y') \text{ za neko } x_0 \in X \\ \implies & (x, y) \succcurlyeq (x, y') \text{ za svako } x \in X. \end{aligned}$$

Ako vrijede oba gornja zahtjeva, onda kažemo da su X i Y *preferencijalno nezavisni* atributi.

Preferencijalna nezavisnost nije prisutna u svakoj situaciji. Tako, na primjer, ako je neka kemijska reakcija uspješnija kad su mase oba reaktanta približno jednake, onda je očito

$$(4, 5) \succcurlyeq (4, 6) \text{ i } (7, 6) \succcurlyeq (7, 5)$$

gdje je prva komponenta masa prvog, a druga komponenta masa drugog reaktanta. Slične se situacije nalaze i u ekonomiji kada se uspoređuju robne košarice od dva ili više artikala koji su korisni samo ako su oba prisutna. Tada je bolja ona košarica koja ima više i jednog i drugog artikla. Ako promatramo attribute maksimalna brzina i sigurnost/pouzdanost automobila, onda je nerazumno tvrditi da su nezavisni.

Preferencijalna nezavisnost atributa omogućava konstrukciju *marginalnih* relacija preferencija \succcurlyeq_X i \succcurlyeq_Y na skupovima X i Y koje mogu imati i svoje reprezentacije v_X i v_Y . Pitanje je uz koje dodatne uvjete njihova suma $(x, y) \mapsto v_X(x) + v_Y(y)$ reprezentira \succcurlyeq na produktu $X \times Y$.

Formulirajmo aksiom nezavisnosti i u općem slučaju:

Aksiom nezavisnosti atributa.

Neka je \mathcal{A} skup atributa, $\mathcal{J} \subset \mathcal{A}$ i $\mathcal{J}' = \mathcal{A} \setminus \mathcal{J}$ skup komplementarnih atributa. Reći ćemo da je \mathcal{J} *preferencijalno nezavisan* skup atributa u \mathcal{A} ako za bilo koja 4 ispitanika a, b, c, d za koja vrijedi

$$\begin{aligned} v_i(a) &= v_i(b), & \forall i \in \mathcal{J}' \\ v_i(c) &= v_i(d), & \forall i \in \mathcal{J}' \\ v_i(a) &= v_i(c), & \forall i \in \mathcal{J} \\ v_i(b) &= v_i(d), & \forall i \in \mathcal{J} \end{aligned}$$

zaključujemo da je $a \succcurlyeq b$ ako i samo ako je $c \succcurlyeq d$.

Drugim riječima, \mathcal{J} je *preferencijalno nezavisan* skup atributa ako preferencija u kvaliteti između dva ispitanika ne ovisi o vrijednostima atributa izvan \mathcal{J} .

Nezavisnost atributa. Primjer. Pretpostavimo da je rezultat ocjenjivanja nekog broja ispitanika dan skupom profila prema nekim kriterijima (atributima) \mathcal{A} i da je \succcurlyeq relacija preferencije na skupu profila izvedena iz parcijalnih bodova v_i gdje je $i \in \mathcal{A}$. Tako na primjer, u tablici 7 atributi I i II su nezavisni ako $(a \succcurlyeq b \iff c \succcurlyeq d)$. Kad bi pravilo agregacije bilo takvo da iz svakog profila zbrojimo tri najbolje ocjene, i rangiramo ih na temelju te sume, takvo bi pravilo razbilo nezavisnost atributa. U većini slučajeva (ispita) zahtijeva se nezavisnost atributa. Posljedica toga je da funkcija agregacije ima jednostavan oblik:

$$\sum_{i=1}^n v_i(x_i),$$

	I	II	III	IV
a	49	60	88	71
b	53	58	88	71
c	49	60	33	46
d	53	58	33	46

Tablica 7: Nezavisnost atributa I i II.

gdje je n broj komponenti koje agregiramo i x_i ocjena dobivena iz i -te komponente. Kako osigurati da nezavisnost atributa bude prisutna u ocjenjivanju je priča za sebe, v. French (1998); Roberts (1979).

Funkcije v_i određuju se tako da ispitivač procijeni koliko bodova jedne komponente vrijedi jedan (ili više bodova) druge komponente. Taj 'trade-off' (trgovina) ne mora biti konstantan, može ovisiti i o samim vrijednostima već dobivenih bodova. Drugim riječima v_i može biti nelinearna funkcija. Na primjer, u matematici ili fizici, praktični zadaci i teorijska pitanja mogu imati jednu trade-off vrijednost ako je ispitanik nešto iznad praga, a drugu vrijednost ako je ispitanik kandidat za visoku ocjenu. Razlog tome je iskustvena činjenica da male vrijednosti ispod praga prolaznosti sugeriraju kandidatu da traži dodatne ustupke kako bi prošao prag.

Preferencijalna zavisnost ocjena. Preferencijalna nezavisnost atributa je vrlo strog uvjet koji je rijetko ispunjen u praktičnim primjenama. Evo jednog primjera iz školske prakse.

U tablici 8 dane su ocjene iz fizike (F), matematike (M) i ekonomije (E) za 4 učenika a, b, c, d . Uobičajena tehnika računanja zajedničke ocjene za sve ispite je uz pomoć formule za srednju ocjenu,

Tablica 8: Preferencijalna zavisnost ocjena

	F	M	E
a	18	12	6
b	18	7	11
c	5	17	8
d	5	12	13

u kojoj su svi predmeti jednako važni. Srednje ocjene studenata a, b su jednake i iznose 12, a isto tako su jednake i srednje ocjene studenata c, d i iznose 10.

Pogledajmo sada kako razmišlja nastavnik. Napomenimo još da je maksimalna moguća ocjena 20 i da je prag prolaznosti 10. Tablica sugerira da studente a i b rangiramo ispred studenata c i d . Student c je podbacio u dva predmeta i razumno ga je staviti na dno liste. Student a ima veći zbroj ocjena iznad praga od studenta b pa nastavnik smatra da bi studenta a trebalo rangirati ispred b . Slijedeći istu nit razmišljanja student d je bolji od studenta c jer ima ocjene iz matematike i ekonomije iznad praga, dok student c ima samo jednu takvu ocjenu. Stoga je razumno studenta d rangirati ispred c . Rezultat ovakvog razmišljanja je rangiranje

$$a > b > d > c$$

koje se ne može dobiti računanjem srednjeg općeg uspjeha (bez obzira na težine predmeta). Ostavljamo čitatelju da sam dokaže navedenu tvrdnju.

Aritmetička sredina zadovoljava princip *preferencijalne nezavisnosti* jer ako

... dva studentska profila imaju istu vrijednost ocjene po nekom kriteriju, onda se promjenom težine tog kriterija ne može promijeniti odnos preferencija u ukupnom poretku izračunatom na temelju svih kriterija.

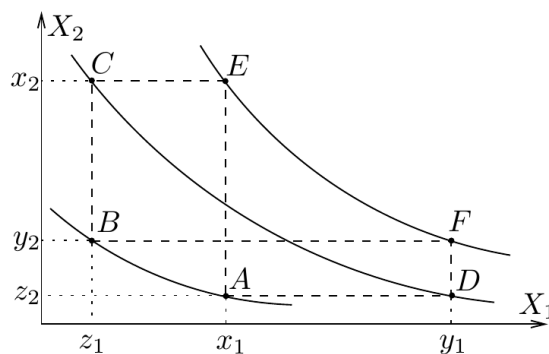
Jedno svojstvo aritmetičke sredine je mogućnost kompenzacije, tj. da student može podbaciti u jednoj komponenti i nadoknaditi taj neuspjeh uspjehom u drugim komponentama. Ako ispitivač želi kažnjavati nerad, onda će uvesti prolazni prag u ocjenjivanje koje generira nelinearnu skalu. Takvo je razmišljanje bilo ugrađeno u rješavanje gornjeg primjera.

Thomsenov uvjet. Kažemo da preferencija \succsim na produktu $X \times Y$ zadovoljava *Thomsenov uvjet*: ako za svaki $x_1, y_1, z_1 \in X$ i svaki $x_2, y_2, z_2 \in Y$

$$\left. \begin{array}{l} (z_1, y_1) \sim (x_2, z_2) \\ (z_1, y_2) \sim (x_1, z_2) \end{array} \right\} \implies (x_1, y_1) \sim (x_2, y_2). \quad (1)$$

Ako uređeni par (x, y) interpretiramo kao kvocijent x/y onda je Thomsenov uvjet zapravo pravilo dvostrukog kraćenja razlomaka. Ilustracija Thomsenovog uvjeta dana je na slici 5.

Slika 5: Grafička ilustracija Thomsenovog uvjeta.



Teorem egzistencije aditivne reprezentacije može se naći knjizi Krantz et al. (1971).

7.9.2 Ilustracija MAVT.

Pretpostavimo da se radi o vrednovanju kuća na prodaju. Kriterij vrednovanja može biti novčana vrijednost (cijena), površina izražena u m^2 (pov), udaljenost od stanice vlaka u minutama hoda (udaljenost), dodatna ulaganja (ulog) i osjećaj bivanja u toj kući (bivanje). Svi kriteriji su mjerljivi u konkretnim, mjernim jedinicama osim posljednjeg.

Dodatna ulaganja mogu se mjeriti količinom dodatnog novca i truda potrebnog da se kuća dovede u 'željeno stanje'. Grubo razmišljanje bi takav dodatni trošak pribrojilo cijeni kuće i tako umanjilo broj kriterija. Međutim, 'željeno stanje' nije jasno definiran pojam, a to 'stanje' ovisi i o konkretnom stanu i o sposobnosti kupca da vizualizira takvo stanje. Osim toga, dodatno ulaganje je teško izraziti precizno; najbolje što se može reći je neki interval cijene od–do, a osim financijskog troška može zahtijevati i osobni angažman koji se mjeri trajanjem i intenzitetom osobnog uloga koji uključuje i fizički rad. Ozbiljnija financijska konstrukcija bi uvažila i mogući dobitak/gubitak zarade na radnom mjestu ili nekom drugom poslu za vrijeme osobnog angažmana. Ako DO posjeduje još i znanje i mogućnost unapređivanja stanovanja s ciljem učinkovitijeg korištenja energije ima razloga ostaviti taj atribut kao zaseban.

Nadalje, sama cijena stana je podatak koji podliježe subjektivnoj reviziji, ovisno o kupovnoj moći donositelja odluke (DO). Kupovna moć je ovdje vrlo važna, jer porastom kupovne moći dolaze do izražaja drugi, subjektivni kriteriji koji su posljedica kupčevih životnih navika i stavova. Iz tog razloga treba reći nešto i o kupcu kako bismo se mogli uživiti u njegovo razmišljanje.

DO odluke, ovdje je to šesteročlana obitelj s bakom kao najstarijom, bračnim parom i troje djece. Djeca imaju 6, 14 i 18 godina. Najstarija kćer sprema se upisati medicinu. Mama i tata su zaposleni, a baka je još radno sposobna žena od 58 godina u mirovini. Voli spremati zimnicu i uređivati vrt oko kuće. Otac povremeno radi kod kuće, a postoje periodi kad putuje radi održavanja sustava za

napajanje solarnom energijom. Mama radi kao fizioterapeut i razmišlja o tome da otvori svoj salon za njegu tijela i proizvodnju kozmetičkih preparata u suradnji s prirodom.

Svaka analiza odluke o kupnji kuće ima eliminatorne kriterije. To su izbor kvarta, udaljenost kuće od posla, škole, fakulteta, minimalna površina, mogućnost dobivanja kredita i još mnogo sličnih. U širem izboru bilo je desetak kuća od kojih je nakon eliminacije ostalo samo 4; označimo ih s K_1, K_2, K_3, K_4 , a njihove karakteristike prema navedenim kriterijima dane su u tablici 9.

Tablica 9: Tablica odlučivanja za kupnju stana

	cijena (10^3€)	pov (m^2)	ulog (€)	bivanje
K_1	150	240	15–25	A
K_2	140	270	80–90	D
K_3	130	210	90–100	C
K_4	200	300	50–70	B

Atribut *bivanje* je sveukupni doživljaj budućeg življenja u toj kući i kao takav nosi i neke popratne aspekte drugih atributa koji nisu potpuno izraženi samim atributom. Tako na primjer atribut cijena je jasan sam po sebi, ali njegov popratni efekt 'osjećaj dobre kupnje' se može uklopiti u osjećaj bivanja u tom objektu.

Dobivena tablica odlučivanja mogla bi se analizirati MAVT metodom uz pretpostavku preferencijalne nezavisnosti atributa ali je glavna prepreka što *bivanje* nije kvantificirano. Ima još jedan problem što se tiče revizije vrijednosti površine o čemu će kasnije biti riječi, a poteškoća kod uloga je taj što je intervalnog tipa.

U svrhu lakše kvantifikacije bivanja korisno je odrediti kategorizaciju (*A, B, C, D, E* na primjer) i opisati svaku klasu te kategorije. Tako na primjer u klasu *A* može se smjestiti solidno građena kuća s dobrom toplinskom izolacijom i izoliranim temeljima, dobrom zvučnom izolacijom i daleko od zagađivača (promet, industrija, elektromog), stupnjem urbanizacije kvarta...

Jednostavni model. Opisani pristup je brz i zasniva se na reskaliranju ulaznih vrijednosti prema intuiciji. To je najučestaliji pristup u poslovanju ako se želi donijeti brza i donekle interpretabilna odluka. Mogli bismo je nazvati *ad hock* metodom jer zaobilazi bilo kakvu raspravu o njenoj opravdanosti. Postupci se čine prihvatljivim, ali se ne sagledavaju njihove posljedice.

Atribut *ulog* je intervalnog tipa i mogli bismo ga svesti na numeričke vrijednosti uzimanjem srednje vrijednosti intervala i pribrojiti cijeni što je izraženo u stupcu *izdaci* u tablici 10.

Površinu bismo mogli izraziti u novčanim jedinicama i dobiti njen ekonomski ekvivalent. Međutim, površina objekta ima i subjektivnu vrijednost (korisnost) za DO koja nema novčani ekvivalent. Štoviše, taj atribut uključuje, osim same površine, i raspored prostorija kao i mogućnost njihove prenamjene. Mogli bismo uvesti novi atribut 'arhitektura' za takvo nešto ali je za naše ilustrativne potrebe dovoljna i takva površina. Za vrijednost površine mogli bismo zahtijevati da bude rastuća funkcija koja je još i konveksna, što znači da marginalna vrijednost površine (derivacija) raste. Revidirana površina dana je u stupcu *rev_pov*.

U tablici 10 varijabla *izdaci* ujedinjuje cijenu i ulaganje izraženi u €, *bivanje* je irażeno na simbo-

Tablica 10: Transformirana tablica odlučivanja za kupnju stana prilagođena jednostavnoj i brznoj analizi.

	izdaci (€)	rev_pov	bivanje
K_1	170	250	A
K_2	225	290	D
K_3	225	210	C
K_4	260	340	B

ličkoj skali 1–5, a revidirana vrijednost površine je izražena tako da uvažava konveksnost i raspon vrijednosti počev od minimalne. Izdaci, u principu, također podliježu subjektivnoj reviziji ovisno o kupovnoj moći DO²⁵.

Dobivena tablica je jednostavnija od početne ali još uvijek heterogena. Pitanje je kako odredite *trade-off* među atributima i ima li potrebe za davanjem težina atributima.

Srednja cijena jedinične površine, prema tablici 9, iznosi $c = 0.6\text{€}$ što ujedno predstavlja i relativan odnos težina atributa *izdaci* i *rev_pov*. Uz pretpostavku preferencijalne nezavisnosti atributa i Thomsenovog uvjeta, kojeg nismo provjeravali, sada možemo zaključiti analizu na temelju ta dva atributa, što izdvaja kuću K_1 kao najbolju, tablica 11.

Tablica 11: Odluka o izboru kuće bez atributa *bivanje*.

	izdaci (€)	rev_pov	→	$0.6 * rev_pov - izdaci$
K_1	170	250	→	-18.03922
K_2	225	290	→	-48.72549
K_3	225	210	→	-97.35294
K_4	260	340	→	-53.33333

Isto tako omjer (zaokruženo na 2 decimalna mjesta)

$$0.6 * rev_pov / izdaci = \begin{matrix} K_1 & K_2 & K_3 & K_4 \\ 0.89 & 0.78 & 0.57 & 0.79 \end{matrix}$$

predstavlja dobit po jedinici ulaganja za svaku kuću. Očito se najbolji omjer dobivenog i uloženog postiže za kuću K_1 i iznosi 0.8938870. Odluka je stabilna jer se za široki interval vrijednosti jedinične cijene c generira ista odluka, a ista odluka se generira i bez revidiranja vrijednosti površine.

Primijetimo da je na temelju atributa *bivanje* kuća K_1 također najbolja čime je dodatno potvrđen njen izbor.

8 Subjektivno mjerenje

8.1 Ordinalna mjera

Sljedeći teorem govori o važnosti tranzitivnosti osobnih preferencija za njihovu numeričku reprezentaciju.

Teorem 8.1 (Savage). *Neka je S konačan skup. Relacija \succsim je slaba preferencija na S ako i samo ako postoji funkcija $V : S \rightarrow \mathbb{R}$ koja je u skladu s \succsim tj.*

$$x \succsim y \iff V(x) \geq V(y). \quad (2)$$

Teorem ovog tipa, govoreći općenito, daje numeričku reprezentaciju relacije uz određene pretpostavke. Takve teoreme nazivamo **teoremima reprezentacije**. Očito je da ako postoji funkcija V i relacija \succsim koja zadovoljava (2), onda je to slaba preferencija koja je jaka ako u (2) stoji stroga nejednakost s desne strane. Svaku funkciju V koja je u skladu s relacijom preferencije (slabom ili jakom) nazivamo *ordinalnom funkcijom vrijednosti* ili *ordinalnom mjerom*.

Nadalje, ako V zadovoljava (2), onda je ona jedinstvena do na kompoziciju sa strogo rastućom funkcijom. Stoga ćemo se u teoremima koji slijede koncentrirati samo na egzistenciju funkcije V , a ne i na njenu jedinstvenost.

²⁵Principima takve revizije bavili su se Kahneman and Tversky (1979) i razvili Teoriju izglednosti za koju je Kahneman dobio i Nobelovu nagradu. Jednostavnosti radi mi ćemo stupac *izdaci* ostaviti takvim kakav jest.

Dokaz. Neka je \succsim slaba preferencija. Definirajmo

$$V(x) := \#S(x) \quad (3)$$

gdje je $\#$ oznaka za broj elemenata skupa i $S(x) = \{y \in S \mid x \succsim y\}$. Tvrdimo da V zadovoljava (2). Očito $x \succsim y \implies S(y) \subseteq S(x)$ (tranzitivnost), što povlači nejednakost $V(x) \geq V(y)$. Za dokaz obrnute implikacije u (2) pretpostavimo $V(x) \geq V(y)$. Na isti način kao gore vrijedi $x \succ y \implies S(y) \subset S(x)$. Ako nije $x \succsim y$, onda je (zbog potpunosti) $y \succ x$, što povlači $S(x) \subset S(y)$ odnosno $V(y) > V(x)$, a to je u kontradikciji s pretpostavkom.

Za dokaz obrata u teoremu pretpostavimo da postoji $V : S \rightarrow \mathbb{R}$ koja zadovoljava (2). Tada je relacija \succsim slaba preferencija jer je očito tranzitivna i potpuna. \square

Napomena. Ako s χ_R označimo indikatorsku (karakterističnu) funkciju relacije R (kao podskupa od $S \times S$), onda tu karakterističnu funkciju možemo zapisati kao matricu s elementima

$$\chi_R(a, b) = \begin{cases} 1, & \text{ako } aRb \\ 0, & \text{inače} \end{cases} \quad (4)$$

onda je $V(a)$ suma elemenata te matrice u retku a . Naravno, takva matrična reprezentacija postoji ako je S konačan ili prebrojiv skup.

Konstrukcija slabe preferencije u praksi. U mnogim praktičnim situacijama donositelj odluke je u stanju preferirati samo one objekte koji se dovoljno razlikuju po nekom atributu, na primjer, glasnoća zvuka. To znači da je definirao relaciju R na skupu objekata koja nije nužno potpuna. U slučaju glasnoće ona je u najboljem slučaju asimetrična i negativno tranzitivna²⁶. Nepotpunost sugerira da se promatra relacija (S, E)

$$xEy \iff y \neg Rx \ \& \ x \neg Ry \quad (5)$$

koju nazivamo *simetričnim komplementom* od R . Očito je E relacija ekvivalencije (što opravdava i njen naziv) i neprazna ako i samo ako je R nepotpuna. Nadalje $R \cap E = \emptyset$, a dodavanjem te relacije relaciji R , tj. $W := R \cup E$, dobiva se potpuna i tranzitivna relacija koju nazivamo *slabo zatvorenje* od R . Relacija R je tada asimetrični dio, a relacija E je simetrični dio od W . Dokaze gornjih tvrdnji ostavljamo čitatelju.

Vrijedi sljedeći teorem koji naglašava praktičnost negativne tranzitivnosti relacije R , a time i egzistenciju ordinalne mjere (bez dokaza).

Teorem 8.2. *Asimetrična relacija R je negativno tranzitivna ako i samo ako je tranzitivna i relacija E definirana formulom (5) je relacija ekvivalencije.*

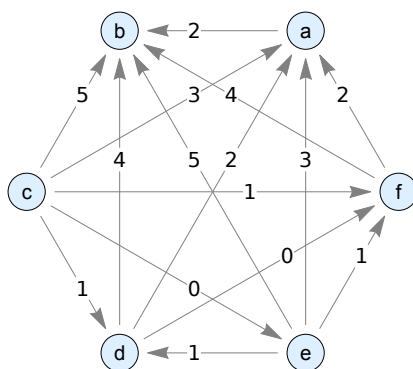
8.2 Graf preferencije

Zamislite sebe kako razmišljate o nekim svojim postupcima iz prošlosti i pokušavate ustanoviti stupanj vlastite slobode odabira u svakom od njih. U nekim situacijama bili ste 'primorani', a u nekim drugim situacijama ste mogli postupiti i drugačije, bilo vam je dano na 'volju'. Jedan od načina da to učinite je taj da sami odredite skalu od 0–10 (ili 0–100) i na toj skali poredate postupke (opcije) po vlastitom nahođenju. Kod takvog vrednovanja vi se oslanjate na svoj osjećaj, iskustvo,... nazovite to kako želite, ali misaoni proces u pozadini takvog vrednovanja vam je nedostižan.

Druga mogućnost je da pokušate razmotriti dvije opcije i prednost date onoj u kojoj je bilo više slobode odabira. Ako možete tu prednost izraziti brojem još bolje. Učinite to za svake dvije opcije ili samo za one za koje ste sigurni u procjeni. Recimo da su vaše usporedbe u parovima

²⁶tj. njena negacija (komplementarna relacija) je tranzitivna.

Slika 6: Šest opcija



kao na slici 6, gdje je strelica usmjerena prema opciji kojoj dajete prednost. Na toj slici postoji određena konzistentnost u procjenama jer za svake odabrane tri opcije, napr. $\{c, f, b\}$ suma brojeva duž zatvorenog puta $cfbc$ jednaka je $1 + 4 - 5 = 0$. Na tom putu ste dionicu bc prolazili u suprotnom smjeru od strelice i njen doprinos sumi iznosi -5 . Upravo zbog navedene konzistentnosti nije teško zaključiti da je jedno moguće vrednovanje tih opcija

a	b	c	d	e	f
4	6	1	2	1	2

Za dokaz gornje tvrdnje povucimo paralelu prikazanog grafa s električnom mrežom u kojoj je razlika potencijala između dva čvora dana brojem uz odgovarajuću žicu (strelicu). Dovoljno je primijetiti da možemo fiksirati vrijednost potencijala V jednog čvora, napr. $V(c) = 1$, a potencijale ostalih čvorova odredimo uvažavajući danu razliku potencijala.

Što ako konzistentnost koju smo opisali nije zadovoljena? Tada se stvari kompliciraju (v. odjeljak 11.1).

Sliku koje ste dobili, u literaturi nazivamo *usmjerenim grafom s težinama*. Taj graf može poslužiti za daljnju analizu vaših podataka. Jedna od njih je *metoda potencijala*, autorova zamisao, o kojoj će kasnije biti riječi.

U drugim konkretnim situacijama umjesto 'osjećaja' prepoznamo kvalitete: duljinu, poželjnost, riskantnost, vrijednost, a entiteti koje uspoređujemo su fizički objekti ili misaoni konstrukti kao što su: akcije, scenariji, opcije...

U osnovi mjerenja stoji uspoređivanje objekata u parovima, što znači da za svaka dva entiteta ili opcije, mjerni instrument ili čovjek donositelj odluke treba biti u stanju reći koji je objekt dulji, poželjniji, riskantniji, kvalitetniji, vrijedniji. Rezultate takvih usporedbi nazivamo, u najapstraktnom smislu, *relacijom* na skupu entiteta; mi ćemo iz razloga konzistentnosti s literaturom tu relaciju nazivati *preferencijom* kao što smo to činili i u odjeljku 7.4 na str. 23.

8.3 Mjerenje razlike preferencije

Pretpostavimo da je (S, \succ) relacija preferencije. Na prvi pogled izgleda razumno reći da donositelj odluke preferira a u odnosu na b više nego što preferira c u odnosu na d ako i samo ako je spremniji odustati od b u zamjenu za a nego odustati od d u zamjenu za c .

Označimo s $(a \leftarrow b)$ prihvaćanje a u zamjenu za b ili kraće *zamjena*. Donositelj odluke sada ima dvije relacije slabe preferencije s kojima se mora suočiti, jedna je relacija \succ na objektima iz S , a druga je relacija \succ_e na skupu zamjena $(S \leftarrow S)$ koji je zapravo kartezijev produkt $(S \times S)$. Svaka relacija za sebe generira funkciju vrijednosti (uz određene uvjete), s napomenom da je prva definirana

na skupu objekata, a druga na skupu zamjena. Nas zanima njihova usklađenost, odnosno odgovor na pitanje je li moguće naći takvu funkciju vrijednosti v na skupu objekata S koja zadovoljava:

$$a \succ b \iff v(a) \geq v(b) \quad (6)$$

$$(a \leftarrow b) \succ_e (c \leftarrow d) \iff v(a) - v(b) \geq v(c) - v(d). \quad (7)$$

Takvu funkciju, ako postoji, nazvat ćemo *izmjerivom funkcijom vrijednosti*. Očito je v ordinalna funkcija vrijednosti, dok (7) zahtijeva da je razlika $v(a) - v(b)$ ordinalna funkcija vrijednosti na skupu zamjena usklađena s relacijom \succ_e na zamjenama.

Mnogi teoretičari smatraju da ne treba donositelje odluke prisiljavati na zamišljanje zamjena jer da oni imaju urođeni osjećaj za intenzitet preferencije. Zamišljamo li relaciju preferencije kao usmjeren graf, onda je intenzitet preferencije broj koji je pridružen svakom luku u grafu kao što je to prikazano na slici 6. Intenzitet tako postaje funkcija na skupu zamjena koja može i ne mora biti funkcija vrijednosti na skupu zamjena usklađena s \succ_e . Takva prezentacija slabe preferencije umnogome olakšava donositelju odluke da se koncentrira na parove i svakom paru (a, b) pridruži intenzitet preferencije²⁷. Na kraju krajeva nije bitno kako ljudi zamišljaju zamjene, bitno je da to rade na konzistentan način.

9 Snadi se kako znaš

9.1 Rangiranje profila

Evo jedan detaljniji. Pretpostavimo da je ocjenjivač bodovao testove tri ispitanika A, B, C na temelju profila u tablici 12. Najveći mogući broj bodova po predmetu je 100. Kvantitativni odnos između 66

Tablica 12: Bodovanje ispitanika.

	Predmet			zbroj
	I	II	III	
A	63	59	66	188
B	66	57	69	192
C	69	56	68	193

i 63 pokazuje da je kandidat B postigao bolji uspjeh od kandidata A na testu ako se gleda samo predmet I. U posljednjem stupcu tabele dana je suma bodova svakog kandidata koja predstavlja *opći uspjeh* kandidata za sva tri predmeta. Takav zbroj (ili srednja vrijednost) je načešće korištena mjera kvalitete općeg uspjeha.

Pretpostavimo da je neki drugi ispitivač također ocijenio iste testove i rezultat njegovog ocjenjivanja je skup profila predstavljen u tablici 13. Primijetimo da numeričke vrijednosti pridijeljene us-

Tablica 13: Drugo moguće bodovanje ispitanika.

	Predmet			zbroj
	I	II	III	
A	59	65	46	170
B	60	61	48	169
C	61	58	47	166

pjehu ispitanika, po predmentima, predstavljaju isti kvalitativni poredak za oba ocjenjivača. Drugi ispitivač je bio nešto blaži u ocjenjivanju za drugi predmet i stroži za preostala dva. Suma bodova međutim daje inverzni poredak za opći uspjeh. Razlog tome je što zbroj bodova, kao kriterij općeg

²⁷Bilo po svom nahođenju, prihvaćanjem nekih egzaktno izmjerenih vrijednosti ili njihovom revizijom.

uspjeha, uvažava i kvantitativnu razliku među bodovima, a ne samo poredak? Ako netko želi uvažavati samo kvalitativni odnos među učenicima kod agregacije onda sumacija bodova nije rješenje.

Ponavljam, ako broj bodova za svaki predmet shvatimo kao odraz *poretka* kandidata na ispitu, onda je računanje prosjeka očigledno *kvantitativno besmislena* operacija jer su oba ocjenjivača dala isti poredak kandidata za svaki ispit. Uzrok toj besmislenosti leži u izboru skale na kojoj mjerimo kandidate. Sada je to *ordinalna* skala jer reflektira samo poredak kandidata i ne uvažava njihove međusobne razlike na toj skali kao dodatnu mjeru kvalitete.

Skala koja uvažava razlike je *intervalna skala*. Što se ocjenjivača tiče, nije sasvim sigurno jesu li u stanju napraviti procjene koje uvažavaju razliku vrijednosti na skali, tj. nisu u stanju 'mjeriti' ispitanike na intervalnoj skali (French and Vassiloglou, 1986). Ashby and Townsend (1984) argumentiraju da ako su podaci dani na ordinalnoj skali onda su sredine, varijance kvantitativno besmislene i bilo kakvi statistički zaključci na njima bazirani nisu prikladni. Samo neparametarski testovi bazirani na rangovima mogu biti primjereni. Nužni uvjet da bi parametarski testovi mogli biti korišteni je da su podaci mjereni na intervalnoj skali.

Postoje i drugačija mišljenja, Gaito (1980), pa čak i uzrečica: *brojevi ne znaju odakle dolaze*. Ipak, opreznost je majka mudrosti. U članku French (1989) dana su dva primjera koja pokazuju da statistički testovi nad ordinalnim podacima imaju smisla ovisno o pitanjima koja zanimaju analitičara.

Da bi numerički odnosi među brojevima²⁸ bili kvalitativno interpretabilni oni svakako moraju biti kvantitativno smisleni. Međutim, još jedan uvjet treba biti ispunjen. Ti odnosi trebaju biti *semantički smisleni* tj. trebaju biti interpretabilni u kvalitativnom smislu u granicama korisnikove percepcije.

9.2 Dualnost u ocjenjivanju

Ovdje ćemo ilustrirati primjenu metode potencijala u ocjenjivanju učenika na temelju više atributa (predmeta). Tu će biti uvedeni i neki pojmovi a to su: (1) *operator adaptacije* i (2) uvažavanje *sinergije* bez eksplicitne definicije što bi sinergija trebala biti. Sinergija i adaptacija ne idu jedno bez drugog i teško ih je separirati jednako kao što je teško separirati prostor i vrijeme.

Primjer. Ocjenjivanje učenika. U tablici 14 dane su ocjene učenika na ispitu iz Fizike, Tjelesnog i Muzičkog, a znak * znači da učenik nije ocijenjen iz odgovarajućeg predmeta. Predmeti učenja su

Tablica 14: Ocjene učenika na završnom ispitu.

	Fizika	Tjelesni	Muzički
Ivica	3	5	4
Marica	*	1	3
Trsek	2	*	2
Grozdek	4	3	*

atributi, a učenici su *alternative*.

Ako odlučimo nekom metodom rangirati učenike po uspješnosti tada treba precizirati važnost atributa. Ako nije drugačije rečeno pretpostavljamo da predmeti podučavanja imaju podjednaku važnost jer bi se u suprotnom neki profesori mogli uvrijediti.

A što ako odlučimo rangirati predmete na temelju dobivenih ocjena konkretnih učenika? To je jednako moguće, ali koji je smisao toga? Takav postupak bi se mogao interpretirati kao rangiranje *atributa* (predmeta) po stupnju ovladavanja gradivom. Što veća vrijednost na rang listi to je gradivo dotičnog predmet bolje savladano.

²⁸Prosjeck ocjena na primjer.

Ako *attribute* shvatimo kao primarne onda svaki stupac možemo shvatiti kao metar, a vrijednosti u tom stupcu su mjere odgovarajućih alternativa (učenika) mjerene tim metrom. Ako *alternative* shvatimo kao primarne onda baždarimo metre (*attribute*) na konkretnim objektima mjerenja, a vrijednosti u odgovarajućem retku čine baždarene vrijednosti metara kriterija.

U svakom od spomenutih postupaka (mjerenje i baždarenje) potrebno je unaprijed zadati važnost (težinu) *atributa* odnosno *alternativa*. Rangiranje alternativa u prvom slučaju ćemo nazvati *primarnim*, a rangiranje kriterija u drugom slučaju zvati ćemo *dualnim* postupkom.

Sinergija i adaptacija Postoji i treća mogućnost, a ta je da provedemo primarni postupak za neke početne vrijednosti težina atributa, izračunamo težine alternativa i provedemo dualni postupak. Zatim ponovimo primarni postupak s novim, promijenjenim težinama, ponovimo dualni postupak i tako dalje. Osnovni korak u toj proceduri, a to je jedan primarni i jedan dualni postupak, nazivamo *operatorom samoprocjene*. To je adaptacija o kojoj smo govorili u početku.

Može se dokazati da operator samoprocjene, u slučaju da se radi o metodi potencijala kao metodom rangiranja, ima fiksnu točku (Čaklović, 2011), a to su one težine atributa koje ostaju nepromijenjene prilikom samoprocjene. Te parametre interpretiramo kao važnost *atributa* i *alternativa* koje tablica nosi u sebi.

Jedan od aspekata sinergije je i sinkronicitet (koherentnost) kako je to nazvao Jung. To je lijepa zamisao ali je s algoritamskog gledišta neupotrebljiva. Mi u svakodnevnom životu konstatiramo sinkronicitet ali na to gledamo kao na slučajnost ili na čudo.

U gotovo svim današnjim modelima učenja koristi se statistika i vjerojatnostni modeli koji stoje u njenoj pozadini. Vjerojatnost je također jedna funkcija vrijednosti²⁹ u koju je ugrađeno svojstvo, u literaturi poznato kao *nezavisnost o irelevantnoj alternativni*. U najkraćim crtama, to svojstvo karakterizira procese u kojima dodavanje nove opcije (alternative) u račun ne mijenja poredak starih opcija. Takvim procedurama je u samom startu otežana mogućnost da modeliraju i objasne one pojave u kojima se pojavljuje sinergija. Sinergija je pojava (konstelacija) koja drastično mijenja stanje svijeta ako je prisutna i teško ju je definirati u operativnom smislu.

Jednostavan primjer sinergije je detekcija bolesti pomoću simptoma. Za neke bolesti, ako se određeni simptomi pojave zajedno, iskusan liječnik će bez ikakvih laboratorijskih nalaza odrediti o kojoj se bolesti radi na temelju vlastitog iskustva.

Metoda potencijala omogućava uvažavanje sinergije ili na način da ju procjenjuju ljudi ili na način da proizlazi iz samih podataka. To je omogućeno operatorom samoprocjene (adaptacije) o kojem će sada biti rečeno nekoliko riječi.

Adaptacija je sposobnost sustava da procijeni stanje vlastitih parametara na temelju promijenjene okoline i svog vlastitog ustrojstva.

Svaki sustav (čovjek) reagira na okolinu koja mu nudi vrijednosti polaznih parametara ili mu određuje dinamiku promjene tih parametara. Operator adaptacije revidira, uvažavajući strukturu sustava, vrijednosti polaznih parametara i izračunava nove vrijednosti koje sustav koristi u procjeni vlastitog stanja i omogućava mu reakciju na promijenjeno stanje okoline.

Adaptacija nosi u sebi dualnost u odnosu između polaznih parametara i mogućih reakcija, tj. odgovora sustava na promjenu parametara. Tehnički termin za polazne parametre je *atributi*, a za moguće odgovore *alternative*. Evo konkretnog primjera.

Za podatke iz tablice 14 fiksna točka operatora samoprocjene je:

Muzički	Tjelesni	Fizika	Ivica	Grozdek	Marica	Trsek
0.339528	0.330731	0.32974	0.264151	0.261712	0.240225	0.233911

Prema rezultatu samoprocjene ispada da je Muzički bolje ovladan nego ostali predmeti. Velika razlika u težinama nastavnih predmeta ukazivala bi na problem čije uzroke treba tražiti van ove tablice.

²⁹u smislu da za dva događaja A, B , $A \supseteq B \iff P(A) \geq P(B)$.

9.3 Agregacija ispitnih bodova

Nešto kompliciranija situacija za modeliranje je rangiranje učenika na temelju više kriterija (formi ispita). Nastavnici koriste takve tablice jer u njih upisuju rezultate (bodove) dobivene na temelju domaćih zadaća, školskih zadaća (testova) i završnih ispita. Evo jednog jednostavnog primjera. U

Tablica 15: Tablica A prikazuje broj stečenih bodova učenika na: domaćim zadaćama, testovima i završnom ispitu.

Tablica A			
učenik	dom zad	test	završni
A1	90	75	70
A2	60	42	65
A3	30	70	
A4	45	45	75
A5	79	55	
A6	85	33	50
A7	60	70	40
A8	90	70	35
A9	90	70	
maks	100	100	100
broj	10	4	1
tež.	0.5	1	1

Tablica 16: Tablica B daje poredak uspješnosti učenika na temelju podataka iz tablice 15 računat pomoću metode potencijala. U tablici C nepostojeći podaci su nadopunjeni s 0 i zatim je računat prosjek brojeva u retku. U tablici D računat je prosjek postojećih podataka.

Tablica B	Tablica C	Tablica D
(kako jest)	(nadopuna s 0)	(nepotpuni prosjek)
A1 20.00	A1 76	A9 76.67
A9 17.69	A8 60	A1 76.0
A8 10.20	A4 57	A5 63.0
A5 8.24	A7 56	A8 60.0
A7 6.17	A2 54.8	A4 57.0
A3 3.86	A6 50.2	A3 56.67
A4 3.00	A9 46	A7 56.0
A2 2.25	A5 37.8	A2 54.8
A6 0	A3 34	A6 50.2

tablici 15 dani su rezultati provjere znanja 9 učenika iz domaćih zadaća (10), testova (4) i završnog ispita (1). Maksimalni broj bodova u 10 zadaća iznosi 100 (10 po svakoj zadaći), jednako kao i u 4 testa i 1 završnom ispitu. Broj ispita i broj bodova u praksi može varirati, jednako kao i relativan odnos maksimalnog broja bodova za različite vrste ispita. Prije agregacije potrebno je odrediti odnos među mjernim jedinicama za svaki stupac (*trade off*) i važnost (težinu) pojedine varijable-stupca. U praksi se *trade off* često zaboravlja i indirektno se provodi određivanjem težina što može dodatno pomutiti sliku³⁰. U ovom primjeru je namjerno pojednostavljena situacija na način da je maksimalan broj

³⁰Istraživanja pokazuju se 40% poslovnih odluka pokazalo krivim zbog nerazumijevanja težina ili zbog loših manipulacija s njima.

bodova za sve kategorije isti i isnosi 100 i da je mjerna jedinica (bod) za svaki stupac jednako vrijedan i umanjena je važnost domaćih zadaća u odnosu na ostale ispite za 50%.

Agregacija ocjena. Najčešća metoda ocjenjivanja na temelju prethodne tablice je takva da nastavnik za svaki ispit daje ocjenu na skali 1–5, na primjer, a zatim radi aritmetičku sredinu tih ocjena. U takvoj agregaciji se bodovi prevode u ocjenu, a ocjene se zatim agregiraju putem aritmetičke sredine.

Nekoliko je nejasnoća u takvom pristupu. Nije jasna procedura kako se bodovi prevode u ocjene jer je to dio nastavnikovog iskustva i treba mu vjerovati. Mnogi nastavnici preskaču davanje bodova i direktno donose ocjene na pojedinom ispitu po vlastitim kriterijima koji nisu dostupni javnosti. I na kraju, agregacija vrijednosti (ocjena) na ordinalnoj skali zbrajanjem je upitna jer nije invarijantna na dozvoljene transformacije skale. Što bi se radilo da ocjene nisu izražene brojevima, nego slovima A, B, C, D, E, F na primjer?

Agregacija bodova. Primijetimo da učenici A₃, A₅, A₉ nisu izašli na završni ispit. U takvoj situaciji, kad nemamo potpunu tablicu, pribjegava se *ad hock* rješenju u kojem se nepostojeći podaci popunjavaju s nulom (o), ili nekim drugim brojem, a zatim se računa težinski prosjek vrijednosti po retcima. Težine stupaca dane su u posljednjem retku tablice 15. Rezultati takvog računa dani su u tablici 16C.

Naišao sam i na metodu u kojoj se računa aritmetička sredina postojećih podataka u retku. Rezultati takve procedure dani su u tablici 16D. Primijetimo da se učenici A₈ i A₉ razlikuju u tome što je učenik A₈ pristupio završnom ispitu, A₉ nije. Metoda potencijala i metoda nepotpunog prosjeka, daju rangiranja koja se manje razlikuju od metode prosjeka nadopunjene tablice.

Rangiranje provedeno metodom potencijala dano je u tablici 16B, s tim da su uvažene težine varijabli (stupaca). Detalji procedure opisani su u odjeljku 11.3 na str. 55. Napomenimo samo da svaki stupac za sebe generira jedan graf preferencije i potencijal u tablici 16B je potencijal grafa preferencije koji je nastao agregacijom pripadnih grafova stupaca. Metoda potencijala daje iste rezultate kao metoda aritmetičke sredine samo u slučaju kad je tablica potpuna (teorem 11.2). Ovdje to nije slučaj i učenici A₉ i A₈ zamijenili mjesta u tablici B i tablici C.

Primjedbe. Ovakve tablice nastavnici danas uveliko koriste u nastavi, bez obzira na vrstu agregacije. One su *post festum*, tj. nisu više podložne promjenama nakon što je ispit obavljen. Drugim riječima, tablica nije formativnog karaktera u smislu da odražava trenutno stanje usvojenog znanja koje će možda već sutra biti drugačije.

Druga je primjedba, odnosno pitanje koje se prirodno nameće, zašto bi se nedolazak na završni ispit 'kažnjavalo' s o bodova? Jedna mogućnost je koristiti metodu potencijala za konstrukciju bodovne skale koja ne zahtijeva dodatno upotpunjavanje tablice.

Druga mogućnost je da se na temelju postojećih podataka "predvidi" broj bodova kojeg bi učenik postigao na završnom ispitu, a zatim koristiti metodu aritmetičke sredine. Što se upotpunjavanja nepostojećih podataka tiče postoji mnogo numeričkih procedura u obradi statističkih podataka koje bi bilo dobro istražiti u ovom kontekstu, ali to prelazi doseg ove knjige. Također nije sasvim jasno bi li nastavnici, učenici i njihovi roditelji prihvaćali metodu koja njima nije jednostavna i razumljiva. Njima je prihvatljivije kažnjavanje nedolaska s o (nula) bodova nego nešto što je van njima shvatljive domene.

Dodatno pitanje u vezi bodovanja je po čemu se bodovi razlikuju od ocjena, tj. zbog čega možemo smatrati da su bodovi dani na intervalnoj skali na kojoj je aritmetička sredina dozvoljena agregacija? Ovdje nema zadovoljavajućeg odgovora. Jedan od argumenata je taj da duljina intervala na bodovnoj skali, napr. 9 bodova, ima substancijalnog smisla. U terminima bodovanja konkretne pisane provjere znanja ili testa, to bi značilo da svaki pojedinačni bod (ili dva) za dio riješenog zadatka odražava jednak kvalitativni doprinos znanju. To je pitanje uravnoteženosti bodovanja. Potrebno je višegodišnje iskustvo u kalibraciji testova kako bi se tvrdilo da mjeri na intervalnoj skali.

Jedini donekle egzaktan način kalibracije testa je provesti ga na odabranoj grupi učenika i iskazati subjektivne preferencije na način kako se to zahtijeva u odjeljku pod nazivom *Mjerenje razlike preferencija* knjige (Čaklović, 2014, str. 147). Drugi, lakše provediv način, je specificirati unaprijed određen broj bodova za svaku komponentu (zadatak) testa, napraviti kompozitni potencijal za grupu učenika i provjeriti jesu li aposteriori zadovoljeni aksiomi egzistencije izmjerive funkcije vrijednosti. Još je jedan zahtjev kojeg treba zadovoljiti prije samog testiranja, a možda i najbitniji, je taj da uzorak učenika za testiranje treba biti dovoljno reprezentativan za učeničku populaciju na koju se test želi primjenjivati. Dodatno je pitanje, što je mjera reprezentativnosti?

9.4 Subjektivna i egzaktna procjena

U praksi se obično izbjegava agregiranje subjektivnih i egzaktnih preferencija iz jednostavnog razloga što se subjektivne preferencije ne uklapaju u tabličnu formu ulaznih podataka. Većina ljudi pokušava podatke uklopiti u tablicu i zatim računati težinsku sumu stupaca ili nešto slično tome.

Ako je subjektivna relacija preferencije potpuna i tranzitivna onda bi se mogao koristiti teorem reprezentacije 8.1 koji bi 'izmjerio' objekte i te vrijednosti mogu se ubaciti kao novi stupac tablice. Naravno, nakon toga je nužno provesti *trade off* među mjernim jedinicama stupaca i tek nakon toga može se nastaviti s agregacijom. Neke metode izbjegavaju *trade off* na način da revidiraju egzaktne vrijednosti (u stupcu) tablice i prebace ih na subjektivnu skalu, najčešće je to skala od 0–10. Dodatno se mogu uvažiti i težine varijabli (stupaca) i zatim izvršiti agregaciju računanjem težinskog prosjeka, na primjer. To su uvriježene *ad-hock* ekonomske procedure koje se naveliko koriste.

Metoda potencijala ide suprotnim putem. Svakom stupcu s egzaktnim podacima pridruži se graf preferencije i ti se grafovi agregiraju, zajedno s grafovima subjektivnih preferencija, u grupni graf čiji potencijal definira grupni poredak učenika. To je, grubo govoreći, metoda za donošenje zaključne ocjene na kraju godine na temelju svih sakupljenih informacija o učenikovim postignućima.

Primjer. Za ilustraciju rečenog promatrajmo egzaktne podatke koje su učenici A, B, C, D, E postigli na dva testa i subjektivne preferencije dobivene na temelju projektnih aktivnosti koje nije baš moguće bodovati, već su preferencije dobivene od samih učenika koji su sami vrednovali svoju aktivnost tokom izrade projekta.

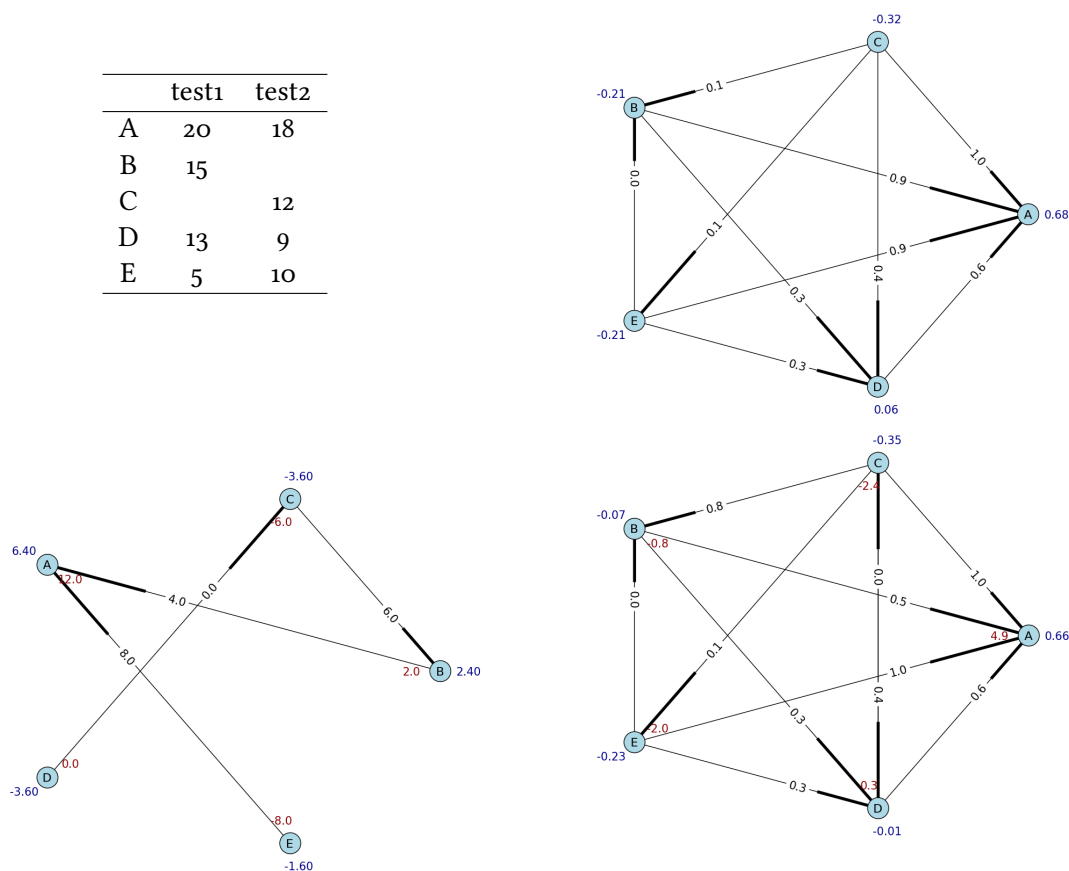
Na slici 7 dani su egzaktni podaci (tablica) i graf subjektivnih preferencija (ispod tablice). Desno od tablice je prikazan je tablični graf, a ispod njega je kompozitni graf nastao agregacijom tabličnog grafa i subjektivnog grafa. Potencijal kompozitnog grafa dan je u tablici 17. Kod određivanja

Tablica 17: Potencijal kompozitnog grafa

	potencijal
A	0.66
D	-0.01
B	-0.07
E	-0.23
C	-0.35

intenziteta preferencije u subjektivnom grafu treba paziti da budu sumjerljiv (istog reda veličine) s podacima iz tablice. U suprotnom treba svakako raditi *trade off* među mjernim jedinicama. U ovom primjeru *trade off* je napravljen na način da su maksimalne komponente tokova u tabličnom grafu i subjektivnom grafu fiksirane na vrijednost 1. Potencijal kompozitnog (agregiranog) grafa zapisan je uz same čvorove (graf na slici 7 dolje desno).

Slika 7: Egzaktni i subjektivni pokazatelji uspjeha



9.5 Zaključna ocjena kao pripadnost klasi

Ovdje ćemo dati primjer donošenja zaključne ocjene na temelju svih sakupljenih podataka o napretku učenika tokom godine. U tablici 18 vidimo bodove od 10 učenika jednog razreda sakupljenih na temelju domaćih zadaća (5), testova (4), nečeg što bi mogle biti laboratorijske vježbe (2), bodova postignuća izvučenih iz taksonomske tablice i projektnog zadatka (1). Učenici su imenovani s u_1 – u_{10} , a reci označeni s I, II, III, IV, V, VI su granice klase u koje su učenici svrstani prema njihovom uspjehu.

Primijetimo da su učenici raspoređeni u 6 klase, koje odgovaraju skali od 6 ocjena, a ne 5 kao što je sadašnja praksa u Hrvatskoj. Klase su definirane na način kako je navedeno u tablici 21. Zadovoljavajuće ocjene, ako ih tako možemo nazvati su: *dobro*, *vrlo dobro*, *odlično* i *izuzetno*, a *nedovoljno* i *loše* nisu zadovoljavajuće.

U tablici³¹ 18 neki učenici nisu prisustvovali u timskom radu izrade projekta i laboratorijskim vježbama (svjetlo sive brojke). Ti bodovi nisu uvaženi u analizi tablice. Ako ih uvažimo dolazi do minorne promjene u poretku učenika. Softver omogućava provoditi simulacije i razmatranja posljedica nakon promjene ulaznih podataka. Na primjer, kad bismo učenicima koji nisu bili sudionici u projektu ili laboratorijskim vježbama dali o bodova došlo bi do drastičnijih promjena.

U retku *norm* određena je vrijednost norme atributnog grafa kako je naznačeno. To je učinjeno s namjerom da mjera jednog boda bude ista za sve atribute. Nešto manja važnost dana je atributu *Projekt* jer 4 boda tog atributa vrijedi kao 1 bod ostalih atributa (v. redak *trade off*). Nadalje, u retku *attr-w* određene su vrijednosti težina za atribute. U principu, ako su norme grafova dobro podešene onda težine ne treba puno podešavati. Ovdje je težina atributa *DZ* ipak umanjena na $1/2$

³¹Račun je proveden uz pomoć softvera za metodu potencijala koji još nije javno dostupan jer je trenutno u fazi dorade.

Tablica 18: Zaključna ocjena agregirana metodom potencijala

	A	B	C	D	E	F	G	H
1	Edit	DZ	Testovi	Nešto	Postignuća	Projekt	X	w
2	VI max	50	100	20	100	10	0.592	0.085
3	u1	50	95	20	94	10	0.54	0.082
4	V	48	95	19	95	10	0.536	0.082
5	u3	48	88	20	90	8	0.452	0.077
6	IV	42	85	17	85	8	0.388	0.074
7	u8	48	87	12	65	5	0.316	0.07
8	u7	50	100	18	15	9	0.166	0.063
9	III	32	65	13	65	6	0.127	0.062
10	u4	38	68	18	25	5	-0.028	0.055
11	u6	13	76	13	17	6	-0.119	0.052
12	II	18	35	7	35	4	-0.245	0.048
13	u5	25	40	15	12	6	-0.249	0.048
14	u2	30	35	13	17	8	-0.26	0.047
15	u10	18	13	18	30	9	-0.366	0.044
16	I	8	15	3	15	2	-0.505	0.04
17	u9	5	8	3	8	1	-0.634	0.036
18	min	0	0	0	0	0	-0.71	0.035
19	attr_w	1	2	2	2	2		
20	norm	5	10	2	10	4		
21	tradeOff	1	1	1	1	4		

u odnosu na težine ostalih sributa. Bodovi u stupcu *Postignuća* su bodovi generirani iz učenikovih taksonomskih postignuća.

Stupci *X* i *w* su izračunati metodom potencijala, a moguće je stupac *X* dodatno reskalirati tako da bodovi budu izraženi u rasponu od 0–100. Napomenimo da metoda potencija dozvoljava da učenik postigne određenu klasu iako su neki od postignutih bodova ispod vrijednosti za tu klasu. To je slučaj učenika *u7* koji upada u klasu IV iako u stupcu *Postignuća* ima manje bodova (15) nego što je granica za tu klasu.

9.6 Važnost poznavanja skale

Medijan je jedina razumna metoda agregacije ordinalnih vrijednosti. 'Numerički jezik' koji u praksi koristimo za ocjenjivanje studenata, umjetničkog dojma, potresa ili kvalitete vina, treba biti objašnjen u svakom konkretnom slučaju. Danski obrazovni sustav, koji se prilagodio ECTS³² sustavu, dobar je primjer koji definira ocjene kao u tablici: Klasu *A* dobiva student koji je po svom uspjehu u gornjih 10% studenata. Sljedećih 25% dobiva klasu *B* itd... Ocjene 2, 4, 7, 10, 12 odabrane su tako da odražavaju, uz male korekcije, intervale (postotaka) u drugom stupcu. Tako, na primjer, interval [2, 3] (klasa *E*) je 10% duljine intervala [2, 12], interval (3, 5.5] (klasa *D*) je 25% duljine intervala [2, 12] itd...

Brojevi sami za sebe ne nude nikakvu informaciju dok se ne definira njihov smisao. Stoga ih je bolje izbjegavati, pogotovo u ocjenjivanju, jer samo mogu zbuniti. Za razliku od izborne procedure, numeričke skale kod ocjenjivanja studenata se generacijama usklađuju i dograđuju i poprimaju

³²European Credit Transfer and Accumulation System

Tablica 19: Tablica ocjena Danskog obrazovnog sustava

klasa	postotak*	ocjena	opisna ocjena
A	10%	12	odličan
B	25%	10	vrlo dobar
C	30%	7	dobar
D	25%	4	zadovoljava
E	10%	2	slab
F		0	nedovoljan

* postotak maksimalnog broja bodova potreban za ulazak u višu klasu.

osobine 'intervalne skale na mala vrata'.

Stupac *postotak* predstavlja aproksimativne vjerojatnosti normalne distribucije sredine $\text{mean}=50$ i standardne devijacije $\text{sd}=25$ za kvantile iz intervala $[0, 20]$, $[20, 40]$, $[40, 60]$, $[60, 80]$, $[80, 100]$. Točne vrijednosti su: 10, 23, 32, 24, 9 (%). Odgovarajući postotci koji proizlaze iz stupca *ocjena* su: 8, 21, 25, 21, 8 (%) što je približno vrijednostima u stupcu *postotak*.

Hrvatski obrazovni sustav, bar što se matematike tiče, nema neki unificirani prijedlog ocjenske tablice – to je ostavljeno školama da odrede same. Neke škole tu tablicu objavljuju na svojim web stranicama, za neke druge škole to je interni dogovor koji se prenosi usmenim putem. Primjer jedne takve interne tablice je dan u tablici 20 u kojoj prvi redak (*ocjena*) predstavlja numeričku vrijednost

Tablica 20: Primjer bodovne skale iz jedne hrvatske osnovne škole.

<i>ocjena</i>	2	3	4	5
<i>uzv (%)</i>	40	60	77	90
<i>rp (%)</i>	32	55	77	91

ocjene, drugi redak (*uzv*) predstavlja bodovni prag, izražen u postocima, za ulazak u višu klasu za 'usvojenost znanja i vještina', a treći redak (*rp*) predstavlja to isto ali za 'rješavanje problema'.

9.6.1 Primjer ocjenske skala

Ocjenska skala A, B, C, D, E, F, predložena u tablici 21, je samo sugestija. Ima nekoliko argumenata za takvu ocjensku skalu. Osnovu skale čini bodovna skala u rasponu 0–100 kako je naznačeno u

Tablica 21: Ocjene

klasa	naziv	bodovne skale			
		100	50	20	15
A	odlično	94	47	19	14
B	vrlo dobro	81	40	16	12
C	dobro	59	30	12	9
D	dovoljno	35	18	7	5
E	nedovoljno	16	8	3	2
F	loše	0	0	0	0

stupcu 100 te tablice. Brojevi u stupcima predstavljaju bodovni prag (u postocima) za ulazak u višu klasu. Tako na primjer, učenik koji na testu dobije 94 boda ili više dobiva ocjenu odličan, dok je 34 posto od ukupnog broja bodova (ili manje) nedovoljno za prolaznu ocjenu.

Jedan od argumenata za uvođenje takve ocjene je inflacija petica u sadašnjem sustavu u RH koja onemogućava ocjenskoj skali dovoljno razlikovanje učeničke populacije. Druga novina u skali je

uvodenje dviju ocjena koje nisu 'prolazne', to su: E i F . Ocjena F je iskaz potpunog neuspjeha, a E je ono što još 'nije dovoljno', a za što Dweck ima zgodan izraz 'not yet'.

U preostalim stupcima 50, 20, 15 izračunate su granice pojedinih klasa ovisno o maksimalnom (tom) broju bodova.

Na internetskim forumima gdje nastavnici izmjenjuju svoja iskustva naišao sam na raspravu o prolaznom pragu, tj. što se smatra granicom za *loše*. Nekima se granica prolaznosti od 45% čini velikom, a nekima je prihvatljiva granica prolaznosti od 20%. Sklon sam razmišljanju da bi to moglo ovisiti o stupnju obrazovanja, s tim da ta granica bude niža u osnovnoj školi, a visoka na sveučilištu. Netko od forumaša je komentirao da se ne bi osjećao ugodno kad bi znao da ga operira netko tko je na studiju medicine prolazio s 20% 'znanja', tj. s 80% 'neznanja'.

9.7 Post-hock analiza rezultata pismene provjere znanja

U ovom odjeljku dana je analiza rezultata pismene provjere znanja koja ima namjeru analizirati 'kvalitetu' i strukturu zadataka postavljenih na pismenom ispitu. Konkretno, radi se o zadacima koji provjeravaju znanje učenika 6. razreda jedne osnovne škole u Hrvatskoj. Analiza je ilustrativnog karaktera, a bazirana je na provjeri znanja od pedesetak učenika. Svaki zadatak donosi određen broj bodova ovisno o njegovoj složenosti i potrebnoj vještini za njegovo rješavanje.

Tekst zadataka

Z1. Poredajte po veličini od najmanjeg do najvećeg broja: $4\frac{1}{4}$, $5\cdot 4$, $\frac{27}{6}$, $4\cdot 5$, 0.45 , $\frac{14}{3}$.

Z2. Ivica ima 6 godina. Prije 4 godine Ana je bila dvostruko starija od Ivice. Koliko Ana ima godina?

Z3. Izračunaj vrijednost algebarskog izraza:

a) $4(a - 1) - (b - 1)$ ako je $a = 7$, $b = 3$.

b) Provjeri je li $y = 3$ rješenje jednadžbe $y + 3 = 3y - 3$.

Z4. 1cm na geografskoj karti predstavlja 10 000cm u prirodi. Ako su dva sela u prirodi udaljena 1km kolika je njihova udaljenost na karti?

Z5. Pojednostavi algebarske izraze:

a) $3a - 2(a - 1) + 2(a + 1) =$

b) $x - (x + 4) + 4 =$

c) Izračunaj na najkraći način: $\frac{17}{21} \cdot \frac{3}{5} - \frac{2}{3} \cdot \frac{17}{21} + \frac{17}{21} =$

Z6. Usporedi:

a) $\frac{\frac{7}{5}}{\frac{3}{18}}$ $8\frac{1}{5}$.

b) $\frac{1}{x}$ x , ako je $x = 0.01$.

Z7. Riješi jednadžbu:

a) $2x - 1 = 13$.

b) $x - 1 = 2x - 15$.

c) $\frac{1}{3} \left(\frac{x}{3} + 2 \right) = x + 1$.

Z8. Pretvorite:

$$\frac{3}{4} \text{ h} = \underline{\hspace{2cm}} \text{ min}, \quad 2\frac{1}{5} \text{ min} = \underline{\hspace{2cm}} \text{ s}$$
$$12 \text{ min} = \underline{\hspace{2cm}} \text{ h}, \quad 80 \text{ s} = \underline{\hspace{2cm}} \text{ min}$$

Z9. U jednakokrakom trokutu čije su stranice a, b, b opseg je 5 m. Stranica b je dvostruko dulja od stranice a . Kolika je duljina stranice a ?

Z10. Broj turista u Zagrebu pao je u odnosu na prethodnu godinu za 25% kada ih je bilo 250 000. Koliko turista je bilo ove godine? Koliko iznosi pad turista? Zaokružiti na znamenku tisućicu.

Z11. Maja je dobila džeparac od 200 kn. 50% potrošila je na zabavu, a 30% na slatkiše. Koliko kuna je uštedjela?

Z12. Temperatura zraka u periodu od 10 dana kretala se (mjereno u stupnjevima Celsiusa): 5, 6, 12, 13, 13, 14, 4, 3, 2, 2.

a) Nacrtajte linijski dijagram kretanja temperature (uz pomoć ravnala).

b) Navedite dane u kojima je temperatura doživjela naglu promjenu u odnosu na prethodni dan. Kolika je ta promjena?

c) Kolika je prosječna temperatura:

c1) u prva dva dana? _____

c2) u posljednja 4 dana? _____

c3) za cijeli period? _____

Kraj zadataka

U tablici 22 dana je raspodjela bodova (po zadacima) za svakog učenika (prvih 6 redaka samo).

Tablica 22: Bodovi na testu po učenicima; prvih 6 redaka

uč.	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10	Z11	Z12
1	1	3	1	2	0	2	2	0	0	0	0	2
2	2	3	4	4	2	2	4	0	0	0	0	1
3	2	3	4	3	0	2	4	0	2	0	0	0
4	1	3	2	2	2	1	2	0	0	2	0	2
5	2	3	4	5	1	1	1	0	0	3	0	6
6	1	3	2	2	0	2	4	0	0	4	4	2

Faktorska analiza tablice 22. Faktorska analiza je metoda grupiranja stupaca tablice 22 (zadataka) prema latentnim dimenzijama (faktorima) koje su 'skrivenе' u tablici podataka. Izračun faktora, u ovom slučaju, proveden je uz pomoć programskog jezika R koji je open source i postoji za sve operacijske sustave. Broj faktora predlaže korisnik, a postoje i metode koje predlažu broj faktora iz samih podataka. Ovdje je broj faktora 4.

U tablici 23 prikazana je takozvana matrica punjenja. Stupci su faktori, a reci predstavljaju zadatke. Tablica se interpretira na način da je svaki redak (zadatak) jedan vektor u 4-dimenzionalnom prostoru faktora, a brojevi u retku predstavljaju komponente tog vektora u bazi faktora. Brojevi koji su manji od 0,3 nisu prikazani iz razloga jer nas zanimaju jedna ili dvije dominantne komponente koje ukazuju na dominaciju faktora za određeni zadatak. Na primjer, zadatak Z1 i Z4 imaju dominantnu komponentu drugog faktora kojeg smo naknadno nazvali Mjere jer ti zadaci zahtijevaju uspoređivanje i pretvaranje mjernih jedinica. Zadaci Z2, Z6, Z7 su računskog karaktera i prvi faktor iz tog razloga nosi naziv Račun. Zadaci Z3, Z5 zahtijevaju sređivanje algebarskog izraza pa je drugi

Tablica 23: Matrica punjenja.

	Račun	Mjere	Alg.izraz	Form.prob
Z1		0.41		
Z2	0.69			-0.30
Z3		0.69	-0.32	
Z4		0.78		
Z5		0.43	0.49	
Z6	0.64			0.36
Z7	0.68			
Z8	-0.55	0.47		
Z9				0.61
Z10		0.32		-0.51
Z11				-0.38
Z12			0.83	

faktor prozvan Algebarski izraz. Isti faktor dominantan je i za zadatak Z12, a zašto je to tako nije sasvim jasno. Razlozi mogu biti višestruki. Taj zadatak zahtijeva prikaz podataka u kartezijevom koordinatnom sustavu i naknadno je ubačen u test. Moj je dojam da broj i struktura bodova za taj zadatak nije promišljeno učinjena jer ga je bilo teško usporediti s ostalim zadacima. Možda bi za korektniju faktorsku strukturu zadataka trebalo taj zadatak izbaciti iz analize. Zadaci Z9, Z10, Z11 su tekstualni zadaci koji provjeravaju sposobnost modeliranja i formuliranja problema i u njima je dominantan četvrti faktor koji je prozvan Formulacija problema. Taj faktor je donekle prisutan i u zadatku Z2 (opravdano) i u zadatku Z6 možda iz razloga što taj zadatak predstavlja izvjestan mentalni napor kao i ostali zadaci unutar tog faktora.

Faktorska analiza nudi još neke podatke za finiju analizu od kojih ćemo spomenuti samo *uniquenes*, pojam koji bismo mogli prevesti, u ovom kontekstu, kao *cjelovitost*. *Uniquenes* je broj pridružen

	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8	Z9	Z10	Z11	Z12
u	0.64	0.45	0.52	0.29	0.37	0.41	0.46	0.40	0.58	0.65	0.81	0.30

svakom zadatku koji se interpretira kao dekompozabilnost zadatka. Što je taj broj veći to se zadatak slabije uklapa u faktorski model – teže ga je razložiti na faktore. Takav zadatak je Z11 čiji je $u(Z_{11}) = 0.81$. Taj zadatak zahtijeva razumijevanje postotka i pažljivo čitanje teksta – traži se razlika između dobivenog i potrošenog novca. Kod ispravljanja testa uočio sam da je to bio kamen spoticanja, a mnogi učenici nisu znali kako započeti rješavanje tog zadatka. Pretpostavljam da je taj tip zadataka slabije uvježban.

Programski kod s podacima za obradu nalazi se na web adresi <https://>. Analize takvog tipa su dobrodošle za buduće preoblikovanje testa i/ili u razmatranju koje nastavne teme možda nisu dovoljno dobro obrađene.

10 Eksterna evaluacija

Nije pitanje misle li strojevi, pitanje je čine li to ljudi.

B. F. Skinner (biheviorist)

Ovaj tekst je inspiriran iskustvima provođenja državnih ispita u Engleskoj, specijalno GCSE³³. Navedene procedure u vrednovanju ispitnih rezultata odnose se na glavne koordinatore i njihovu tehničku podršku. U procesu vrednovanja ispitanika, a u svrhu donošenja opće ocjene, ispitivačke

³³General Certificate of Secondary Education

komisije trebaju apsorbirati, analizirati i usporediti milijune ocjena. Sadašnje procedure su mješavina lokalno razvijenih statističkih procedura po uzoru na psihometrijske procedure. Neke su smislene i adekvatne, a neke su nejasne i dvojbene.

Kao prvo, glavni koordinatori trebaju analizirati podatke i provjeriti jesu li instrumenti vrednovanja korišteni kako je zamišljeno i jesu li ocjenjivači poštovali zamišljenu proceduru. U tu svrhu, mnogi autori sugeriraju korištenje alata deskriptivne statistike. Drugo, glavni ispitivači trebaju prosuditi o sposobnostima svakog ispitanika i na kraju ga ocijeniti. Za ovu fazu edukativna literatura sugerira normativne mjerne tehnike, to su tehnike višekriterijskog odlučivanja, koje vode nepristranom i konzistentnom vrednovanju.

Postavlja se pitanje da li se teorija primjenjuje u praksi. U strogom smislu ne. Ipak, navedena pitanja i sugestije imaju za posljedicu da se rasprava o metodi fokusira oko budućeg praktičnog i teorijskog istraživanja posebno prema razvoju alata za potporu ispitivačima i njihovom prosuđivanju.

10.1 Procedura ocjenjivanja

Rezultat vrednovanja svake komponente je određen skup bodova — tablica profila. Nakon toga, daljnje vrednovanje se isključivo bazira na toj tablici profila. Bodovi su brojevi i tu leži opasnost u manipulaciji s njima. Oni se mogu zbrajati, može se računati srednja vrijednost, mogu biti podvrgnuti statističkoj analizi čak i ako je to posve besmisleno.

Sintezu (dodjelu ocjena na temelju profila) radi glavni koordinator sa suradnicima. To je dvofazna procedura. U prvoj fazi daje se ocjena svakom kandidatu po nekoj uhodanoj proceduri, a nakon toga se analiziraju i pregledavaju testovi onih ispitanika čiji bodovi su blizu zacrtane granice za tu ocjenu, v. Forrest (1981.) Christie and Forrest (1981). Pri tome se koriste *deskriptivne* i *normativne* procedure.

Deskriptivne procedure služe za utvrđivanje činjenica i njihovu analizu. Jedna takva deskriptivna procedura je računanje zajedničke distribucije i kovarijance među zadacima i među ispitanicima u svrhu naknadne procjene težine zadataka ili mjerenje mogućeg odstupanja ocjenjivača od dogovorenih standarda u ocjenjivanju.

Normativne procedure su posve drugačije. One nastoje mijenjati i činiti konzistentnijim vjerovanja i preferencije pojedinca. Pojedinca se pita za njegove subjektivne stavove i njegovi odgovori se analiziraju i ispituju jesu li u skladu s unaprijed utvrđenim pravilima 'racionalnog' ponašanja. Njegova vlastita inkonzistentnost i nepoštivanje pravila mu se predočuju i on sebe shvaća, sagledava i evoluira u svojim vjerovanjima i preferencijama i teži sve većoj konzistenciji.

Uzmimo za primjer, ocjenjivača koji ocjenjuje određen broj testova ispitanika. Prije svega, on utvrđuje neku grubu shemu bodovanja koja bi trebala reflektirati kvalitetu samih testova. Pretpostavimo da je ocijenio određen broj testova i, kako napreduje, on profinjuje utvrđenu shemu ocjenjivanja jer postaje svjestan finesa u proceduri ocjenjivanja. U nekom trenutku, ocjenjivač postaje nezadovoljan rezultatom svog ocjenjivanja jer smatra da je određen broj ispitanika dobio više bodova nego je 'zaslužio', a neki drugi ispitanici su dobili manje od 'zasluženog' broja bodova. On radi pauzu u ocjenjivanju i razmišlja što da učini. Moguća su dva zaključka: (1) njegova je shema ocjena neprikladna; (2) njegova holistička procjena precjenjuje neke aspekte ispitanikovog uratka koje je već nagradio ili ih je predvidio ali ih je zaboravio nagraditi. U prvom slučaju on revidira shemu ocjenjivanja, a u drugom modificira mentalne procese koje koristi u holističkoj procjeni testova. U oba slučaja prisutna je inkonzistentnost između *numeričkih reprezentacija* njegovih procjena i samih *procjena* (mentalnih procesa). Kvaliteta njegovog rada evoluira.

Jednako tako, normativna je i procedura vrednovanja ispitanikovog 'općeg uspjeha' koja u prvom redu pomaže ocjenjivaču formirati procjenu na pošten i konzistentan način što je bolje moguće.

10.2 Principi, snaga i ograničenost ljudskog prosuđivanja

Postoje neki opći principi kojima je podložno ljudsko prosuđivanje i koji imaju utjecaj kako na organizaciju javnog testiranja tako i na organizaciju ispravljanja testova sve do trenutka utvrđivanja liste profila, odnosno prije faze evaluacije i donošenja sveukupne ocjene (Greatorex, 2007).

1. princip. U procesu procjene, ocjene se donose uspoređivanjem u parovima. Apsolutne procjene se ne traže ni u kom slučaju. Laming (2004) Laming (2004) zaključuje da su sve procjene koje ljudi čine usporedbe jednog objekta s drugim i da su te procjene *nešto bolje od ordinalnih*. Osim toga, ljudski um nije u stanju držati precizan i stabilan referentni okvir u memoriji, a može razlikovati najviše pet kategorija danog kontinuuma ako nema neku dodatnu podršku. Stoga, kad ljudi donose niz uzastopnih procjena, njihove nedavne procjene su pod utjecajem prijašnjih procjena pa postaju zbunjeni i nije začuđujuće da to vodi greškama u procjenjivanju. Procjene postaju preciznije kada se objekti uspoređuju u parovima umjesto da se iznose apsolutne procjene koristeći prihvaćene standarde.

2. princip. Ocjenjivači donose odluke nezavisno jedan od drugoga. Procjene pojedinca u grupi podliježu grupnoj dinamici pa osciliraju. Ljudi često donose pogrešne procjene samo da bi se uklopili u gomilu.

3. princip. Ako je ikako moguće, smanjiti prisutnost nevažnih informacija. Na primjer, procjena kvalitete likovnog djela je pod velikim utjecajem (mogućeg) autorstva tog djela.

4. princip. Zahtijeva se adekvatno iskustvo ocjenjivača. Neposredno iskustvo koristi se kao referentna točka, a prošla iskustva čine procjenu pouzdanom i ponovljivom što je važno kod utvrđivanja granice među ocjenama.

5. princip. Ocjenjivači procjenjuju jedan zadatak (komponentu) testa. Za donošenje ukupne ocjene postoji softverska podrška. Ljudi donose preciznije i pouzdanije procjene ako se koncentriraju na komponentu umjesto na cjelinu, Dawes & Corrigan (1974) Dawes and Corrigan (1974). Na primjer, statistička kombinacija znakova bolesti koje dijagnosticira liječnik je preciznija nego njegovo sveobuhvatno mišljenje, Laming (2004) Laming (2004). Jednako tako, eksperti su dobri u pronalaženju onoga što traže ali nisu dobri u povezivanju parcijalnih informacija u cjelokupnu procjenu kvalitete ispitanika.

6. princip. Odgovori ispitanika (prikazani ispitivaču) trebaju biti u obliku koji olakšava razumijevanje ispitanikovih odgovora. Na primjer, čitanje testova na monitoru otežava njihovo razumijevanje zbog nemogućnosti prelistavanja i istovremenog gledanja svih listova (testova).

7. princip. Ocjena koju donese ocjenjivač mora se moći rekonstruirati. Drugim riječima, da bude tako formirana da ju drugi ocjenjivač može ponoviti.

Jedan od zaključaka temeljen na tim principima je da ocjenjivač ne bi smio znati kojoj školi ili grupi ispitanika pripada test kojeg ocjenjuje. Prema principu 3. to je suvišna i nevažna informacija.

10.3 Određivanje granica među ocjenama

Glavni koordinator (s ekipom) koji ocjenjuje opći uspjeh na temelju profila svakog ispitanika može krenuti u fazu određivanja granica između dvije ocjene ako je siguran da su pojedine komponente ispita korektno bodovane. To još uvijek ne znači da pojedini testovi neće biti ponovno pregledani ako se nalaze na granici između dvije ocjene.

Pretpostavimo da je svakom ispitaniku dodijeljen njegov profil s ocjenama po komponentama kao u tablici 12 ili tablici 13 i da želimo svakom ispitaniku dati opću ocjenu na ispitu u skupu {1, 2, 3, 4, 5}. Opća ocjena nastaje *agregacijom* ili *akumulacijom* parcijalnih bodova po nekom principu. Kao što smo vidjeli u tablicama 12 i 13 prosjek nije dobar postupak agregacije bodova jer nije kvantitativno smislen. Problem je u tome što ne postoji dobar i smislen algoritam koji bi profilima automatski dodijelio ocjene. Mi ćemo, radi jednostavnijeg razumijevanja cijele procedure opisati rudimentarnu proceduru koja lako može biti implementirana. Vidi također Greatorex (2009) za druge metode određivanja ocjenskih granica.

Metoda se provodi u dvije faze. U prvoj fazi se svakom profilu dodijeli ocjena u skupu {1, 2, 3, 4, 5} ili dvije susjedne ocjene ako postoji nedoumica. Tu klasifikaciju radi ocjenjivač ekspert ili neki jednostavan algoritam koji još uvijek nije dovoljno inteligentan da razgraniči spada li pojedini test ispitanika u kategoriju (klasu) 'čiste ocjene'. Testovi kojima pripadaju dvije ocjene spadaju u tzv. *granične* testove. U drugoj fazi se granični testovi rangiraju na temelju uspoređivanja u parovima nekom od metoda. Proceduru uspoređivanja po parovima može učiniti jedan ekspert ili grupa njih. U slučaju grupne odluke rangiranje je dano konsenzusom grupe (Čaklović, 2005). Nakon toga se skup graničnih testova podijeli u gornju i donju klasu koje se pridruže gornjoj i donjoj 'čistoj ocjeni'.

U predloženoj proceduri, kriterij prema kojem se testovi uspoređuju u parovima je *opći dojam*, dakle jedan kriterij. Tehnički nije teško zakomplicirati proceduru na način da se pojedinim komponentama daju težine i rangiranje provede u odnosu na komponente ispita kao kriterije. Pitanje je koliko je to smisleno, no to je stvar odluke glavnog koordinatora i njegovih pomagača.

10.3.1 Neke nove/stare sugestije

Postoji li uopće alternativa sadašnjoj praksi? Pollitt and Elliott (2003a,b) sugeriraju da ocjenjivanje, ovakvo kakvo je ovdje opisano, može biti zamijenjeno 'širokopojsnim' uspoređivanjem u parovima. Takva razmišljanja postoje već neko vrijeme u literaturi ali je nejasno kako to sprovesti. Uspoređivanje u parovima je zamorno i dugotrajno. Takve bi procedure trebalo moderirati i standardizirati (za interno ocjenjivanje) i standardizirati (za vanjsko ocjenjivanje). Oni sugeriraju da se u takvo uspoređivanje u parovima uključe i testovi ranijih generacija radi održavanja nivoa i kvalitete ocjenjivanja umjesto da se koristi slabo kriterijsko referenciranje.

Kod elektroničkog bodovanja, statističke procedure mogu biti korištene za naknadno određivanje granica među ocjenama. Na primjer, za svako zasebno pitanje potrebno je odrediti granicu za svaku ocjenu i zatim sve te granice uklopiti u granicu za ispitanika. Druga mogućnost je da se pitanja sortiraju po težini i da se odredi koje je najteže pitanje, za svaku ocjenu, na koje ispitanik treba korektno odgovoriti. Prednost takvog pristupa je da ocjenjivači eksperti ne trebaju procjenjivati težinu pitanja, što inače (kako pokazuju istraživanja) rade loše. Mana takvog pristupa je da ne dozvoljava kompenzaciju.

10.4 Državna matura u Hrvatskoj

Čini se da državna matura u Hrvatskoj proživljava 'dječje bolesti' državnih matura zemalja za koje se može reći da imaju više iskustava. Osnovni problem u svakoj zemlji izgleda da je nesporazum između zakonodavca i obrazovne struke. Zakonodavac nema jasne i transparentne namjere a struka nema jasne definicije i stavove. Postoji samo ogromno iskustvo. Dva su otvorena pitanja:

1. Što se maturom mjeri?
2. Tko, zašto i na koji način koristi te informacije?

Pretraživanjem hrvatskog web prostora naišao sam na svega par javnih informacija o državnoj maturi. Jedna je *Pravilnik o polaganju državne mature*, a druga je komentar jednog člana *Povjerenstva za provedbu mature* u javnim glasilima, koji potvrđuje ne samo neslaganje između države i struke nego i nerazumijevanje i/ili nepoznavanje sadašnjeg trenutka u znanstvenoj i obrazovnoj literaturi u svijetu.

10.4.1 Iz pravilnika o polaganju državne mature, NN 87/08

Članak 2. *Cilj je državne mature provjera i vrjednovanje postignutih znanja i sposobnosti učenika, stečenih obrazovanjem prema propisanim općeobrazovnim nastavnim planovima i programima.*

Članak 10. *Ispitni sadržaji te način provjere i ocjenjivanja znanja i sposobnosti na ispitima uređuju se ispitnim katalozima prema nastavnim planovima i programima iz općeobrazovnih predmeta.*

Koliko je meni poznato katalozi (iz matematike) daju primjere i način bodovanja pojedinih zadataka. Nigdje se ne govori o zaključivanju ocjene na temelju profila.

10.4.2 Iz javnih glasila

Pitanje: Što će se promijeniti uvođenjem državne mature?

- Pa prvo više neće biti paušalnih dijeljenja ocjena jer je matura zapravo postupak vanjskog vrednovanja, možemo reći nekakav mjerni uređaj koji će objektivno procijeniti nečije znanje, ali i školu koju je netko pohađao te nastavnike. Rezultati mature zapravo će biti zlatni rudnik informacija.

Vedran Mornar (bivši član povjerenstva za provedbu mature, *Večernji list*, 17. 01. 2010.)

10.4.3 Eksterna evaluacija nužnost ili zabluda

Naslov je polemički i nakon svega napisanog možda bi bolje promijeniti ga u Eksterna evaluacija, nužnost i zabluda. Riječ 'nužnost' zadržavam jer smatram da eksterna evaluacija na dužu stazu vodi ka kvalitetnijem obrazovnom sustavu, a akcentiram 'zabludu' jer zakonodavac, prema riječima spomenutog člana povjerenstva, pokušava riješiti jednadžbu

evaluacija rezultata maturalnog ispita = evaluacija školstva.

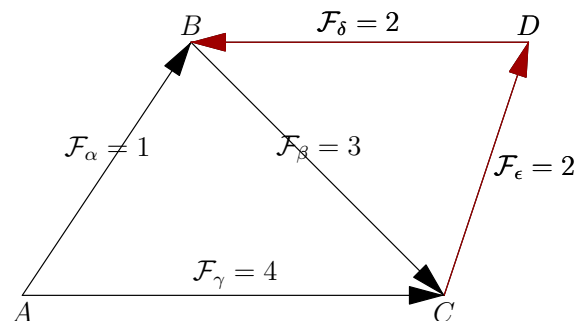
U zabludi su i profesori u srednjim školama koji školske aktivnosti usmjeravaju ka 'uspjehu na maturi' umjesto poticanju znanja i sposobnosti.

11 Dodatak

11.1 Metoda potencijala

Metoda potencijala sjedinjuje fleksibilnost uspoređivanja u parovima i konstrukciju izmjerive funkcije vrijednosti. Umjesto da izgrađuje preferencije na zamjenama, donositelju odluke nudi se mogućnost iskazati to isticanjem *intenziteta preferencije* za zadani par alternativa kao što je to prikazano ne slici 8. Slika pokazuje *graf preferencije* koji nije potpun jer alternative A i D nisu uspoređene.

Slika 8: Primjer grafa preferencije



Umjesto da govorimo o paru alternativa (B, A) mi ćemo govoriti o strelici (luku) koja izlazi iz A i ulazi u B. Odgovarajući intenzitet preferencije izražavamo brojem F_α koji na slici iznosi 1. Vrijednost $F_\alpha = 0$ izražavala bi činjenicu da donositelj odluke jednako preferira te dvije opcije. Katkada ćemo luk α s odgovarajućom težinom s označavati kao $B \stackrel{s}{\leftarrow} A$.

Intenzitet preferencije je nenegativan³⁴ broj kojeg donositelj odluke pridjeljuje preferenciji na nekoj skali u njegovoj mentalnoj slici. Za subjektivne preferencije uvriježeno je davati intenzitet u skupu $\{0, 1, 2, 3, 4\}$, što odgovara verbalnom intenzitetu: *jednaka, slaba, jaka, izrazita* i *apsolutna* preferencija, ali nisu isključeni niti drugačiji rasponi vrijednosti. Iskustvo pokazuje da je korisno povećati raspon skale intenziteta ako je broj alternativa veći.

Dakle, svakom luku $\alpha \in \mathcal{A}$ iz skupa lukova \mathcal{A} donositelj odluke pridjeljuje nenegativan broj $\mathcal{F}_\alpha \geq 0$. Funkciju \mathcal{F} definiranu na skupu lukova nazivamo *tokom preferencije* ili kraće *tokom*.

11.1.1 Konzistentnost toka.

Za razliku od preciznosti, koja je smisljena kod mjerenja s mjernom jedinicom, ovdje ima smisla *konzistentnost* ulaznih podataka. Podgraf (na slici 8) određen lukovima α, β, γ je konzistentan jer je $\mathcal{F}_\alpha + \mathcal{F}_\beta = \mathcal{F}_\gamma$, dok podgraf određen lukovima β, ϵ, δ nije konzistentan jer je $\mathcal{F}_\epsilon + \mathcal{F}_\delta + \mathcal{F}_\gamma = 7 \neq 0$. Tok je konzistentan ako je za svaki ciklus (zatvoren put) suma intenziteta preferencije duž lukova tog ciklusa jednaka nuli.

Algebarski se to može izraziti i postojanjem funkcije $X : V \rightarrow \mathbb{R}$ takve da je za svaki luk $\alpha = (B, A)$

$$X(B) - X(A) = \mathcal{F}_\alpha. \quad (8)$$

U sustavu jednadžbi (8) imamo toliko jednadžbi koliko ima lukova u grafu. Donositelj odluke ne mora nužno dati konzistentan graf. Matematika kaže da bez obzira na konzistentnost toka, sljedeća jednadžba

$$A^T A \mathcal{F} = A^T \mathcal{F} \quad (9)$$

uvijek ima rješenje X i to rješenje predstavlja skalu s vrijednostima X_A, X_B, X_C, X_D na koju su smješteni objekti. Funkciju X nazivamo *potencijalom*, a metodu koja konstruira tu skalu nazivamo *metodom potencijala* (L. Čaklović, A. S. Kurdija (2017) *A universal voting system based on the Potential Method*).

Metoda potencijala omogućava mjerenje bilo kakvih objekata čija je međusobna 'udaljenost' izražena putem grafa preferencije.

Teorem 11.1. *Za potpuni tok vrijedi formula*

$$X = \frac{1}{n} A^T \mathcal{F}. \quad (10)$$

Dokaz. Matrica $A^T A$ je simetrična matrica na čijoj dijagonali je broj $n-1$, a izvan dijagonale stoje -1 . Iskoristimo li uvjet $\sum_i X_i = 0$ i dodamo ga svakom retku jednadžbe (9), dobije se gornja formula. \square

Korisnost formule (10) je u tome što se potpuno zaobilazi rješavanje sustava $A^T A X = A^T \mathcal{F}$.

11.2 Agregacija grafova preferencije

Osnovna procedura kod grupnog i višekriterijskog odlučivanja je ujedinjenje pojedinačnih grafova (tokova) u jedan jedini koji predstavlja grupnu preferenciju. Pretpostavimo da je u problemu prisutno više kriterija $C_i, i \in \{1, \dots, k\}$. Svaki od njih generira svoj vlastiti graf preferencije $G_i = (V, \mathcal{A}_i)$ i vlastiti tok \mathcal{F}_i . U proceduri agregacije potrebno je poznavati težine w_i svakog individualnog kriterija C_i (toka \mathcal{F}_i).

U prvom koraku, za svaki par $\alpha = (u, v)$, koji može i ne mora biti luk grafa G_i , računamo

$$F_\alpha := \sum_{\substack{i=1 \\ \pm \alpha \in \mathcal{A}_i}}^k w_i \mathcal{F}_i(\alpha) \quad (11)$$

³⁴Broj koji nije negativan. Može biti jednak 0 ili pozitivan.

gdje doprinos $w_i \mathcal{F}_i(\alpha)$ uvažavamo samo ako je $\alpha \in \mathcal{A}_i$ ili $-\alpha \in \mathcal{A}_i$, pri čemu je $\mathcal{F}_i(\alpha)$ negativan broj ako je suprotan luk $-\alpha = (v, u) \in \mathcal{A}_i$. Ako je suma u (11) nenegativna, tada par α proglašavamo lukom u grupnom grafu i definiramo vrijednost grupnog toka \mathcal{F} na luku α kao

$$\mathcal{F}(\alpha) := F_\alpha.$$

Ako je suma u (11) negativna, tada je $-\alpha := (v, u)$ luk grupnog grafa (mijenjamo njegovu orijentaciju) i stavljamo

$$\mathcal{F}(-\alpha) := -F_\alpha.$$

Ako F_α nije definiran, onda u i v nisu spojeni u grupnom grafu. Tok \mathcal{F} je sada dobro definiran i nazivamo ga *kompozitnim tokom*. Zbog linearne veze između toka i potencijala odgovarajući *kompozitni potencijal* X jednak je linearnoj kombinaciji odgovarajućih individualnih potencijala. Dakle, vrijede formule

$$\mathcal{F} = \sum_{i=1}^n w_i \mathcal{F}_i, \quad (12)$$

$$X = \sum_{i=1}^n w_i X_i. \quad (13)$$

11.3 Potencijal tablice odlučivanja

U tablici odlučivanja kompozitni tok, definiran na grafu čiji su vrhovi akcije a_1, \dots, a_m , je prema formuli (11)

$$\mathcal{F}_{kj} = \sum_i P(\theta_i)(v_{ki} - v_{ji}), \quad k, j = 1, \dots, m \text{ i } k \neq j, \quad (14)$$

gdje je \mathcal{F}_{kj} komponenta toka na luku (a_k, a_j) ; dakle sumira se atributima. U slučaju da je neki od podataka u tablici nepoznat, npr. v_{ki} ili v_{ji} , tada u sumi (14) zanemarimo i -ti sumand.

Sljedeći teorem pokazuje da su metoda potencijala i kriterij očekivane vrijednosti ekvivalentni, ne samo što se rangiranja tiče, već potencijal i očekivana vrijednost U

$$U_i = U(a_i) := \sum_j P(\theta_j)v_{ij}$$

daju iste (do na konstantu) numeričke vrijednosti. To je svakako zanimljiv rezultat jer u slučaju nepotpune tablice očekivanu vrijednost nije moguće izračunati dok potencijal i dalje postoji.

Teorem 11.2. *Neka je \mathcal{F} tok izračunat po formuli (14) i X odgovarajući potencijal. Tada su očekivana vrijednost U i potencijal X vezani formulom*

$$X_k = U_k - \frac{1}{m} \sum_j U_j, \quad (15)$$

tj. razlikuju se do na konstantu.

Dokaz. Iz formula (14) i (10) je

$$X_k = \frac{1}{m} \sum_j \mathcal{F}_{kj} = \frac{1}{m} (U_k - U_j) = U_k - \frac{1}{m} \sum_j U_j.$$

□

Teorem 11.2 također vrijedi i u slučaju višeatributnog odlučivanja u kojem su atributi preferencijalno nezavisni.

Napomena. U formuli (14) za kompozitni tok tablice figuriraju vjerojatnosti $P(\theta_i)$ atributa ($i = 1, \dots, n$). Postavlja se pitanje kako odrediti težinu atributa ako vjerojatnosti nisu poznate. U principu, određivanje težina je složena procedura i jedan od načina je odrediti trade-off među mjernim

jedinicama atributa. Ako mjerne jedinice nisu dane, napr. ako su vrijednosti jednog atributa izraženi na ordinalnoj skali, moguće je iskoristiti metodu potencijala na način da se konstruira subjektivni graf preferencija među atributima i izračuna potencijal X . Poznavanje potencijala X još uvijek ne znači poznavanje težina jer su neke vrijednosti potencijala negativne. Bez velike argumentacije, težine atributa se računaju po formuli

$$w = \frac{a^X}{|a^X|_1}, a > 1, \quad (16)$$

i normirane su na način da im je suma jednaka 1. Broj $a > 1$ određuje brzinu rasta eksponencijalne funkcije i modulira odnos među težinama (trade-off).

Alternativna mogućnost je da su atributi jednake težine, ali se njihovi grafovi preferencije normiraju na način tako da se maksimalna vrijednost toka preferencije postavi na neku vrijednost ovisno o atributu. Tu vrijednost nazvat ćemo *norma* grafa. Nakon renormiranja tablice na taj način prilazi se računanju kompozitnog toka tablice po formuli (14) uvažavajući zahtjev da je $P(\theta_i) = 1, i = 1, \dots, n$.

11.3.1 Dualnost u ocjenjivanju

Jedna od primjena metode potencijala u kojoj dolazi do izražaja njena fleksibilnost je agregacija ocjena. Gotovo isti pristup koristi se u kontekstu strojnog učenja i umjetne inteligencije što nadilazi ciljeve ove knjige. U samoj jezgri tih problema nalazi se algoritam koji detaljizira Helmholtzovu ideju da je mozak stroj za predviđanje i da revidira svoje iskustvo sadržano u memoriji na temelju ulaznih podataka (trenutne percepcije).

Metoda potencijala ima dvije novine koje joj omogućavaju tu fleksibilnost. To su: (1) *operator adaptacije* i (2) mogućnost uvažavanja *sinergije* bez eksplicitne definicije što bi sinergija trebala biti. Sinergija i adaptacija ne idu jedno bez drugog i teško ih je separirati jednako kao što je teško separirati prostor i vrijeme.

Samoprocjena. Adaptacija. Adaptacija je sposobnost sustava da procijeni stanje vlastitih parametara na temelju promijenjene okoline i svog vlastitog ustrojstva.

Svaki sustav (čovjek) ragira na okolinu koja mu nudi vrijednosti polaznih parametara ili mu određuje dinamiku promjene tih parametara. Operator adaptacije revidira, uvažavajući strukturu sustava, vrijednosti polaznih parametara i izračunava nove vrijednosti koje sustav koristi u procjeni vlastitog stanja i omogućava mu reakciju na promijenjeno stanje okoline.

Adaptacija nosi u sebi dualitet u odnosu između polaznih parametara i mogućih reakcija, tj. odgovora sustava na promjenu parametara. Tehnički termin za polazne parametre je *atributi*, a za moguće odgovore *alternative*. Evo konkretnog primjera.

Sinergija. Jedan od aspekata sinergije je i sinkronicitet (koherentnost) kako je to nazvao Jung. To je lijepa zamisao ali je s algoritamskog gledišta neupotrebljiva. Mi u svakodnevnom životu konstatiramo sinkronicitet ali na to gledamo kao na slučajnost ili na čudo.

U gotovo svim današnjim modelima učenja koristi se statistika i vjerojatnostni modeli koji stoje u njenoj pozadini. Vjerojatnost je također jedna funkcija vrijednosti u koju je ugrađeno *svojstvo aditivnosti* događaja, slično nezavisnosti o irelevantnoj alternativni. U najkraćim crtama, to svojstvo zahtijeva da ako događajima A, B pridružimo njima disjunktni događaj C onda odnos vjerojatnosti od $A \cup C$ i $B \cup C$ ostaje nepromijenjen. Takvim procedurama je u samom startu otežana mogućnost da modeliraju i objasne one pojave u kojima se pojavljuje sinergija. Sinergija je pojava (konstelacija) koja drastično mijenja stanje svijeta ako je prisutna i teško ju je definirati u operativnom smislu.

Jednostavan primjer sinergije je detekcija bolesti pomoću simptoma. Za neke bolesti, ako se određeni simptomi pojave zajedno, iskusan liječnik će bez ikakvih laboratorijskih nalaza odrediti o kojoj se bolesti radi na temelju vlastitog iskustva.

Metoda potencijala omogućava uvažavanje sinergije ili na način da ju procjenjuju ljudi ili na način da proizlazi iz samih podataka. U kontekstu strojnog učenja metoda potencijala koristi operator samoprocjene (adaptacije). Detalje ostavljamo za neku drugu priliku.

11.4 Ocjenjivanje učenika. Primjer

U tablici 14 dane su ocjene učenika na ispitu iz Fizike, Tjelesnog i Muzičkog, a znak '*' znači

Tablica 24: Ocjene učenika na završnom ispitu.

	Fizika	Tjelesni	Muzički
Ivica	3	5	4
Marica	*	1	3
Trsek	2	*	2
Grozdek	4	3	*

da učenik nije ocijenjen iz odgovarajućeg predmeta. Predmeti učenja su *atributi*, a učenici su *alternativa*.

Pitanje koje se postavlja je možemo li na temelju tih podataka rangirati učenike po uspješnosti? Takvo pitanje je nametnuto civilizacijskim dostignućem ili, da budem sarkastičan, nametnutim ponašanjem koje nam nudi naš obrazovni sustav.

11.4.1 Kompozitni tok tablice

Za sada nije preciziran nikakav odnos između atributa i alternativa, ali ako odlučimo rangirati učenike po uspješnosti tada treba precizirati važnost atributa. U ovom slučaju pretpostavljamo da predmeti podučavanja imaju podjednaku važnost jer bi se u suprotnom neki profesori mogli uvrijediti. Primijetimo da je tablica ocjena nepotpuna i kao takva nije pogodna kao input za sve do sada spominjane metode bez nekih dodatnih friziranja ulaznih podataka.

Kompozitni tok tablice nastaje agregacijom tokova za svaki predmet, tj. toka \mathcal{F}_F za fiziku, toka \mathcal{F}_T za tjelesni i toka \mathcal{F}_M za muzički. Svaki od tih tokova ima iste čvorove, a to su alternative (učenici) koje ćemo jednostavnosti radi označiti s I, M, T, G . Lukovi toka \mathcal{F}_F su: $\{I \xleftarrow{1} T, G \xleftarrow{1} I, G \xleftarrow{2} T\}$, lukovi toka \mathcal{F}_T su: $\{I \xleftarrow{4} M, I \xleftarrow{2} G, G \xleftarrow{2} M\}$, a lukovi toka \mathcal{F}_M su: $\{I \xleftarrow{1} M, I \xleftarrow{2} T, M \xleftarrow{1} T\}$. Agregirani graf ima vrhove I, M, T, G a njegove lukove čini unija svih lukova od atributnih tokova. Taj kompozitni graf ima paralelne lukove. Postoje tri načina kako analizirati taj graf: (1) kao multigraf u kojem paralelni lukovi ostaju, (2) paralelne lukove zamijeniti jednim koji ima intenzitet jednak **sumi** paralelnih lukova i (3) paralelne lukove zamijeniti jednim koji ima intenzitet jednak **srednjoj vrijednosti** paralelnih lukova. Svaka od navedenih metoda ima za i protiv, što nadilazi ciljeve ovog teksta. Mi ćemo uzetu sumu.

Rješenje Laplaceove jednadžbe dobivenog kompozitnog grafa (9) daje potencijal prikazan u tablici 25

Tablica 25: Rangiranje učenika iz tablice 24 na temelju jednako važnih atributa.

Učenici	Ivica	Grozdek	Trsek	Marica
potencijal	2.25	0.75	-1.5	-1.5
težina	0.665688	0.235356	0.0494776	0.0494776

11.4.2 Samoprocjena

A što ako odlučimo rangirati predmete na temelju ocjena (mjerenja) konkretnih učenika? To je jednako moguće, ali koji je smisao toga? Takav postupak bi se mogao interpretirati kao rangiranje *atributa* (predmeta) po stupnju ovladavanja gradivom. Što veća vrijednost na rang listi to je gradivo bolje savladano.

Ako *attribute* shvatimo kao primarne onda smo u stanju 'mjeriti' alternative, a atribut ovdje ima ulogu metra kojim mjerimo učenika. Ako *alternative* shvatimo kao primarne onda baždaramo metre (*attribute*) na konkretnim objektima mjerenja. U svakom od spomenutih postupaka (mjerenje i baždarenje) potrebno je unaprijed zadati važnost (težinu) *atributa* odnosno *alternativa*. Mjerenje ćemo nazvati *primarnim*, a baždarenje *dualnim* postupkom.

Postoji i treća mogućnost kako iskoristiti podatke u tablici, a ta je da provedemo primarni postupak, za neke početne vrijednosti težina atributa, izračunamo težine alternativa i provedemo dualni postupak. Zatim ponovimo primarni postupak s novim, promijenjenim težinama, ponovimo dualni postupak i tako dalje. Osnovni korak u toj proceduri, a to je jedan primarni i jedan dualni postupak, nazivamo *operatorom samoprocjene*. To je adaptacija o kojoj smo govorili u početku.

Može se dokazati³⁵ da operator samoprocjene ima fiksnu točku a to su vrijednosti težina atributa, koje ostaju nepromijenjene prilikom samoprocjene. Te parametre interpretiramo kao važnost *atributa* i *alternativa* koje tablica nosi u sebi. Za podatke iz tablice 24 fiksna točka operatora samoprocjene je:

Muzički	Tjelesni	Fizika	Ivica	Grozdek	Marica	Trsek
0.339528	0.330731	0.32974	0.264151	0.261712	0.240225	0.233911

Prema rezultatu samoprocjene ispada da je Muzički, posmatrajući s gledišta ove grupe učenika, bolje ovladan od drugih predmeta. Velika razlika u težinama nastavnih predmeta ukazivala bi na problem čije uzroke treba tražiti van ove tablice.

11.5 Ordinalno rangiranje s metodom potencijala

Za razliku od drugih, metoda potencijala kao input koristi intenzitet preferencije što je u slučaju tablice 24 razlika ocjena. Ordinalnost, kao pristup, zaboravlja intenzitet preferencije i uvažava samo prioritet. Jedan od načina da se to uradi je taj da se svi intenziteti preferencije u kompozitnom grafu normiraju na 1. Ovdje je to učinjeno na način da je kompozitni graf dobiven sumacijom, a zatim

Tablica 26: Ordinalno rangiranje učenika iz tablice 24 na temelju jednako važnih atributa.

Učenici	Ivica	Grozdek	Marica	Trsek
potencijal	3.75	1.25	-1.25	-3.75
težina	0.824028	0.145669	0.0257509	0.00455215

su svi intenziteti preferencije normirani na 1. Razlog tome je što normiranjem toka prije računanja kompozitnog grafa može dovesti do situacije da se u kompozitnom grafu dokinu intenziteti na 0. U tom smislu ordinalnost metode potencijala, što se analize tablice tiče, je 'kontrolirana'.

Posljedica toga je rangiranje učenika dano u tablici 26. U usporedbi s tablicom 25 vidimo da ordinalno rangiranje daje prednost Marici o odnosu na Trseka. Ostali odnosu na rang-listi ostaju nepromijenjeni.

³⁵Čaklović, L. (2011). *Conflict Resolution. Risk-As-Feelings Hypothesis*. Labsi Working Papers, (35):1-16. (<http://www.labsi.org/wp/labsi35.pdf>).

11.6 Agregacija unutar hijerarhije

O hijerarhiji i njejoj strukturi već je rečeno u odjeljku 5, a aksiomatika hijerarhije opisana je u (Čaklović, 2014, str. 454).

Vrednovanje elemenata u hijerarhiji odvija se silazno po razinama. Cilj dobija težinu 1. Svaki element odluke koji je već vrednovan, tj. ima težinu, prenosi svoju težinu na svoje listove po nekom pravilu, a suma težina listova jednaka je težini korijena. Ako su svi elementi neke razine vrednovani onda je njihova ukupna težina jednaka težini njihovih korijena. Na dnu hijerarhije su alternative i suma težina svih alternativa jednaka je 1. Vrednovanje u hijerarhiji može se shvatiti kao prenošenje cjelovite informacije od cilja (težina cilja 1), preko svih razina hijerarhije, na alternative, ali tako da nema gubitaka informacije.

Ako ste upoznati s linearnom regresijom onda bismo mogli usprediti nju i metodu potencijala³⁶, te hijerarhijski linearni model i vrednovanje hijerarhije pomoću metode potencijala. Hijerarhija je prirodna struktura koja unosi nelinearnost u agregaciju grafova preferencije.

11.6.1 Zavisnost ocjena

U primjeru koji slijedi analizira se tablica 8 koju ovdje ponovno prepisujemo:

	<i>F</i>	<i>M</i>	<i>E</i>	<i>Total</i>
<i>a</i>	18	12	6	36
<i>b</i>	18	7	11	36
<i>c</i>	5	17	8	30
<i>d</i>	5	12	13	30

U ovom problemu je prirodno ugrađena sljedeća hijerarhija. U prvu, ciljnu razinu, stavljamo cilj kojeg ovdje nazivamo "ocjena", u drugu razinu stavljamo dvije klase učenika nazovimo ih: *AB* i *CD* jer grupa učenika $\{a, b\}$ očito dominira grupu $\{c, d\}$. Klasa *AB* kao svoje listove ima učenike *a, b*, a klasa *CD* kao listove ima učenike *c, d*. Potpuna hijerarhija je:

razine:{"ciljna","grupe", "pojedinci"}
elementi razina:{"ciljna":{"ocjena"}, "grupe": {*AB, CD*}, "pojedinci":{*a, b, c, d*}}
listovi:{"ocjena": {*AB, CD*}, *AB*:{*a, b*}, *CD*:{*c, d*}}

Grupa *AB* je bolja od grupe *CD*, tako da bez neke velike torije grupi *AB* možemo dati težini 0.7, a grupi *CD* težinu 0.3. Unutar grupe *AB* student *a* je nešto bolji od studenta *b* pa bez neke velike teorije možemo im dati težine (unutar grupe) 0.6 i 0.4 respektivno. Isto tako je student *c* nešto lošiji od studenta *d*, jer ima dvije ocjene ispod praga od 10, pa im možemo dati težine 0.35 i 0.65 respektivno. Prema pravilu o prenošenju težina u hijerarhiju težine učenika su:

$$w(a) = 0.7 * 0.6 = 0.42, \quad w(b) = 0.7 * 0.4 = 0.28, \\ w(c) = 0.3 * 0.35 = 0.105, \quad w(d) = 0.3 * 0.65 = 0.195.$$

Dakle, $a > b > d > c$ što je isti rezultat kao u ranijoj analizi tablice na str. 32. Ovo razmišljanje se u suštini ne razlikuje od ranije analize iste tablice. U ranijoj analizi, hijerarhija je implicitno uvažena, a ovdje eksplicitno.

11.7 Samodualna hijerarhija

Samodualna hijerarhija je organizirana tako da su elementi odlučivanja u prvoj razini i posljednjoj razini jednaki. Procedura vrednovanja hijerarhije je iterativna i odvija se na sljedeći način:

³⁶U principu se radi o istom tipu minimizacije kvadratičnog funkcionala s tom razlikom što u linearnoj regresiji matricu *A* 'čitamo' iz podataka, a kod metode potencijala to je matrica incidencije grafa preferencije. U linearnoj regresiji desna strana jednadžbe $AX = Y$ je eksplanatorna varijabla, a kod metode potencijala to je tok preferencije.

1. Težine elemenata prve razine se odrede proizvoljno, najčešće se određuju tako da su sve međusobno jednake i normirane tako da im je suma jednaka 1. Vektor težina označimo s w_0 .
2. Sljedeća razina se rangira na način kako je to opisano u odjeljku 11.6 i nastavlja se dalje po nižim razinama sve dok ne bude vrednovana i posljednja razina. Dobivene težine w_1 posljednje razine ovise o početnim težinama, a funkciju Φ koja ih povezuje nazivamo *operatorom samorangiranja*; dakle $w_1 = \Phi(w_0)$.
3. Težine w_1 supstituiraju se umjesto početnih težina w_0 i ponovi se postupak rangiranja s novim revidiranim težinama.

Rezultat iterativnog postupka je niz $w_n = \Phi^n(w_0)$, $n \in \mathbb{N}$ za koji se pokazuje da konvergira uz vrlo blage pretpostavke na preferencije dane u hijerarhiji. Dokaz te činjenice nadilazi matematičku razinu ovog teksta, a može se naći u još neobjavljenom radu Čaklović (2011). U spomenutom članku pokazano je također kako se emocije također mogu uvući kao element odlučivanja ali ne u običnu nego u samodualnu hijerarhiju.

Samodualna hijerarhija je zapravo postupak revizije polaznih težina u sustavu s povratnim vezama. Ako ste upoznati s revizijom vjerojatnosti i Bayesovom formulom u teoriji vjerojatnosti onda je operator samorangiranja Φ njena dobra analogija. Detaljnije o toj vezi može se naći u (Čaklović, 2014, str. 438) u odjeljku *Samorangiranje i revizija prioriteta*.

Kao primjer samorangiranja dat ćemo reviziju državnog proračuna i komunikaciju inteligentnih robota-vozila.

11.8 Revizija državnog proračuna

Ovo je vrlo primitivan model rebalansa državnog proračuna koji se može modificirati i nadograđivati novim elementima odlučivanja.

Pretpostavimo da je vlada države *Južno Veliko More* odlučila reprogramirati javno financiranje zbog smanjenog priljeva novca u državnu blagajnu. Na sjednici vlade je zaključeno da se porezni obveznici dodatno ne opterećuju jer je opći dojam da bi to moglo poremetiti ionako slabo razvijeno tržište kapitala. Premijer želi da se ministarstva sama izjasne o svojim potrebama i zatražio je od

Tablica 27: Samorangiranje. Preferencije ministarstava.

	TU	VO	IN	ZN	TR
TR-VO	-4		2		-1
IN-TU	-2	-3	-2	0	2
TR-IN	-1	2		3	
TR-TU	-2		1		2
IN-VO		-3		-2	
VO-TU		0		2	
ZN-TR	3	0		-2	2
ZN-TU		-2	-1	-1	
ZN-VO			-2		2
ZN-IN				-1	3

ministara *transporta* (TR), *industrije* (IN), *turizma* (TU), *vojske* (VO) i *znanosti* (ZN) da izraze svoje potrebe u formi grafa preferencije. Grupa donositelja odluke je $G = \{TR, IN, TU, VO, ZN\}$, a njihove preferencije dane su u tablici 27. Iz tablice vidimo da TU daje prednost TR u usporedbi TR-VO i

intenzitet te preferencije je 4³⁷. U istoj usporedbi IN preferira VO s intenzitetom 2, dok VO i ZN taj par nisu uspoređivali. Niti jedan individualni graf preferencije nije potpun. Fiksna točka preslikavanja samorangiranja Φ izračunata je u tablici 28.

Postoji i mogućnost analize bliskosti ministarstava prema njihovim preferencijama. Detalji se mogu naći u članku *Graph Distance in Multicriteria Decision Making Context* Čaklović (2003) i na-

Tablica 28: Fiksna točka državnog proračuna.

Ministarstvo	rang	potencijal
IN	0.230	0.210
ZN	0.206	0.051
TR	0.201	0.017
TU	0.189	-0.076
VO	0.173	-0.202

dilaze ambicije ovog teksta.

Ako ministar pokazuje dodatnu mudrost, on može iz razmatranja izbaciti preferencije u kojima ministarstva uspoređuju svoje potrebe s potrebama drugih ministarstava i promatrati stabilnost odluke. Ako ima posla s poštenim ministarstvima koji uvažavaju interese svih u njegovoj zemlji, a ne samo svoju individualnu korist, onda se ta nova odluka dobivena samorangiranjem ne bi smjela drastično razlikovati od prethodne.

11.9 Inteligentni mobilni roboti

Ovo je primjer konfliktne situacije u kojoj se nalaze dva inteligentna mobilna robota (vozila), nazvat ćemo ih R_1 i R_2 , koja istovremeno pokušavaju proći uskim prolazom. Da bi izbjegli sudar, roboti mogu:

skrenuti desno (R), skrenuti lijevo (L), čekati (W).

Ovdje 'inteligencija' znači da roboti imaju ugrađene preferencije između navedenih akcija, ovisno o situaciji. Na primjer, ako prvi robot čeka (W), tada drugi robot preferira kretanje udesno (R) i ulijevo (L) u usporedbi s čekanjem (W) i preferira kretanje udesno (R) nasuprot kretanju ulijevo (L), drugim riječima preferira se kretanje nasuprot čekanju i preferira se vožnja desnom stranom.

Preciznije, pretpostavimo da prvi robot najavljuje skretanje ulijevo (L). Odgovor drugog robota je strukturiran tokom preferencija \mathcal{F}^L :

$$\mathcal{F}_{(L,W)}^L = 1, \mathcal{F}_{(W,R)}^L = 1, \mathcal{F}_{(L,R)}^L = 2. \quad (\text{L-flow})$$

Ako prvi robot pokazuje namjeru skretanja udesno (R), tada je odgovor drugog robota strukturiran tokom preferencija \mathcal{F}^R :

$$\mathcal{F}_{(W,L)}^R = 4, \mathcal{F}_{(R,W)}^R = 2, \mathcal{F}_{(R,L)}^R = 4. \quad (\text{R-flow})$$

Ako prvi robot najavljuje namjeru čekanja (W), onda je odgovor drugog robota strukturiran tokom preferencija \mathcal{F}^W :

$$\mathcal{F}_{(W,L)}^W = 0, \mathcal{F}_{(R,W)}^W = 2, \mathcal{F}_{(R,L)}^W = 3. \quad (\text{W-flow})$$

Zapis $\mathcal{F}_{(R,W)}^W = 2$ znači da je opcija R preferirana u odnosu na opciju W i intenzitet te preferencije iznosi 2. Gornji indeks W označava namjeru prvog robota – čekanje.

³⁷Negativna vrijednost intenziteta preferencije znači da je prednost dana lijevoj opciji, a pozitivna vrijednost da je prednost dana desnoj opciji u paru.

Tablica 29: Akcije oba robota organizirane u hijerarhijski model odluke.

	akcije		
1. razina	R _{1L}	R _{1W}	R _{1R}
2. razina	R _{2L}	R _{2W}	R _{2R}
3. razina	R _{1L}	R _{1W}	R _{1R}

Tako definirani tokovi preferencija su zapisani u memoriju svakog robota i oni pri susretu jedan drugome signaliziraju te preferencije.³⁸

Razine hijerarhijskog modela odluke dani su u tablici 29. Prva i treća razina hijerarhije su jednake, što znači da je to sustav s povratnim vezama. Svaka akcija u prvoj razini je roditelj akcijama u drugoj razini, a svaka akcija u drugoj razini je roditelj akcijama u trećoj razini. Odgovarajući tokovi preferencija dani su u formulama (L-flow), (R-flow) i (W-flow).

Iterativni proces možemo započeti proizvoljnim početnim vektorom težina. Za jednake početne težine akcija u prvoj razini iterativni proces daje sljedeće vrijednosti (s preciznošću³⁹ $\epsilon = 0.0001$). Zaključak je da će prvi robot skrenuti udesno. Drugi robot rješava problem odlučivanja iz tablice

Tablica 30: Fiksna točka za potez robota R₁. Akcija udesno (R) ima najveći rang.

korak	R _{1W}	R _{1R}	R _{1L}
1.	0.295	0.546	0.159
2.	0.271	0.648	0.080
3.	0.257	0.686	0.057
4.	0.253	0.696	0.051
5.	0.252	0.698	0.049
5.
8.	0.252	0.699	0.049

(31) s tokom preferencija (R-flow). Rješenje tog problema dano je u tablici (31). Drugi robot također skreće udesno, što se i očekuje.

Tablica 31: Problem odlučivanja drugog robota R₂.

1. razina	R		
2. razina	L	W	R
rang	0.027	0.276	0.696

Roboti ne moraju imati istu logičku strukturu kao u našem primjeru. Na primjer, jedan od njih može imati viši prioritet (vozilo hitne pomoći, policija) što se reflektira u definiciji komponenata toka (W-flow) tog robota. Situacija se može zakomplicirati i dodavanjem još neke mogućnosti kretanja kao što je kretanje unatrag ili određivanjem polazne pozicije robota u koridoru (lijeva, desna).

11.10 Hijerarhija obrazovnih ciljeva

Ciljevi-postignuća o kojima je bilo riječi u odjeljku 6 pod naslovom *Motivacija za učenjem* odnose se na neposredne ciljeve koji su zacrtani kurikulumom i relativno brzo i lako se postižu. Naši osobni životni interesi se, za razliku od gore spomenutih, lako prepoznaju ali nisu jasno definirani, a samim

³⁸Ne bismo ulazili u tehničke specifikacije te komunikacije. Možemo zamisliti da su te preferencije tipizirane i da robot signalizira drugom robotu samo tip preferencija koji je u njega ugrađen.

³⁹Pod preciznošću se ovdje smatra razlika udaljenosti među težinama u posljednje dvije uzastopne iteracije nakon zaustavljanja algoritma.

tim nije poznato niti vrijeme niti obujam aktivnosti za njihovo postizanje. Većina takvih ciljeva služi razvoju osobnosti i stvaranju vlastite predodžbe o nama samima. Obrazovne ustanove i interesne skupine postavljaju i vlastite ciljeve koji im olakšavaju realizaciju programa i organizaciju nastave.

Mnogi od tih ciljeva imaju zajedničke pojedinosti kojima su obojeni i koje ih određuju. Na primjer, dijete koje je voli planinarenje i penjanje po stijenama i dijete koje zanima botanika imaju mnogo toga zajedničkog. I jedno i drugo će, žele li produbiti svoja znanja i sposobnosti trebati ovladati principima orijentacije u prirodi (čitanje karata, rukovanje kompasom i GPS-om, osnovama fotografije...), poželjno je da uvježbaju pružanje prve pomoći i samopomoći, upoznaju se sa opasnostima u prirodi, upoznaju se sa vlastitim slabostima i ograničenjima, razviju disciplinu kroz postupno prikupljanje vlastitog iskustva, proučavanjem literature i u razgovorima s ljudima koji u njihovom području interesa imaju više iskustva i znanja i koji ih mogu podučiti.

Zajedničke komponente naših interesa i postavljenih ciljeva posjeduju apstraktnije karakteristike koje čine *više ciljeve* i u hijerarhiji ciljeva obično se prikazuju u razini iznad razine osnovnih ciljeva. Tako dobivenu hijerarhiju ciljeva nazivamo *ciljnom strukturom* i ona, u praktičnoj primjeni vrednovanja aktivnosti koje vode ostvarivanju postavljenih ciljeva stoji iznad razine aktivnosti koja čini dno sveukupne hijerarhije (v. odjeljak 5 pod naslovom [Hijerarhija odluke i potciljevi](#)).

U kontekstu obrazovanja osnovnu razinu ciljne strukture *ciljevi-postignuća* i na njima se temelji plan i program izvođenja nastave. U današnjim kurikulumima rijetko kad se ciljna struktura nadograđuje višim obrazovnim ciljevima koji zadiru u osnovne ljudske vrijednosti i etičnost u najširem smislu – žalosno ali je tako.

11.10.1 Dualnost ciljeva i obrazovnih sadržaja

U obrazovnoj literaturi spominju se viši ciljevi obrazovanja ali autor ove knjige nije naišao na njihovu hijerarhijsku organizaciju i praktične upute kako ih iskoristiti u vrednovanju obrazovne strukture i programa. Realnost nije tako jednostavna kako se to čini na prvi pogled. Konkretni nastavni sadržaji su u službi ciljeva ali nije apriori jasna niti relativna važnost ciljeva niti relativna važnost obrazovnih sadržaja. U najjednostavnijem modelu ciljne strukture ciljevi obrazovanja su u višoj razini, a obrazovni sadržaji u nižoj razini. Svaki od ciljeva zastupljen je u nekim (ili svim) obrazovnim sadržajima u onoj mjeri u kojoj je organizator nastave to zamislio. Međutim, svaki obrazovni sadržaj, zbog svoje multidisciplinarnosti, potiče razvoj učenika u smjeru zacrtanih ciljeva i onih ciljeva koji možda nadilaze osnovne ciljeve programa. U tom smislu postoji povratna veza (dualnost) između ciljeva i nastavnih sadržaja koju je moguće i kvantitativno izraziti.

U odjeljku 11.7 pod nazivom [Samodualna hijerarhija](#) (str. 59) detaljno je obrazložena matematička pozadina takve hijerarhijske strukture. Ovdje ćemo dati pojednostavljen primjer u kontekstu obrazovnih ciljeva i sadržaja

Pretpostavimo da se radi o školi koja je društveno orijentirana, s kulturno političkim aspektom i da su zamišljena četiri obrazovna sadržaja: *Književnost* (Knj), *Strani jezik* (StJzk), *Matematika*⁴⁰ (Mat) i *Umjetnost* (Umj), a ciljevi su: *Poticanje kreativnog izražavanja* (KrIz), *Razvoj kritičkog mišljenja* (KrMi), *Uravnotežen individualni i socijalni život* (UrŽ) i *Briga za okoliš* (BrOk). Zanima nas zastupljenost svakog predmeta u ukupnoj nastavi i naglašenost ciljeva na nekoj skali.

Hijerarhiju ćemo organizirati tako da ćemo u višu razinu staviti ciljeve, a u nižu razinu predmete⁴¹. Za svaki od ciljeva potrebno je načiniti graf preferencije među elementima niže razine i obratno, za svaki element niže razine potrebno je načiniti graf preferencije među elementima više razine. Takva hijerarhijska struktura generira operator samorangiranja čija fiksna točka definira težine elemenata u svakoj razini.

Hijerarhijska struktura ciljeva i sadržaja dana je u tablici 32 i ima tri razine koje smo označili slovima C1, S, C2 kao oznake za *Ciljeve* i *Sadržaje*. Razine ciljeva C1 i C2 imaju iste elemente jer se radi o samo-dualnoj hijerarhiji.

⁴⁰Možda bi bilo bolje predvidjeti neki opći predmet prirodno matematičkog usmjerenja, ali za ilustraciju ćemo ga nazvati *Matematika*.

⁴¹ili obratno jer se radi o sustavu s povratnim vezama

Varijabla *children* određuje djecu za svakog roditelja iz više razine, a djeca se nalaze u razini neposredno ispod. Oznaka 1.1: {2.1, 2.2, 2.4} znači da čvor *KrIz(1.1)*, iz prve razine, kao roditelj ima kao djecu čvorove {1, 2, 4} iz druge razine, a to su {*Knj*, *StJzk*, *Umj*}. *Matematika Mat(2.3)* nije dijete od *Kreativnog izražavanja KrIz(1.1)*, ali *Književnost*, *Strani jezik* i *Umjetnost* jesu što znači da kvalitetu roditelja *KrIz Matematika* ne posjeduje, a ostali predmeti je posjeduju. Primijetimo također da *Mat(2.3)* nije dijete čvora *UrŽv(1.3)* iako *UrŽv(3.3)* iz treće razine kao dijete može biti zastupljen u nastavi matematike putem problemskih zadataka⁴².

Kod uspoređivanja predmeta u parovima za određeni cilj (roditelj) u hijerarhiji prednost se daje onom predmetu u kojem je cilj više zastupljen. Za predmete-roditelje iz druge razine i par ciljeva iz treće razine prednost se daje onom cilju s kojim je predmet više usklađen. Rezultati uspoređivanja u parovima dani su u varijabli *preferences*. Za preciznije određivanje intenziteta preferencija potrebno je definirati sadržaje svakog predmeta što je dugotrajan proces. Intenziteti preferencije koji su ovdje dani su isključivo radi ilustracije same procedure.

Metoda potencijala daje sljedeće rangove za ciljeve:

$$\text{UrŽv} \rightarrow 0.315954; \text{KrIz} \rightarrow 0.302689; \text{KrMi} \rightarrow 0.302689; \text{BrOk} \rightarrow 0.302689$$

i sljedeće rangove za predmete:

$$\text{Umj} \rightarrow 0.368061; \text{Knj} \rightarrow 0.284170; \text{Mat} \rightarrow 0.176258; \text{StJzk} \rightarrow 0.171511$$

⁴²...što bi trebalo predvidjeti u nastavnim sadržajima.

Tablica 32: Primjer samodualne hijerarhije obrazovnih ciljeva i sadržaja

```
name: {Ciljevi-Sadržaji} ;
levels: {C1, S, C2} ;
structure:
{C1: {KrIz,KrMi,UrŽv,BrOk}, S: {Knj,StJzk,Mat,Umj},
C2:{KrIz,KrMi,UrŽv,BrOk}} ;

children: {
1.1: {2.1, 2.2, 2.4},
1.2: {2.1, 2.3, 2.4},
1.3: {2.1, 2.2, 2.4},
1.4: {2.1, 2.3, 2.4},
2.1: {3.1, 3.2, 3.3, 3.4},
2.2: {3.1, 3.3, 3.4},
2.3: {3.1, 3.2, 3.3},
2.4: {3.1, 3.2, 3.3, 3.4}
} ;

preferences: {
1.1: {(2.1, 2.2):2, (2.4, 2.1):3, (2.4, 2.2):1},
1.2: {(2.1, 2.3):1, (2.4, 2.1):2, (2.4, 2.3):0},
1.3: {(2.1, 2.2):2, (2.1, 2.4):0, (2.4, 2.2):0},
1.4: {(2.1, 2.3):2, (2.4, 2.1):1, (2.4, 2.3):3},
2.1: {(3.1, 3.2):2, (3.1, 3.3):0, (3.1, 3.4):3, (3.2, 3.3):0,
(3.2,3.4):3, (3.4, 3.3):0},
2.2: {(3.3, 3.1):1, (3.1, 3.4):1, (3.3, 3.4):0},
2.3: {(3.2, 3.1):5, (3.1, 3.3):0, (3.2, 3.3):5},
2.4: {(3.1, 3.2):1, (3.1, 3.3):0, (3.1, 3.4):2, (3.3, 3.2):4,
(3.4, 3.2):2, (3.3, 3.4):4}
} ;

aggregation: {average} ;
htype: {dual} ;
```


Kako interpretirati dobiveni rezultat? *Uravnotežen individualni i socijalni život* (UrŽv) je navažniji kao cilj, a prate ga *Poticanje kreativnog izražavanja* (KrIz), *Razvoj kritičkog mišljenja* (KrMi) i *Briga za okoliš* (BrOk). Što se predmeta tiče najvažnija je *Umjetnost* (Umj), a slijede je *Književnost* (Knj), *Matematika* (Mat) i *Strani jezik* (StrJzk).

U društveno-umjetničkom orijetiranom nastavnom programu to se moglo i očekivati. Ako rezultat nije u skladu s namjerama kreatora programa onda bi trebalo revidirati nastavne programe i ponoviti poroceduru vrednovanja ciljeva i predmeta na gore opisani način.

12 Umjesto zaključka

12.1 Agregacija

Agregacija ocjena nije tako jednostavna kako se čini. Aritmetička sredina nije preporučljiva ako se ocjena shvaća kao podatak na ordinalnoj skali jer razlika sredina nije invarijantna na rastuću transformaciju skale. Dodatni problem je što, ako se i računa srednja ocjena kao aritmetička sredina, nakon agregacije obrazovni sustav 'vraća' (zaokružuje) dobivenu vrijednost natrag na skalu, napr. 3.51 na 4 i daje ocjenu *vrlo dobar*. Takav postupak ničim nije opravdan osim uzrečicom: "Svi tako rade". Takvo ocjenjivanje je stresno za učenike, njihove roditelje i nastavnike. Jedan od načina je prekinuti s takvim načinom što bi se moglo postići reorganizacijom obrazovne prakse. Obrazovna praksa nije uzrok, ona je posljedica dubljih nedorečenosti u samom društvu. Osnovno pitanje je što društvo očekuje od obrazovanog pojedinca i je li obrazovanje koje ono nudi tom pojedincu poticajno ili ne za njegov osobni uzlet?

Vidjeli smo da je cilj metode agregacije donošenje rang-liste. Rang-lista se odnosi na grupu učenika koja je obuhvaćena ocjenjivanjem (razred, škola, država), a problem koji mi se čini otvorenim je kako interpretirati te rang liste u obrazovnom sustavu. To otvara drugo, smjelije pitanje, je li ocjena nužna i čemu zapravo služi. Ako je ona osnova za upis i viši razred i daljnje školovanje onda je za takav nastavak školovanja bitno predznanje, a ne ocjena. Predznanje se u principu odnosi na tematske jedinice i za nastavak školovanja treba provjeriti je li učenik ovladao tim tematskim jedinicama. Na primjer, ako je obrazovni cilj da učenik razumije i zna računati s postotcima, i ako to nije savladao onda nije 'zaslužio' da upiše srednju školu ili fakultet dok to ne savlada. Takvo viđenje stvari traži totalnu reviziju obrazovnog sustava i njegovog razumijevanja. To nije problem raznih kurikula već organizacije školstva.

Većina metoda koje smo prezentirali zadovoljava *aksiom o irelevantnoj alternativni*. To znači da u agregaciji učenikovih ocjena sudjeluju samo njegove ocjene, a ne i ocjene drugih učenika. Takav zahtjev odudara od prirode procedure donošenja individualne ocjene učeniku jer nastavnik uvijek uspoređuje učenike u parovima, pa čak i kad ispravlja test ili domaću zadaću. I procedura bodovanja zadatka na testu nije nikada fiksirana unaprijed jer ispravljač revidira svoje viđenje tokom ispravljanja testa.

12.2 Metoda potencijala

Metoda potencijala, za razliku od drugih metoda, uvažava i nepotpune podatke. Konkretno, ako neki učenik nije dobio ocjenu na jednom od testova ne mora mu se dati o bodova (ili ocjena 1) na tom testu. Podaci se mogu uvažiti i analizirati takvi kakvi jesu. To vrijedi i za ordinalno rangiranje također.

Osnova metode je uspoređivanje u parovima, a to je zahtjevna procedura i kod prikupljanja podataka i za numeričku obradu. Može se raditi i bez popratnog sofvera ali tada zahtijeva znanje iz linearne algebre. U današnjem, informatički rastućem okruženju to više ne bi trebao biti problem.

Svi izračuni, što se metode potencijala tiče, urađeni su pomoću softvera koji je u završnoj fazi izrade ali nije još za komercijalnu upotrebu. Neki dijelovi tog sofvera bit će dostupni akademskoj zajednici.

Ranije smo govorili o socijalnim aksiomima koji se mogu zahtijevati za metodu. Neki od tih aksioma diskutirani su u članku Čaklović, L. and Kurdija, A. S. (2017). *A universal voting system based on the Potential Method*. *European Journal of Operational Research*, 259:677–688. Možda nije na odmet napomenuti da su Bordina metoda i Condorcetova metoda specijalni slučajevi metode potencijala što je također dokazano u gore spomenutom članku.

12.3 Kako vidim budućnost ocjenjivanja

Vratimo se natrag na pitanje i problem donošenja individualne ocjene kao mjeru znanja i ovladavanja raznih umijeća, a za koji smo u samom početku rekli da nije tema ove rasprave. Ovdje pod ocjenom ne mislim na broj, to može biti logička vrijednost ili skup takvih vrijednosti, procjena i samoprocjena izražena nekim kategorijskim varijablama, ili štošta drugo. Neka radni termin za te podatke bude *profil učenika*. Za takvo opisivanje učenika treba precizirana procedura i trening za kreatora takvog profila. Za donošenje prosudbe o tome je li učenik smije zakoračiti u viši stupanj obrazovnog sustava nije više zaslužan niz brojeva nego skup podataka. Takav skup podataka trebao bi biti dinamičan u smislu da se može revidirati kad je učenik na to spreman, a sustav mu to treba omogućiti.

Analiza takvih profila svakako treba biti hijerarhijska i višekriterijska jer su podaci kvalitativno različiti. Metoda potencijala i hijerarhijski modeli regresije trebale bi biti u stanju analizirati ih.

Jedna od uloga današnje ocjene, bar što se tiče upisa na fakultet, ima cilj razlikovanja učenika, ali je inflacija petica ukinula takvu tu njenu zadaću. Gore opisani pristup profiliranja učenika ne samo da onemogućava inflaciju petica već ukida i ocjenu kao mjeru usvojenog znanja i vještina. Nastavak školovanja omogućava učenikov profil, a ne ocjena. Treba znati rukovati s profilima, a ne ocjenama.

Literatura

- Ashby, F. G. and Townsend, J. T. (1984). Measurement scales and statistics: the misconception misconceived. *Psych. Bull.*, 96:394–401.
- Barron, K. E. and Harackiewicz, J. M. (2001). Achievement goals and optimal motivation: Testing multiple goal models. *Journal of Personality and Social Psychology*, 80(5):706–722.
- Biggs, J. B. (1987). *Study Process Questionnaire Manual. Student Approaches to Learning and Studying*. Melbourne.
- Christie, T. and Forrest, G. M. (1981). *Defining Public Examination Standards*. Schools Council Research Studies, MacMillan, London.
- Čaklović, L. (2003). Graph Distance in Multicriteria Decision Making Context. *Metodološki zvezki (Advances in Methodology and Statistics)*, 1(19):25–34.
- Čaklović, L. (2005). Stochastic preference and group decision. *Metodološki zvezki (Advances in Methodology and Statistics)*, 2(1):205–212.
- Čaklović, L. (2011). Conflict Resolution. Risk-As-Feelings Hypothesis. *Labsi Working Papers*, (35):1–16. (<http://www.labsi.org/wp/labsi35.pdf>).
- Čaklović, L. (2014). *Teorija vrednovanja s naglaskom na metodu potencijala*. SLAP, Jastrebarsko, Croatia.
- Dawes, R. and Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81:95–101.
- Dweck, C. (1986). Motivational processes affecting learning. *American Psychologist*, (41):040–1048.

- D'Souza, K. A. and Maheshwari, S. K. (2010). Factors Influencing Student Performance in the Introductory Management Science Course. *Academy of Educational Leadership Journal*, 14(3).
- Eckblad, G. (1981). *Scheme theory: Conceptual framework for cognitive-motivational processes*. Academic Press, New York.
- Edelson, D. C. and Joseph, D. M. (2004). The interest-driven learning design framework: Motivating learning through usefulness. In *Proceedings of the 6th International Conference on Learning Sciences*, ICLS '04, pages 166–173. International Society of the Learning Sciences.
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist*, 34(3):169–189.
- French, S. (1989). Statistical and Decision Theoretic Aspects of Examination Assessment. *Trabajos de estadística*, 4(1):33–66.
- French, S. (1998). *Decision theory - An introduction to the mathematics of rationality*. Ellis Horwood, Chichester.
- French, S. and Vassiloglou, M. (1986). Strength of performance and examination assessment. *Brit. J. Math. Statist. Psych.*, 39:1–14.
- Gaito, J. (1980). Measurement scales and statistics: resurgence of an old misconception. *Psych. Bull.*, 87:564–567.
- Greator, J. (2003). What happened to limen referencing? an exploration of how the awarding of public examinations has been and might be conceptualised. pages 1–15, Edinburgh. A paper presented at the BERA 2003.
- Greator, J. (2007). Contemporary GCSE and A-level Awarding: A psychological perspective on the decision-making process used to judge the quality of candidates' work. A paper presented at BERA 2007.
- Greator, J. (2009). How are archive scripts used in judgements about maintaining grading standards? A paper presented at BERA 2009.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47:263–291.
- Krantz, D. H., Luce, D. R., Suppes, P., and Tversky, A. (1971). *Foundations of Measurement, vol. 1: Additive and polynomial representations*. Academic Press, New York.
- Laming, D. (2004). *Human judgement The Eye of the Beholder*. Cambridge University Press, Cambridge.
- Lengnick-Hall, C. and Sanders, M. (1997). Designing Effective Learning Systems for Management Education. *Academy of Management Journal*, 40(6):1334–1368.
- Linn, R. L., editor (1993). *Educational Measurement*. Oryx Press, Phoenix, USA.
- Locke, E. and Latham, G. (1990). *A theory of goal setting and task performance*. Prentice-Hall, Englewood Cliffs, NJ.
- McClelland, D. C., Koestner, R., and Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, 96:690–702.
- Murphy, P. K. and Alexander, P. A. (2000). A Motivated Exploration of Motivation Terminology. *Contemporary Educational Psychology*, 25:3–53.

- National Research Council (2000). *How people learn: brain, mind, experience and school*. Washington, D.C.
- Pintrich, P. R. (2000a). An Achievement Goal Theory Perspective on Issues in Motivation Terminology, Theory, and Research. *Contemporary Educational Psychology*, (25):92–104.
- Pintrich, P. R. (2000b). An Achievement Goal Theory Perspective on Issues in Motivation Terminology, Theory, and Research. *Contemporary Educational Psychology*, 25:92–104.
- Pintrich, P. R. (2000c). Multiple Goals, Multiple Pathways: The Role of Goal Orientation in Learning and Achievement. *Journal of Educational Psychology*, 92(3):544–555.
- Pollitt, A. and Elliott, G. (2003a). Finding a proper role for human judgement in the examination system. Technical report, Research and Evaluation Division, University of Cambridge Local Examinations Syndicate, 1 Hills Road, Cambridge.
- Pollitt, A. and Elliott, G. (2003b). Monitoring and investigating comparability: a proper role for human judgement. Technical report, Research and Evaluation Division, University of Cambridge Local Examinations Syndicate, 1 Hills Road, Cambridge.
- Roberts, F. S. (1979). *Measurement Theory*. Addison Wesley, Reading, Ma.
- Robustelli, S. (2008). *There Is More Than One Yellow Brick Road to Achievement*. VDM Verlag Dr. Mueller e.K., Saarbrücken, Germany.
- Taylor, S. E., Pham, L. B., Rivkin, I. D., and Armor, D. A. (1998). Imagination: Mental Simulation, Self-Regulation, and Coping. *American Psychologist*, 53(4):429–439.
- Thompson, D. V., Hamilton, R. W., and Petrova, P. K. (2009). When Imagination Hinders Behavior: The Effects of Outcome versus Process Oriented Thinking on Decision Difficulty and Performance. *Journal of Consumer Research*, 36(4):562–574.
- Williams, K. and Williams, C. (2011). Five Key Ingredients for Improving Student Motivation. *Research in Higher Education Journal*, 12:1–23.
- Wilson, J. I. (2009). A two factor model of performance approach goals in student motivation for starting medical school. *Issues in Educational Research*, 19(3):271–281.