

---

<b>Uvodno razmatranje .....</b>	<b>3</b>
<b>1. UČENJE I POUČAVANJE UZ POMOĆ INTELIGENTNIH TUTORSKIH SUSTAVA .....</b>	<b>6</b>
<b>1.1. Inteligentni tutorski sustavi.....</b>	<b>7</b>
1.1.1. Interakcija prirodne sklonosti za učenje i načina poučavanja .....	9
1.1.2. Modeliranje učenika.....	10
<b>1.2. Struktura sustava TEx-Sys.....</b>	<b>12</b>
<b>1.3. Web-orijentirani inteligentni tutorski sustavi .....</b>	<b>13</b>
1.3.1. Preporuke za Web-orijentirano poučavanje .....	13
<b>1.4. Struktura sustava DTEEx-Sys.....</b>	<b>16</b>
<b>2. VREDNOVANJE POSTIGNUĆA U RADU S INTELIGENTNIM TUTORSKIM SUSTAVIMA.....</b>	<b>19</b>
<b>2.1. Pregled provedenih istraživanja.....</b>	<b>19</b>
2.1.1. Povećanje učinkovitosti poučavanja uz pomoć računala .....	20
2.1.2. Smanjenje vremena potrebnog za postizanje ciljeva poučavanja.....	21
2.1.3. Trošak poučavanja uz pomoć računala .....	22
<b>2.2. Metode vrednovanja inteligentnih tutorskih sustava .....</b>	<b>23</b>
2.2.1. Dokazi ispravnosti (eng. proofs of correctness).....	23
2.2.2. Vrednovanje na osnovi kriterija (eng. criterion-based evaluation).....	23
2.2.3. Znanje i ponašanje eksperta (eng. expert knowledge and behaviour).....	24
2.2.4. Potvrđivanje (eng. certification) .....	24
2.2.5. Analiza osjetljivosti (eng. sensitivity analysis).....	24
2.2.6. 'Pilot' testiranje (eng. pilot testing).....	24
2.2.7. Eksperimentalno istraživanje (eng. experimental research).....	25
<b>2.3. Kriteriji za vrednovanje utjecaja inteligentnih tutorskih sustava na proces učenja..</b>	<b>25</b>
2.3.1. Kriteriji za određivanje mjera postignuća .....	25
2.3.2. Afektivne mjere (eng. affective measures).....	27
<b>2.4. Učinkovitost Web-orijentiranog poučavanja.....</b>	<b>27</b>
2.4.1. Nedostaci istraživanja Web-orijentiranog učenja i poučavanja.....	27
<b>3. PRISTUP VLASTITOM ISTRAŽIVANJU .....</b>	<b>31</b>
<b>3.1. Planiranje i realizacija eksperimenta .....</b>	<b>31</b>
<b>3.2. Analiza rezultata.....</b>	<b>32</b>
3.2.1. T-test.....	32
3.2.2. Veličina učinka.....	34
3.2.3. Metričke karakteristike testova .....	35
3.2.4. Komentar rezultata.....	41
<b>4. ZAKLJUČAK I PRAVCI DALJNJEG ISTRAŽIVANJA.....</b>	<b>43</b>
<b>Literatura.....</b>	<b>44</b>
<b>PRILOG .....</b>	<b>48</b>

---

<i>Prilog A - Sadržaj predavanja.....</i>	<i>49</i>
<i>Prilog B - Pre-test.....</i>	<i>66</i>
<i>Prilog C - Post-test.....</i>	<i>70</i>
<i>Prilog D - Izgled baze znanja.....</i>	<i>76</i>
<i>Prilog E - Tablica uspjeha studenata .....</i>	<i>83</i>

## Uvodno razmatranje

Prema nekim definicijama *e-učenje* (eng. e-learning) se poistovjećuje s on-line učenjem, ali se on-line i e-učenje razlikuju. E-učenje je nova paradigma učenja uz pomoć različitih uređaja koji se temelje na elektronskoj tehnologiji [Stankov *et al.*, 2004], a odnosi se na dostavljanje sadržaja učenja učenicima putem svih vrsta elektroničkih medija, uključujući računala, Internet, Intranet, emitiranje putem satelita, audio i video vrpce, interaktivnu televiziju, CD-ROM-ove ili DVD-ove i sl. (<http://www.learnframe.com/aboutelearning>). U ovom radu ćemo se orijentirati samo na jedan aspekt e-učenja, na Web-orijentirano učenje tj. konkretno na učenje i poučavanje uz pomoć Web-orijentiranih inteligentnih tutorskih sustava.

Individualizirano ili tutorsko poučavanje je poučavanje tipa "jedan učitelj – jedan učenik", za razliku od klasičnog poučavanja u razredu gdje na jednog učitelja dolazi oko 20-30 učenika. Već se iz ove rečenice vidi razlika u vremenu kojeg učitelj može posvetiti svakom učeniku. Kod klasičnog predavanja u razredu svaki učenik se mora izboriti za svojih "5 minuta", pod uvjetom da to želi. Često se u velikim razredima događa da učenik pasivno prati nastavu, a učitelj to ne primjećuje. Naravno, nije dovoljno da učenik samo bude prisutan, te da uči tek kada se najavi ispitivanje, učenik mora aktivno sudjelovati kako bi znanje i vještine koje usvoji bile što kvalitetnije. U situaciji kada jedan učitelj poučava samo jednog učenika, on mora biti aktivan, ne može se "sakriti", može dobiti dodatno objašnjenje ako želi, a ne mora se dugo zadržavati na onome što je dobro usvojio.

Razmišljanjem o razlikama između individualiziranog i klasičnog poučavanja čini se da individualizirano poučavanje ima prednosti u odnosu na klasično poučavanje. Isticanje prednosti ili nedostataka nije dovoljno kako bi se uvjerali u veću ili manju učinkovitost nekog načina poučavanja, potrebno je provesti istraživanje. Jedan od primjera uspoređivanja razlika između individualiziranog i klasičnog načina poučavanja je istraživanje koje je provodio Benjamin Bloom [Bloom, 1984; prema Fletcher, 2003] sa svojim studentima na Sveučilištu Chicago. Razlika je išla u korist individualiziranog poučavanja (to su i očekivali), a bilo je iznenađujuće koliko je takvo poučavanje bilo važno. Dobivena je razlika od dvije standardne devijacije ili "2-sigma razlika", jer se obično standardna devijacija označava grčkim slovom sigma. [Fletcher, 2003]. To je približno ekvivalentno povećanju od 50-og do 98-og centila ili razine postignuća. Centil je bilo koja od 99 točaka označenih brojevima koje dijele uređeni skup postignuća na 100 dijelova, od kojih svaki sadrži 1/100 ili 1% cjeline [Fletcher, 2003]. Npr. ako se uspjeh nekog učenika nalazi između 85-og i 86-og centila to znači da je 85% učenika ostvarilo lošiji, a 14% bolji rezultat od njega. Ako neki učenik na testu ostvari uspjeh koji se nalazi iznad 99-og centila to ne znači da je riješio 99% testa, nego je njegov uspjeh među 1% najboljih.

Jasno je da različiti učenici imaju različite sposobnosti, pa tako dok se kod klasičnog poučavanja u razredu bolji učenici dosađuju, učenici slabijih sposobnosti su često izgubljeni u pokušajima praćenja nastave. Iako se ulažu naponi za održavanje zajedničke razine znanja u razredu, tijekom svake školske godine te se razlike povećavaju, otprilike jednu godinu razlike za svaku školsku godinu [Heuston, 1997; prema Fletcher, 2003]. Svi učenici dolaze npr. u osnovnu školu s različitim sposobnostima i predznanjem. Te se razlike tijekom vremena povećavaju. Uspješniji učenici i dalje napreduju, a oni koji su manje uspješni svake godine zaostaju sve više jer je za usvajanje novog znanja često potrebno dobro usvojiti prethodno znanje. Tako npr. ako učenik nije dobro usvojio zbrajanje i oduzimanje cijelih brojeva, teško će se snaći s razlomcima. Individualizirano poučavanje omogućuje učitelju prilagođavanje tempa i sadržaja potrebama i sposobnostima individualnog učenika. Kad učenik nešto usvoji, onda se može napredovati dalje. Kod klasičnog predavanja u razredu učitelj se mora pokušati prilagoditi većini, tj. prosječnim učenicima, jer je obično njih najviše.

Radile su se i usporedbe interakcija učenika i učitelja kod individualizirane i klasične nastave. Primjer je istraživanje koje su provodili Graesser i Person [Graesser & Person, 1994; prema Fletcher, 2003]. Oni su usporedili individualizirano poučavanje s poučavanjem u razredu sljedećih sadržaja: algebra za 7.razred i metode istraživanja za studente. Promatrali su prosječan broj pitanja koje učenici i učitelji postavljaju tijekom poučavanja kao kriterij interaktivnosti. Utvrdili su da u razredu učitelj u prosjeku postavi oko 3.0 pitanja, a učenik oko 0.1 pitanja. Kod individualiziranog, tj. tutorskog poučavanja prosječan broj pitanja koja postavlja učenik (tutor odgovara) je bio 21.1 za metode istraživanja i 32.2 za algebru, a broj pitanja koja postavlja tutor (učenik odgovara) je 117.2 za metode istraživanja i 146.4 za algebru.

Možemo se upitati: Kako bi učitelj u razredu mogao odgovoriti na 20-30 pitanja svakog od 20-30 učenika tijekom svakog sata poučavanja, a kako bi tek mogao pripremiti i postaviti 115-145 pitanja prilagođenih potrebama svakog učenika te još dati povratnu informaciju za svaki odgovor? Naravno, da se na to pitanje ne očekuje odgovor jer se poučavanje u razredu ne može natjecati s individualiziranim poučavanjem u interaktivnosti i prilagodljivosti. Zašto onda svi učenici ne uče na individualizirani način? Zato što je ekonomski nemoguće imati tutora za svakog učenika. Međutim, uključivanjem računala u obrazovanje i pojavom inteligentnih tutorskih sustava otvaraju se nove mogućnosti u poučavanju. Posljednjih desetljeća su se brzina računala te njihovi kapaciteti za spremanje podataka udvostručavali svakih 18-24 mjeseca [Phipps & Merisotis, 1999], a u isto vrijeme se smanjivala njihova cijena. U usporedbi s ljudskim tutorima računala su jeftinija, pa bi se uz njihovu pomoć možda mogao popuniti dio razlike koja postoji između individualiziranog i klasičnog poučavanja (Bloom-ova razlika od dvije standardne devijacije ili "2-sigma razlika" [Fletcher, 2003]). Međutim, tehnologija ne može zamijeniti ljudski faktor u učenju i poučavanju. William Massy [Phipps & Merisotis, 1999] je primijetio kako će se u procesu obrazovanja na akademskoj razini s napretkom učenja uz pomoć računala uloga profesora promijeniti, te će profesori od stručnjaka za sadržaj postati i stručnjaci za oblikovanje procesa učenja, mentori, osobe koje motiviraju, te "ekspertni učenici" (eng. expert learners) tj. osobe koje vode proces učenja pokazujući svojim primjerom kako treba učiti.

Čini se da inteligentni tutorski sustavi mnogo obećavaju što se tiče poboljšanja postignuća učenika, ali potrebno je ispitati ostvaruju li se ta obećanja kod konkretnih inteligentnih tutorskih sustava. To je na neki način bila i inspiracija za ovaj rad. Dalje slijedi opis rada po poglavljima.

U prvom poglavlju kratko objašnjavamo i tumačimo strukturu i organizaciju inteligentnih tutorskih sustava. Često se spominje i učenje "bilo kad i bilo gdje" tj. omogućavanje pristupa nastavnim sadržajima s bilo kojeg mjesta uz pomoć Web-a. Posebno nam je zanimljiva kombinacija Web-orijentiranog učenja i inteligentnih tutorskih sustava, jer je takav sustav DTEEx-Sys (Distributed Tutor Expert System) [Rosić, 2000] na kojem je provedeno istraživanje opisano u ovom radu. Osim toga, opisuje se i sustav TEx-Sys (Tutor-Expert System) [Stankov, 1997] koji je prethodio sustavu DTEEx-Sys jer je na njemu izgrađena baza znanja koja je korištena u istraživanju.

U drugom poglavlju se daje pregled metoda vrednovanja inteligentnih tutorskih sustava, te primjeri provedenih istraživanja koja pokazuju određeni uspjeh takvih sustava. U primjere su uključeni i CAI sustavi (prethodnici inteligentnih tutorskih sustava) zato nema mnogo podataka vezanih samo za inteligentne tutorske sustave. Neka od tih istraživanja su pokušavala dokazati da CAI sustavi mogu oponašati karakteristike ljudskih tutora u individualiziranom poučavanju [Regian & Shute, 1992; prema Yaakub, 1998]. Regian i Shute vjeruju da u nekim aspektima CAI sustavi mogu biti čak i učinkovitiji od ljudskih tutora jer su

---

stalno motivirani, postojani i ne umaraju se. Inteligentni tutorski sustavi bi trebali ostvarivati učinkovitiji od CAI sustava. Mnogo sredstava se ulaže u razvoj takvog načina poučavanja. Prema [Cotton, 1991] učenici vole raditi s računalima zato što su računala beskonačno strpljiva, ne umaraju se i ne ljute, dozvoljavaju učenicima privatnost pri radu, ne zaboravljaju pohvaliti ni ispraviti, poučavaju na individualiziran način, omogućuju isprobavanje različitih opcija, mnogo su objektivnija od učitelja, poučavaju u malim koracima itd.

Potaknuti dobrim rezultatima tih istraživanja odlučili smo ispitati i sustav DTE<sub>x</sub>-Sys kojeg razvija Zavod za informatiku Fakulteta prirodoslovno-matematičkih znanosti i odgojnih područja Sveučilišta u Splitu. Provedeno istraživanje i rezultati su detaljno opisani u trećem poglavlju, uključujući i kratak opis statističkih metoda koje su korištene pri analizi dobivenih rezultata (npr. t-test, veličina učinka).

## 1. UČENJE I POUČAVANJE UZ POMOĆ INTELIGENTNIH TUTORSKIH SUSTAVA

Klasično poučavanje u razredu se u početku održavalo uz pomoć krede i ploče. Kada su postali dostupni npr. filmovi i projektori, radio i televizija, tehnologija se počinje uključivati u poučavanje. Kasnije se nastojalo u poučavanje uključiti i računalnu tehnologiju. Računala su se u početku koristila na isti način kao npr. televizija, filmovi i radio tj. kao pomoćni alat kojim će se na malo drugačiji način prikazati dijelovi sadržaja klasičnog predavanja.

Nastava temeljena na računalu (eng. Computer-Based Instruction, kratica CBI) i obrazovanje temeljeno na računalu (eng. Computer-Based Education, kratica CBE) su najširi termini koji se odnose se na bilo kakvu upotrebu računala u obrazovnim uvjetima, uključujući vježbanje i ponavljanje, tutorsko poučavanje, simulacije, upravljanje poučavanjem, programiranje, izrada baza podataka, pisanje upotrebom uređivača teksta i dr. Ovi termini se mogu koristiti i za učenje isključivo pomoću računala ili za nadopunjavanje materijala kojeg predstavlja učitelj [Cotton, 1991].

Jedan od najranijih primjena računala u poučavanju je sustav PLATO (eng. Programed Logic for Automated Teaching Operations) koji je počeo kao projekt na Sveučilištu Illinois 1960. PLATO je omogućio računalno-zasnovanom poučavanju (eng. Computer-Based Instruction) integraciju teksta i grafike, te je omogućio okolinu u kojoj su učitelji mogli izrađivati svoje PLATO lekcije [Fletcher, 2003].

Forcier (1996) [prema Yaakub, 1998] dijeli upotrebu računala u poučavanju na: upravljanje (eng. management), poučavanje i učenje (eng. instruction and learning), te istraživanje u obrazovanju (eng. educational research). Podkategorije poučavanja i učenja uključuju računalnu pismenost (eng. computer literacy), poučavanje uz pomoć računala (eng. Computer-Assisted Instruction, kratica CAI), te poučavanje upravljano s računalom (eng. computer-managed instruction).

Poučavanje uz pomoć računala (eng. Computer-Assisted Instruction) je uži pojam od CBI, a najčešće se odnosi na upotrebu računala za predstavljanje sadržaja poučavanja u obliku "vježbanja i ponavljanja" (eng. drill and practice), tutorskog poučavanja (eng. tutorial) i simulacija, a sve navedeno se može koristiti samostalno ili kao nadopunjavanje materijala kojeg predstavlja učitelj [Cotton, 1991]. Termin se koristi uz sinonime: računalno-zasnovano učenje (eng. Computer-Based Learning, kratica CBL) i učenje uz pomoć računala (eng. Computer-Assisted Learning, kratica CAL) [Chambers & Sprecher, 1983; prema Yaakub, 1998].

Kod CAI postoji interakcija između računala i učenika pri učenju. Mnoge CAI lekcije su u početku sličile knjigama, gdje je na ekranu bio prikazan samo tekst, a jedina korisnikova interakcija s računalom je bilo pritiskanje tipke za napredovanje kroz program [Steinberg, 1990; prema Yaakub, 1998]. Takvi su se programi nazivali "elektronički okretači stranica" (eng. electronic page turners). Usprkos tome, lekcije su počele sadržavati složenije vježbe, te su omogućavale stjecanje i razvijanje kritičkog razmišljanja i vještina za rješavanje problema. Ako učenik na problem kojeg rješava ne da točan odgovor, onda ga računalo upućuje na ponavljanje odgovarajuće lekcije [Debak, 2002], a ako je odgovor točan onda može rješavati novi problem.

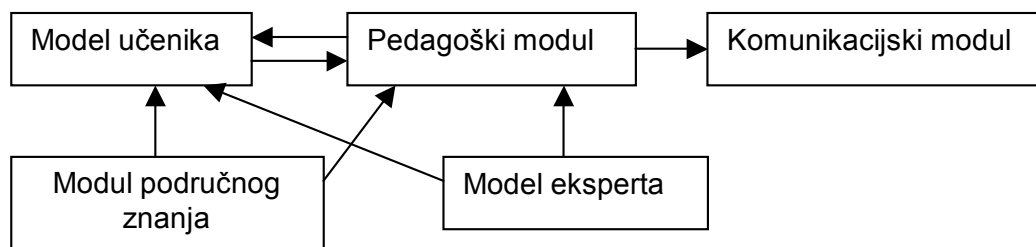
### 1.1. Inteligentni tutorski sustavi

Inteligentno poučavanje uz pomoć računala (eng. Intelligent Computer-Assisted Instruction, kratica ICAI) je tip CAI zasnovan na tri vrste znanja pri rješavanju problema. To su: znanje eksperta (eng. expert knowledge), znanje učenika (eng. student knowledge) i dijagnostičko znanje (eng. diagnostic knowledge). Drugim riječima, ICAI se sastoji od tri modula: modul eksperta (eng. expert module), modul učenika (eng. student module) i tutorski modul (eng. tutorial module). ICAI je također poznat pod nazivom: inteligentni tutorski sustav (eng. Intelligent Tutoring System, kratica ITS) [Burns & Capps, 1988; prema Yaakub, 1998]. Za ICAI sustave se osim naziva "inteligentni tutorski sustavi", koriste još: interaktivne okoline za poučavanje (eng. interactive learning environments), kognitivni tutorski sustavi (eng. cognitive tutoring systems), inteligentni sustavi za obuku (eng. intelligent training systems), i sl. [Regian, 1997]. Dalje u tekstu će se za ICAI sustave koristiti naziv *inteligentni tutorski sustavi*.

Woolf [prema Debak, 2002] navodi komponente ITS-a na drugačiji način tj. povećava broj komponenti koje bi ITS morao imati. To su:

- model učenika (eng. student model)
- pedagoški modul (eng. pedagogical module)
- modul područnog znanja (eng. domain knowledge module)
- komunikacijski modul (eng. communication module)
- model eksperta (eng. expert model)

Na Slici 1-1 vidi se interakcija komponenti ITS-a.



Slika 1-1: Interakcija komponenti ITS-a [Beck *et al.*, 1996]

*Model učenika* sadrži informacije koje određuju trenutno znanje svakog učenika. Može sadržavati i informacije o načinu na koji je učenik usvojio i organizirao novo znanje, te znanje koje je učenik razumio ili nije razumio [Regian, 1997]. Sustav nakon svakog ciklusa učenja i poučavanja, ovisno o rezultatima učenika, mijenja model učenika. Pod tim ciklusom učenja i poučavanja misli se na procese od trenutka generiranja problema kojeg učenik rješava do završetka rješavanja problema kada sustav daje učeniku povratnu informaciju.

*Pedagoški modul* sadrži model koji se odnosi na proces učenja [Debak, 2002]. Ovaj se modul prilagođava učeniku prema njegovim potrebama i stanju modela učenika. Može omogućiti prilagođavanje različitih razina pomoći koju pruža učeniku tijekom učenja. Regian [Regian, 1997] ovaj modul naziva "modeliranjem poučavanja" (eng. instructional modeling). Npr. vođeno učenje putem otkrivanja (eng. guided discovery learning), kada učenik samostalno "otkriva" novo znanje dok ga pri tome ITS nenametljivo vodi ili usmjerava, može

se zamijeniti poučavanjem tipa "jedan-prema-jedan" (eng. one-on-one tutoring) ako učenik zatraži posebnu aktivnost ili mu je potrebno ponavljanje [Woolf *et al.*, 2003].

*Modul područnog znanja* sadrži područno znanje koje ITS koristi za poučavanje tj. informacije koje učenik uči. Ovaj modul je najvažniji jer ITS bez njega ne bi imao što poučavati [Beck *et al.*, 1996]. Potrebno je predstaviti područno znanje na takav način da mu mogu pristupiti ostali dijelovi ITS-a i da mu se lako dodaju nove informacije. Predstavljanje znanja (eng. knowledge representation) je potpodručje umjetne inteligencije o oblikovanju i korištenju sustava za spremanje znanja, činjenica i pravila o nekom subjektu (<http://computing-dictionary.thefreedictionary.com>). Baza područnog znanja (eng. domain knowledge base) je dio modula područnog znanja. Dane su neke definicije baze područnog znanja ili kraće *baze znanja*:

- Baza znanja je vrsta baze podataka koja se koristi za rad sa znanjem (<http://en.wikipedia.org>).
- Baza znanja je skup znanja izražen nekim formalnim jezikom za predstavljanje znanja (<http://computing-dictionary.thefreedictionary.com>).

*Komunikacijski modul* se odnosi na komunikaciju ili međusobnu interakciju između učenika i ITS-a. Komunikaciju između sustava i učenika kontrolira sustav ili učenik, ovisno o tome tko postavlja pitanja tj. da li učenik uči ili sustav provjerava njegovo znanje. ITS bi trebao biti sposoban inicirati interakcije s učenikom, te učinkovito interpretirati i odgovoriti na interakcije koje inicira učenik. Regian [Regian, 1997] ovu komponentu sustava naziva kao sposobnost vođenja dijaloga "mješovite inicijative" (eng. mixed initiative dialog). Pri tome mješovita inicijativa označava da dijalog tj. komunikaciju, kao što je već navedeno, može započeti sustav ili učenik. Također se u literaturi spominje i mogućnost komunikacije ITS-a i učenika prirodnim jezikom. Ministarstvo obrane SAD-a (eng. Department of Defense, kratica DoD) ulaže velika financijska sredstva u razvoj CAI i ITS-ova, a dio odlazi i na istraživanje mogućnosti takvog načina komunikacije. Regian [Regian, 1997] navodi da u laboratorijima DoD-a postoje prototipovi sustava koji koriste komunikaciju prirodnim jezikom, iako u usko ograničenim područjima.

*Model eksperta* se ovisno o literaturi, ponekad smatra dijelom područnog znanja. Prema [Debak, 2002] razlikuje se od područnog znanja ili prikaza podataka jer je to model o tome koliko netko, odnosno ekspert, zna o određenom području. Usporedbom modela eksperta i modela učenika, sustav utvrđuje znanje učenika.

Prema [Regian, 1997] ITS bi uz gore navedena svojstva trebao biti i generativan (eng. generative) tj. trebao bi generirati materijal koji koristi za poučavanje na osnovi uspjeha učenika. To se može primijeniti npr. i na pitanja koja sustav postavlja učeniku prilikom ispitivanja, što nije isto kao kad već postoje unaprijed pripremljena pitanja od kojih se zatim bira jedno po jedno. Park i Lee [Park & Lee, 2003] definiraju ITS-ove kao prilagodljive sustave za poučavanje koji su razvijeni primjenom metoda i tehnika umjetne inteligencije. Upravo metode umjetne inteligencije koje se koriste kod prikaza znanja (npr. semantičke mreže (eng. semantic networks), produkcijska pravila (eng. production rules)) omogućuju generativnost.

Svaki ITS nema sve gore navedene komponente, a kako i CAI sustavi mogu imati neke od njih, općenito se ITS-ovima smatraju sustavi koji imaju većinu komponenti. ITS-ovi su zasnovani na znanju o učeniku, pravilima zaključivanja o mogućim načinima poučavanja sadržaja i dinamičkom generiranju prilagođenih putova kroz znanje kako bi se odgovorilo na ponašanje učenika. Takvi sustavi mogu zaključivati o spremljenom znanju, davati povratne informacije, pomagati i otkrivati greške. ITS-ovi imaju mnoge prednosti za učenje i



poučavanje jer se prilagođavaju posebnim potrebama različitih učenika, a u idealnom slučaju trebali bi biti slični individualiziranom poučavanju kada je učitelj čovjek [Woolf *et al.*, 2003]. Potrebe učenika ovise o njihovim karakteristikama i sposobnostima. Učenje putem otkrivanja više odgovara boljim učenicima tj. učenicima s više *prirodne sklonosti za učenje* (eng. aptitude), dok slabiji učenici zahtijevaju više interakcije odnosno pomoći [Park & Lee, 2003]. Kod učenja uz pomoć simulacije također postoji određena razina pomoći (eng. level of guidance) kojom sustav vodi učenika kako bi simulacija bila učinkovita. Opet je učenicima s manje prirodne sklonosti za učenje potrebno više vođenja nego učenicima s više prirodne sklonosti kojima veća razina vođenja ne odgovara, a čak može biti i kontraproduktivna [Regian, 1997]. Prema ovim primjerima, čini se da postoji određena interakcija između prirodne sklonosti za učenje i načina poučavanja. *Prirodna sklonost za učenje* i njezina interakcija s načinom poučavanja su opisani u slijedećem poglavlju.

### 1.1.1. Interakcija prirodne sklonosti za učenje i načina poučavanja

Cotton [Cotton, 1989] definira *prirodnu sklonost za učenje* (eng. aptitude) kao količinu vremena koja je učeniku potrebna kako bi nešto naučio ili izvršio određeni zadatak pod optimalnim uvjetima poučavanja. Međutim, ta definicija nije u potpunosti korektna.

Cronbach i Snow [Cronbach & Snow 1977; prema Park & Lee, 2003] definiraju *prirodnu sklonost za učenje* (eng. aptitude) kao bilo koju individualnu karakteristiku koja povećava ili umanjuje učenikovu mogućnost uspjeha u danom načinu poučavanja (npr. poučavanje uz pomoć ITS-a). Cronbach 1957. godine [Cronbach, 1957; prema Park & Lee, 2003] ističe kako je poučavanje potrebno prilagoditi razlikama među učenicima. Predlaže da se jedan tip poučavanja koristi za učenika s određenim karakteristikama, a potpuno drugačiji tip poučavanja za drugog učenika s drugačijim karakteristikama. Ta Cronbach-ova strategija se naziva *interakcija prirodne sklonosti za učenje i načina poučavanja* (eng. Aptitude Treatment Interaction, kratica ATI) [Park & Lee, 2003].

Prema [Snow & Swanson, 1992; prema Park & Lee, 2003] sklonosti možemo podijeliti na *spoznajne sklonosti* (eng. cognitive aptitudes) i *osjećajne sklonosti* (eng. affective aptitudes). Spoznajne sklonosti se odnose na:

- intelektualne sposobnosti (eng. intellectual ability) – Npr. logičko razmišljanje, verbalne vještine i sl.
- spoznajne stilove (eng. cognitive styles) – To su načini opažanja, pamćenja, rješavanja problema, razmišljanja i donošenja odluka [Snow & Lohman, 1984; prema Park & Lee, 2003].
- stilove učenja (eng. learning styles) – Misli se na načine na koje učenici uče. Učenici mogu koristiti globalni pristup (usvajaju ukupni koncept, a zatim dodaju detalje), ili mogu koristiti linearni pristup (koncentriraju se na detalje i procedure, uče serijski) ili kombinaciju oba pristupa [Pask, 1976, 1988; prema Park & Lee, 2003].
- prethodno znanje (eng. prior knowledge) – Prethodno znanje ili predznanje koje učenik ima je važno za predviđanje njegovog uspjeha i potreba tijekom poučavanja.

Osjećajne sklonosti uključuju:

- anksioznost (eng. anxiety) – Učenici koji osjećaju veliki strah od ispitivanja ili testa ostvaruju slabiji uspjeh u odnosu na učenike čiji je strah mali.
- motivacija (eng. achievement motivation) – Motivacija može biti unutarnja (eng. intrinsic) ili vanjska (eng. extrinsically). Učenici s unutarnjom motivacijom ostvaruju

bolji uspjeh od onih s vanjskom motivacijom [Lin & McKeachie, 1999; prema Park & Lee, 2003]. To su često stariji učenici koji žele ostvariti uspjeh. Učenici s vanjskom motivacijom uče zato što moraju, motivirani su npr. nagradom ili kaznom, a to su obično djeca.

- samoučinkovitost (eng. self-efficacy) – To je učenikova procjena vlastite sposobnosti ostvarivanja nekog zadatka. Mijenja se ovisno o uspjehu ili neuspjehu kojeg učenik doživljava.

Iako je još 1957. godine Cronbach predložio svoju strategiju, ne postoji mnogo istraživanja s konzistentnim rezultatima u vezi ATI-ja. Teško je izmjeriti individualne karakteristike, a i valjanost testova postaje upitna jer ako se poučavanje prilagođava potrebama učenika, onda bi se i testovi također trebali prilagođavati [Park & Lee, 2003]. Kao što je prije navedeno, u sustavu DTEEx-Sys postoji prilagođavanje pitanja prema odgovorima učenika, ali ako se radi o istraživanju učinkovitosti sustava (opisano u Poglavlju 3.) onda se svi moraju ispitati istim pismenim testovima.

Tobias 1976 [Tobias 1976; prema Park & Lee, 2003] predlaže drugačiji tj. malo jednostavniji pristup kojeg naziva *interakcija postignuća i načina poučavanja* (eng. Achievement-Treatment Interaction). Ispituje se interakcija između prethodnog postignuća i uspjeha kojeg učenik ostvaruje nakon određenog načina poučavanja. Prethodno postignuće se određuje pre-testom. Sposobnosti koje je potrebno ispitati su kod ovog pristupa mnogo jasnije. Kod nekih istraživanja se za ovaj pristup također koristi naziv ATI (Aptitude Treatment Interaction) iako ATI pokriva šire područje što se tiče karakteristika učenika. Npr. Albacete i VanLehn [Albacete & VanLehn, 2000] u svom istraživanju ispituju interakciju postignuća i načina poučavanja (poučavanja uz pomoć inteligentnog tutorskog sustava *Conceptual Helper*), ali to nazivaju ATI. Odredili su medijan (srednju vrijednost) svih rezultata eksperimentalne grupe na pre-testu. Učenike koji su ostvarili uspjeh iznad medijana su nazvali "boljim", a učenike s uspjehom ispod medijana "slabijim". Zatim su ispitali razliku u postignuću (eng. gain) između boljih i lošijih učenika, a pokazalo se da su slabiji učenici više napredovali. Istraživači su bili zadovoljni rezultatima zato što je slabijim učenicima ipak potrebno više pomoći ili podrške tijekom poučavanja. Tu hipotezu potvrđuju i druga istraživanja [Tobias, 1973; prema Park & Lee, 2003].

Međutim, iako je ovaj pristup jednostavniji za primjenu od ATI pristupa i dalje nije riješen problem promjena učenikovih karakteristika ili sposobnosti koje se događaju tijekom samog procesa poučavanja. Kod tog problema djelomično pomažu ITS-ovi jer mogu bilježiti akcije učenika tijekom procesa učenja i poučavanja koje spremaju u *model učenika*, pa se tako npr. može odrediti kako je učenik došao do rješenja zadanog problema. Ono što ipak nedostaje, za razliku od ATI pristupa, su informacije o motivaciji, intelektualnim sposobnostima, strahu od testa i sl. Model učenika je opisan u sljedećem poglavlju.

### 1.1.2. Modeliranje učenika

Za ITS-ove je važno prilagođavanje curriculum-a i podrška interaktivnom načinu rješavanja problema, posebno za učenje na daljinu jer tu učitelj nije prisutan kako bi vodio učenika tijekom učenja [Shareef & Kinshuk, 2003]. Zato se kod ITS-ova koristi modeliranje učenika kako bi se postigla prilagodljivost (eng. adaptivity). Za svakog učenika se stvara individualni model koji određuje trenutno znanje učenika, a zatim ITS prema tome prilagođava curriculum i pomaže učeniku.

Informacije koje se nalaze u modelu učenika se mogu podijeliti na dvije vrste: informacije specifične za područno znanje (eng. domain-specific information) i informacije neovisne o područnom znanju (eng. domain-independent information). *Informacije specifične za područno znanje* predstavljaju trenutno stanje i razinu znanja učenika u vezi određenog koncepta [Brusilovsky, 1994; prema Shareef & Kinshuk, 2003]. *Informacije neovisne o područnom znanju* mogu uključivati ciljeve učenja, način na koji učenik uči (npr. kako odgovara, postavlja pitanja, reagira na pomoć), motivaciju učenika, i sl.

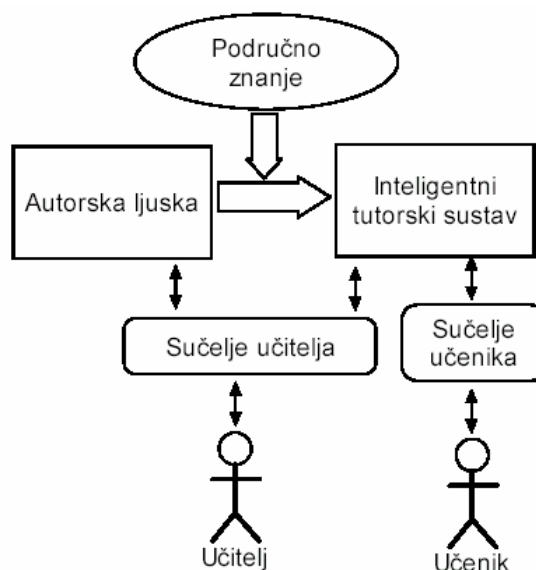
Model učenika bi trebao omogućiti samostalnost ITS-a, sa što manje intervencija čovjeka. Tijekom vremena dok učenik uči, ITS bi trebao pratiti što se događa i u skladu s tim usklađivati model učenika. Odabir modela učenika ovisi o područnom znanju te o tome što se kod učenika želi modelirati određenim ITS-om [Zhou & Evens, 1999]. Npr. mogu se procjenjivati vještine učenika u rješavanju problema ili promatrati greške koje učenik radi.

ITS-ovi najčešće imaju dvije baze za spremanje znanja, jedna sadrži znanje eksperta, a druga znanje učenika [Stankov, 1996]. Najčešće vrste razlika između tih baza znanja su: nedostatak znanja (eng. missing conception) i pogrešno znanje (eng. misconception). Kod *nedostatka znanja*, znanje učenika je pravi podskup znanja eksperta tj. razlikuju se u onome što učeniku nedostaje, a takav model učenika se naziva *model s prekrivanjem* (eng. overlay model) [Debak, 2002]. Nedostatak tog pristupa je što znanje učenika ne mora biti podskup znanja eksperta tj. učenici mogu neke pojmove ili područja pogrešno shvaćati. U tom slučaju se radi o *pogrešnom znanju*, a model učenika se zove *model s prekrivanjem i "buggy" proširenjem*. Model učenika tada sadrži i knjižnicu grešaka (eng. bugs) ili pogrešnih odgovora koje učenik može napraviti.

Prema [Stankov, 1996] modeliranje učenika također uključuje i mjerenje učenikovog znanja. ITS bi trebao provjeriti znanje i vještine učenika te zaključiti kako učenik razumije dano gradivo kako bi mogao prilagoditi poučavanje na odgovarajući način. Takvo zaključivanje se naziva postavljanje dijagnoze [Amižić *et al.*, 2002]. Sam model učenika je skup podataka, a dijagnoza je manipulacija tim podacima kako bi se odredilo znanje učenika.

Kao i svaka druga tehnologija, ITS-ovi se s vremenom usavršavaju i pronalaze put do sve većeg broja korisnika. Logičan slijed događaja je omogućavanje jednostavnije izgradnje ITS-ova. Tu na scenu dolaze *hipermedijske autorske ljuske za izgradnju inteligentnih tutorskih sustava* [Stankov, 2003]. S njihovom pojavom, izgradnja ITS-ova postaje bliža i neprogramerima. Autorska ljuska (eng. authoring shell) je alat za razvoj ITS-ova koji se obično sastoji od korisničkog sučelja, tutorskog modula i struktura za rad s područnim znanjem, ali bez područnog znanja tj. *autorska ljuska* "zna kako poučavati", ali joj nedostaju materijali za poučavanje [Smith, 1997; Goodkovsky, 1997].

Na Slici 1-2 se vidi odnos ITS-a i autorske ljuske. Sustav Tutor-Expert System (TEx-Sys) [Stankov, 1997; Stankov *et al.*, 2001] je inteligentna hipermedijska autorska ljuska za izgradnju inteligentnih tutorskih sustava u po volji odabranom područnom znanju.



Slika 1-2: Autorska ljuska i inteligentni tutorski sustav [Stankov, 2003]

## 1.2. Struktura sustava TEx-Sys

Inteligentna autorska ljuska Tutor-Expert System (TEx-Sys) omogućuje učenicima stjecanje znanja i vještina, zatim mjerenje i dijagnosticiranje njihovog znanja, te ocjenjivanje s preporukom za daljnji rad. Struktura sustava je modularna tj. sastoji se od sljedećih modula:

- 1) Prijavni modul (*Login*) – služi za registraciju korisnika za rad u sustavu.
- 2) Ljuska za izgradnju ITS-ova (*Developing* ili *T-Expert*) – omogućuje izgradnju baze proizvoljno odabranog područnog znanja.
- 3) Modul za učenje i poučavanje (*Learning ang Teaching*) – omogućuje učenje i poučavanje na odabranoj bazi područnog znanja.
- 4) Modul za ispitivanje (*Testing*) – omogućuje vrednovanje znanja učenika.
- 5) Modul za ocjenjivanje i pregled učenikovih ocjena (*Evaluation*) – omogućuje dijagnozu učenikovog znanja i pregled ocjena.
- 6) Modul kviz (*Quiz*) – znanje učenika se ispituje putem testova (ili "kvizova") u kojima se učeniku nudi skup pitanja s ponuđenim točnim i netočnim odgovorima. Učenik bira odgovore koje smatra točnima, a nakon rješavanja, test se ocjenjuje.

U sustavu TEx-Sys moguće je izgrađivati baze područnog znanja bez poznavanja programiranja, a sustav pruža podršku pri učenju i poučavanju na tim bazama. Baza područnog znanja koja je korištena za istraživanje opisano u Poglavlju 3. je izgrađena uz pomoć sustava TEx-Sys. Znanje u sustavu TEx-Sys je prikazano u obliku semantičkih mreža s okvirima. Osnovni elementi semantičkih mreža s okvirima su čvorovi (eng. nodes) i veze (eng. links). Čvorovi služe za prikaz objekata ili koncepata područnog znanja, a vezama se prikazuje odnos između objekata. Svakom čvoru se mogu pridružiti neki od strukturnih atributa kao što su: slika, zvuk, video, tekst, URL adresa, prezentacije MS Powerpoint-a i sl. Naravno, moguće su i kombinacije navedenih elemenata. Strukturnim atributima se dodatno opisuju objekti. Za svaki je čvor moguće dodati i opis u obliku hiperteksta čiji je unos i ažuriranje osiguran u modulu *T-Expert*. Dodavanje ostalih gore navedenih strukturnih atributa je također osigurano u modulu *T-Expert*, a izrađuju se npr. uz pomoć programskih alata iz okruženja Microsoft Windows-a.

Sve češće učitelji i učenici u procesu učenja i poučavanja postaju i fizički udaljeni [Stankov, 2003]. Napredak računalne komunikacije tj. računalnih mreža, omogućuje povećanje fizičke udaljenosti. Dakle, učenje i poučavanje postaje distribuirano, a u skladu s tim ITS-ovi postaju Web-orijentirani.

### ***1.3. Web-orijentirani inteligentni tutorski sustavi***

Distribuirano učenje i poučavanje je oblik poučavanja koji ne zahtijeva da se učenici okupljaju u određeno vrijeme, na određenim mjestima radi učenja. Učitelj (misli se na čovjeka) može biti prisutan u isto vrijeme kad i učenici, ali i ne mora. Poučavanje se može događati bilo kad i bilo gdje pa takve tehnologije možemo smatrati i asinkronima [Fletcher, 2003]. Pod pojmom distribuiranog učenja i poučavanja ili učenja i poučavanja na daljinu danas se većinom misli na Web-orijentirano poučavanje.

Općenito, obrazovanje na daljinu je poučavanje koje se dostavlja do jednog ili više udaljenih korisnika koji se nalaze na jednoj ili više različitih lokacija [Phipps, Wellman, & Merisotis 1998; prema Lewis *et al.*, 1999]. Prema toj definiciji obrazovanje na daljinu postoji još od 1830-tih s pojavom dopisnih tečajeva [Hanson *et al.*, 1997; prema Lewis *et al.*, 1999].

Međutim, učenici nisu previše oduševljeni tradicionalnim učenjem na daljinu jer nije dovoljno privlačno i nema mnogo interakcija, ali Web-orijentirano poučavanje bi se u tom smislu moglo pokazati mnogo boljim. [Yang *et al.*, 2002]. Web-orijentirani ITS-ovi postaju glavni predmet istraživanja. Mogu se instalirati i održavati s jednog mjesta dok ih koriste tisuće učenika širom svijeta [Brusilovsky *et al.*, 1996; prema Yang *et al.*, 2002]. Arhitektura klasičnih ITS-ova se može koristiti i za Web-orijentirane ITS-ove.

Web-orijentirano poučavanje nudi učenicima pristup obrazovnim sadržajima kakav ne mogu dobiti u klasičnom poučavanju u razredu, te omogućuje interaktivno i učinkovito poučavanje [Kahn, 2001; prema Wisher & Olson, 2003]. Takav način poučavanja omogućuje i brzo povezivanje s drugim izvorima znanja na Internetu i učenje "bilo kad, bilo gdje" [Fletcher & Dodds, 2001; prema Wisher & Olson, 2003]. Učenici putem Interneta mogu međusobno komunicirati i zatražiti savjete od predavača. Web postaje moćan alat za učenje i poučavanje na daljinu. Mogućnosti za on-line učenje su neograničene. Web-orijentirano obrazovanje prelazi ograničenja vremena i prostora, omogućujući studentima pristup sadržajima učenja danju i noću iz svakog dijela svijeta. Sve više se uključuje i u klasično predavanje kada predavač npr. na Web-u objavljuje zadatke ili projekte koje učenici moraju napraviti, njihove ocjene, popis literature i sl. U ovom se slučaju radi o kombiniranim ili "mješovitim" predavanjima (eng. blend courses) koja ne možemo vrednovati na isti način kao predavanja koja su potpuno on-line.

#### ***1.3.1. Preporuke za Web-orijentirano poučavanje***

Istraživači iz instituta "The Institute for Higher Education Policy" [Phipps *et al.*, 2000] su provodili istraživanje o preporukama (eng. benchmarks) koje su se koristile pri razvoju općenitog obrazovanja na daljinu kako bi odredili može li se nešto od toga primijeniti i na Web-orijentirano učenje i poučavanje u procesu obrazovanja na akademskoj razini. Najprije su prikupili podatke iz literature koji se odnose na učenje na daljinu, a zatim su potražili fakultete koji su bili spremni na suradnju. Posjetili su šest fakulteta, vidjeli čime se bave, obavili ankete i razgovore. Na kraju su prema rezultatima istraživanja izdvojili neke od

preporuka koje su smatrali najvažnijima za kvalitetu Web-orijentiranog poučavanja. Preporuke su rasporedili u sedam kategorija:

#### 1. *Preporuke podrške institucije* (eng. Institutional Support Benchmarks)

Preporuke u ovoj kategoriji uključuju one aktivnosti institucije (npr. fakulteta) koje pomažu osiguravanju okoline u kojoj je moguće održavati kvalitetu obrazovanja na daljinu i planove koji potiču razvoj Web-orijentiranog učenja i poučavanja. Također se odnose i na probleme tehnološke infrastrukture, te poticaja predavača jer možda svi predavači neće biti oduševljeni novim pristupom koji zahtjeva određene promjene. Potrebno je:

- Planirati tehnologije za elektroničke mjere sigurnosti (npr. pristup s lozinkom, šifriranje, zaštitne kopije).
- Osigurati pouzdanost tehnologije koja se koristi.
- Osigurati centralizirani sustav za podršku izgradnje i održavanja infrastrukture sustava za obrazovanje na daljinu.

#### 2. *Preporuke razvoja kolegija* (eng. Course Development Benchmarks)

Ova kategorija uključuje preporuke za razvoj curriculum-a kojeg izrađuju sami predavači, grupa predavača na fakultetu ili stručnjaci odgovarajućih područja.

- Potrebne su upute o tome koji su minimalni standardi za razvoj i oblikovanje kolegija, te predstavljanje sadržaja poučavanja. Rezultati učenja, a ne dostupnost tehnologije određuju koja će se tehnologija koristiti za predstavljanje sadržaja.
- Materijali koji se koriste za predavanje se povremeno moraju pregledavati kako bi se osigurali standardi.
- Kolegiji se organiziraju na način da se i studenti uključuju u analizu, sintezu i procjenu onoga što se od njih zahtijeva u tom kolegiju.

#### 3. *Preporuke učenja/poučavanja* (eng. Teaching/Learning Benchmarks)

U ovoj se kategoriji radi o nizu aktivnosti povezanih s pedagogijom i načinom predavanja. Uključuju se preporuke za interaktivnost, suradnju i modularno učenje.

- Važna karakteristika učenja i poučavanja na daljinu je međusobna interakcija studenata te interakcija studenata i predavača (npr. putem elektroničke pošte).
- Potrebno je omogućiti studentima dobivanje povratne informacije u vezi zadataka koje obave ili pitanja koja postavljaju predavačima, na konstruktivan način i u pravo vrijeme.
- Studente je potrebno uputiti na odgovarajuće metode istraživanja i procjene valjanosti izvora podataka koji su im potrebni za zadatke.

#### 4. *Preporuke strukture kolegija* (eng. Course Structure Benchmarks)

Ove preporuke se odnose na proces učenja i poučavanja. Uključuju ciljeve kolegija, dostupnost izvora podataka, vrste materijala koji se daju studentima, vrijeme davanja odgovora studentima na njihove upite, te očekivanja studenata. Preporuke uključuju sljedeće:

- Prije nego što Web-orijentirano predavanje započne, potrebno je ispitati jesu li studenti motivirani za učenje na daljinu (da li to žele) i imaju li pristup minimalnoj tehnologiji koja je potrebna za praćenje kolegija.
- Studentima se daju dodatne informacije o kolegiju s opisom ciljeva, koncepata, ideja i rezultata koji se očekuju, u pismenom obliku, na jasan i izravan način.
- Studenti bi trebali imati pristup izvorima podataka koji mogu uključiti "virtualne knjižnice" dostupne na Web-u.

- Predavači i studenti bi se trebali složiti o očekivanim vremenima koja su potrebna za završavanje zadataka i dobivanja povratne informacije od predavača.

#### 5. Preporuke podrške studentima (eng. Student Support Benchmarks)

U ovoj se kategoriji uključuje obučavanje i pružanje pomoći studentima za korištenje Interneta.

- Treba poučiti i pomoći studentima pri traženju informacija iz elektroničkih baza podataka i drugih izvora informacija.
- Tijekom trajanja kolegija studenti trebaju imati pristup tehničkoj podršci, uključujući detaljne upute u vezi elektroničkog medija kojeg koriste, potrebno im je omogućiti prilike za vježbanje prije nego što kolegij započne (kako ne bi imali problema prilikom učenja). Naravno, tu se uključuje i pristup osoblju za tehničku podršku.
- Na pitanja studenata je potrebno odgovarati brzo.

#### 6. Preporuke podrške predavačima (eng. Faculty Support Benchmarks)

Kako svi predavači nemaju vještine i znanje potrebne za učenje na daljinu putem Interneta, potrebno im je pomoći pri tome, uključujući podršku za prijelaz s klasičnog poučavanja u razredu na potpuno on-line predavanje, te pomoć tijekom trajanja cijelog perioda poučavanja.

- Tehnička podrška treba biti dostupna predavačima tijekom pripremanja ili oblikovanja kolegija.
- Potrebno je pomoći predavačima prilikom prijelaza s klasičnog načina predavanja na on-line način predavanja, te procijeniti da li to rade na odgovarajući način.
- Obučavanje i pomoć predavačima uključuje i pomoć njihovih kolega koji su bolje upoznati s takvim načinom predavanja, ako je potrebno i tijekom cijelog vremena trajanja kolegija.
- Potrebno je predavačima dati pismene izvore podataka kako bi se mogli nositi s problemima koji nastaju kad studenti koriste elektronički dostupne podatke.

#### 7. Preporuke za vrednovanje(eng. Evaluation and Assessment Benchmarks)

Ove se preporuke odnose na procedure kojima institucije vrednuju (ako vrednuju) Web-orijentirano poučavanje na daljinu. Uključuju vrednovanje dobivenih rezultata.

- Vrednuje se učinkovitost učenja i poučavanja. Određuju se metode i standardi za proces vrednovanja.
- Podaci o troškovima, uključenim predavačima i studentima, te rezultatima se koriste za vrednovanje učinkovitosti programa.
- Rezultati učenja se redovito pregledavaju. Trebaju biti jasni.

Prema [Phipps *et al.*, 2000] namjena istraživanja je bila pronaći preporuke ili upute koje bi mogle biti korisne osobama i organizacijama koje se bave razvojem učenja na daljinu uz pomoć Interneta. Jedan od predavača s fakulteta na kojem su autori provodili istraživanje je istaknuo kako izgleda da bi za obrazovanje na daljinu trebalo više nadgledanja, vrednovanja i detaljnog izrađivanja kolegija nego što se inače radi kod klasičnog načina predavanja, a jedan od studenata je izjavio kako on-line predavanja zahtijevaju više discipline, ali omogućuju rad vlastitim tempom. Studenti moraju bolje razumjeti uloge i odgovornosti koje imaju prilikom on-line predavanja. Npr. kod klasičnog načina predavanja u razredu se često treba brinuti o tome hoće li studenti dolaziti na predavanja i izvršavati svoje obaveze. Ako im se pri tome omogući određena sloboda npr. da ne moraju dolaziti na predavanja ako

ne žele, onda oni to često i koriste ne imajući u vidu eventualne probleme prilikom učenja za ispit, a ako će se predavač brinuti o njihovim dolascima, onda preuzima dio odgovornosti na sebe.

Web-orijentirani inteligentni tutorski sustavi pokazuju prednosti naspram inteligentnih tutorskih sustava jer prelaze ograničenja prostora i vremena, pa tako postaju sve zanimljiviji. Međutim, teško se pronalaze podaci o provedenim istraživanjima Web-orijentiranih ITS-ova jer ih u ovom trenutku vjerojatno nema tako mnogo kao npr. za CAI sustave ili ih možda ima, ali nisu lako dostupni (pokušavali smo pretraživanjem Interneta pronaći više podataka, ali bez nekog uspjeha). U svakom slučaju, ako već ne možemo učiti na iskustvu drugih, onda smo se odlučili na samostalno istraživanje konkretnog Web-orijentiranog ITS-a u stvarnim uvjetima. Za ovaj rad smo odabrali Web-orijentirani inteligentni tutorski sustav DTEEx-Sys.

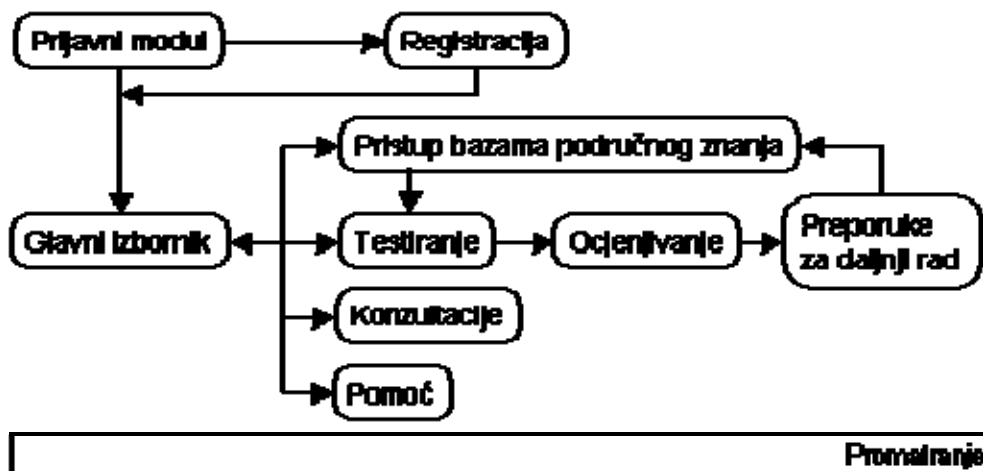
#### **1.4. Struktura sustava DTEEx-Sys**

Često Web-orijentirani ITS-ovi nastaju nadogradnjom jednokorisničkih ili samostalnih (eng. stand-alone) sustava [Yang *et al.*, 2002]. Na sličan način je i sustav DTEEx-Sys nastao na naslijeđu sustava TEx-Sys. Distributed Tutor Expert System (DTEEx-Sys) je Web-orijentirani inteligentni tutorski sustav [Rosić *et al.*, 2001]. DTEEx-Sys ne ovisi o platformi, vremenu i mjestu pristupa, ima jedinstveni način spremanja informacija za različita područna znanja, mogućnost prilagođavanja potrebama korisnika, učenje u hipermedijskoj okolini, mogućnost praćenja napretka učenika, te mogućnost testiranja i savjeta za nastavak. Sustav je zasnovan na oponašanju ljudskih učitelja i omogućuje svakom učeniku učenje u bilo kojem trenutku, a također učenici mogu koristiti usluge sustava koliko im je potrebno kako bi stekli željenu razinu znanja.

Korisnici sustava mogu odabrati sljedeće usluge: pristup bazama područnog znanja, testiranje znanja ili zatražiti rezultate testova. Modul testiranja se prilagođava učeniku procjenjujući svaki korak testiranja. Nakon svakog koraka sustav generira nova pitanja ovisno o djelomičnim rezultatima učenika, te na kraju uz ocjenu daje i preporuke za daljnji rad. Sustav DTEEx-Sys je generativan jer generira pitanja tijekom izvođenja, a težina pitanja se određuje ovisno o učenikovim odgovorima.

Na Slici 1-3 se vidi funkcionalna struktura sustava. Znanje je predstavljeno u obliku semantičkih mreža s okvirima (kao i u sustavu TEx-Sys). Sustav također podržava slike, animacije, prezentacije, URL adrese i opise u hipertekstu.



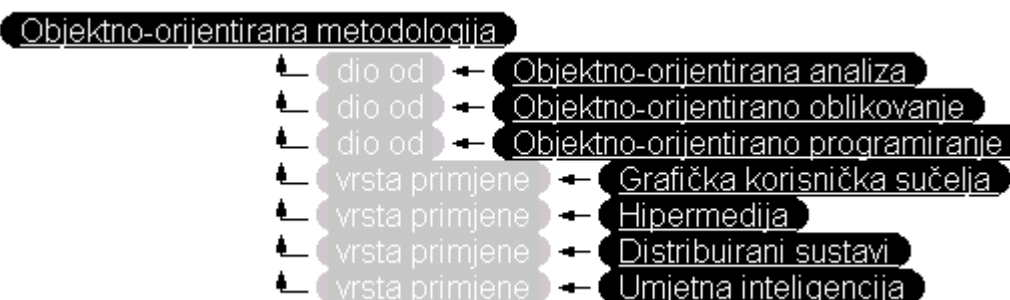
Slika 1-3: Funkcionalna struktura sustava DTEEx-Sys [Rosić *et al.*, 2001]

Arhitektura sustava DTEEx-Sys:

DTEEx-Sys je implementiran kao 3-slojna klijent-server arhitektura [Rosić *et al.*, 2001]:

- 1) Korisničko sučelje – Sloj korisničkog sučelja sustava DTEEx-Sys pruža sučelje za pristup prijavi na sustav, bazama znanja, rješavanju testa i pregledu ocjena koje učenici postižu.
- 2) Logika aplikacije – Srednji sloj generira dokument koji se distribuira korisniku prema njegovom zahtjevu, ovisno što korisnik želi raditi (npr. učiti, rješavati test).
- 3) Baze podataka – Sloj baza podataka se sastoji od baza područnog znanja i baza podataka koje sadrže podatke o korisnicima sustava. Znanje je predstavljeno semantičkim mrežama s okvirima čiji su osnovni elementi čvorovi i veze. Čvorovi se koriste za predstavljanje objekata područnog znanja, a veze prikazuju odnose među njima (kao i kod sustava TEx-Sys). Baze podataka o korisnicima sustava sadrže elementarne podatke o korisnicima, njihovom ponašanju pri korištenju sustava i rezultatima njihovog testiranja.

Na Slici 1-4 je prikazan primjer kako izgleda semantička mreža u sustavu DTEEx-Sys (ono što korisnici vide).



Slika 1-4: Primjer

Za svaki ITS kojeg se namjerava koristiti kod učenja i poučavanja potrebno je ispitati njegovu učinkovitost, zatim radi li ono što bi trebao raditi i kako bi trebao raditi. Uvijek se teži prema poboljšanju učenja i poučavanja pa je važno otkriti da li određeni ITS pri tome pomaže ili ne. Ako se npr. dokaže da učenici koji inače slušaju klasično predavanje u razredu postižu bolje rezultate uz primjenu ITS-a, onda postoji opravdan razlog za njegovu upotrebu, a također je prema rezultatima istraživanja moguće od više ITS-ova odabrati učinkovitiji.

---

ITS-ovi se provjeravaju i dok traje njihov razvoj kako bi se uočili nedostaci, planirali daljnji koraci i sl. U sljedećem se poglavlju opisuju neki od načina vrednovanja inteligentnih tutorskih sustava. Sve osobe ili organizacije koje se bave izradom ITS-ova također nastoje i ispitati njihovu učinkovitost, a prikupljanjem i analizom podataka o provedenim ispitivanjima pojedinih ITS-ova dobiva se učinkovitost ITS-ova općenito.

## 2. VREDNOVANJE POSTIGNUĆA U RADU S INTELIGENTNIM TUTORSKIM SUSTAVIMA

Prema [Woolf *et al.*, 2003] vrednovanja postignuća učenika s inteligentnim tutorskim sustavima pokazuju da se uz njih može postići poboljšanje postignuća koje je blizu poboljšanju koje se postiže individualiziranim poučavanjem. Postoji mnogo istraživanja o učinkovitosti CAI sustava. Prema [Regian, 1997] pokazalo se da je prosječno poboljšanje postignuća kod CAI sustava od 50-tog do 65-og centila tj. *veličina učinka* (eng. effect size, kratica ES) je 0.39 standardnih devijacija, dok je ušteda vremena potrebnog za postizanje ciljeva poučavanja 24%. Za ITS-ove postoji manji broj istraživanja, a pokazalo se poboljšanje od 50-tog do 84-og centila, tj. veličina učinka je 1.00 standardna devijacija, a ušteda vremena 55%. Prema [Woolf *et al.*, 2003] može se za CAI sa sigurnošću reći da rade prilično dobro u različitim uvjetima i različitim područjima poučavanja. Također se može reći i da ITS-ovi rade dobro, mogu biti znatno učinkovitiji od CAI, a mogli bi biti puno više učinkoviti. Što se tiče uštede vremena potrebnog za postizanje ciljeva poučavanja, opet se može sa određenom sigurnošću reći da CAI smanjuju to vrijeme u različitim uvjetima za mnogo različitih područnih znanja, a ITS-ovi također smanjuju vrijeme, ali vjerojatno više od tradicionalnih CAI sustava. Prema jednom primjeru [Anderson, 1990, Anderson & Reiser, 1985; prema Woolf *et al.*, 2003], studenti koji su koristili LISP tutor na Sveučilištu Carnegie Mellon, su završili vježbe iz programiranja za 30% manje vremena od onih koji su koristili klasično poučavanje u razredu i postigli su 43% više uspjeha na završnom ispitu.

Iako ovi rezultati ohrabruju, ostaje činjenica da testovi u razredu često nisu odgovarajuća mjera uspjeha ovih sustava jer materijal nije isti kao kod klasičnog poučavanja u razredu i sadržaj sustava se ne može lako integrirati u klasični curriculum. Npr. prema [Anderson *et al.*, 1985, 1995; prema Woolf *et al.*, 2003] tutori *Geometry* i *Algebra* su u područnim znanjima iz geometrije i algebre orijentirani na rješavanje problema dok su se računske operacije (npr. zbrajanje i oduzimanje) obavljale automatski. Klasični curriculum u razredu se više bavio računskim operacijama ili manipulacijom simbola (npr. + i -). Zbog toga je bilo teško na kraju procijeniti rezultate učenja.

### 2.1. Pregled provedenih istraživanja

Nema mnogo dostupnih istraživanja koja se bave istraživanjem učinkovitosti ITS-ova, pa se ovdje navode i istraživanja koja se bave učinkovitošću CAI sustava. Yaakub [Yaakub, 1998] je došao do rezultata da su ICAI sustavi (tj. ITS-ovi) učinkovitiji od CAI sustava. Prema tome, možemo od ITS-ova očekivati čak i bolje rezultate.

Za početak je bilo potrebno istražiti da li općenito CAI nešto radi, tj. da li uopće poučava, a ne je li dobra za poučavanje i poučava li ono što je potrebno. Istraživanja su pokazala da zaista "nešto radi". Npr. istraživanja koje su provodili: Crotty (1984), Verano (1987) i Allan (1989) [prema Fletcher, 2003]. Oni su uspoređivali interaktivno multimedijско poučavanje s placebo načinima poučavanja gdje nisu predstavljani materijali za poučavanje. Prosječna *veličina efekta* je bila 1.38 standardnih devijacija (ili sigma), što otprilike predstavlja prosječno povećanje postignuća od 50-og do 92-og centila. Dakle, CAI ipak utječe na poučavanje.

Često se u literaturi spominje "ukupna veličina učinka" (eng. overall effect size) CAI sustava i ITS-ova. Naravno da ne postoji neki općeniti CAI ili ITS od kojeg bi se izračunala

ukupna veličina učinka. Ta veličina se dobije iz rezultata više istraživanja postupkom koji se naziva *meta-analiza*.

Pojam "meta-analiza" (eng. meta-analysis) je izmislio Gene V. Glass 1976 [Glass, 1976; prema Yaakub, 1998]. Glass definira meta-analizu kao "analizu analizâ" (eng. analysis of analyses). Meta-analizom se analiziraju rezultati istraživanja koja su već provedena, ali nisu unaprijed planirana radi toga. Podaci koji se analiziraju su obično aritmetičke sredine, standardne devijacije, te rezultati statističkih analiza, a ne izvorni podaci (npr. uspjeh nekog učenika na post-testu).

Meta-analiza omogućuje istraživačima skupljanje rezultata mnogih istraživanja koja pokušavaju odgovoriti na zajedničko pitanje, npr. kao što je učinak upotrebe CAI u poučavanju, te veličinu tog učinka u jedinicama standardnih devijacija, upotrebom zajedničke mjere koja se naziva "*veličina učinka*" (eng. effect size). Cooper i Hedges (1994) [prema Yaakub, 1998] definiraju meta-analizu kao statističku analizu skupa rezultata analiza iz individualnih istraživanja sa svrhom integriranja nalaza.

Istraživači koji odluče provesti meta-analizu moraju koristiti objektivne metode za pronalaženje odgovarajućih istraživanja koje će analizirati [Glass, McGaw i Smith, 1981; prema Yaakub, 1998]. Kod odabira istraživanja koja će se analizirati meta-analizom ima smisla uzeti u obzir samo ona istraživanja koja sadrže sve potrebne podatke, te potpuni tekst. Istraživanja koja se uspoređuju moraju imati iste nezavisne varijable (npr. metode poučavanja: ITS i klasično poučavanje u razredu) i zavisne varijable (npr. postignuće učenika). Važno je da istraživanja nemaju grešaka. Npr. ako ispitanici nisu slučajnim odabirom podijeljeni na eksperimentalne i kontrolne grupe ili ako nisu korišteni pre-test i post-test te statistička analiza razlika među grupama, onda takva istraživanja ne dolaze u obzir. Također je bitno ispitati valjanost i pouzdanost mjerenja tj. testova kojima se ispituje postignuće učenika. Kontrolna i eksperimentalna grupa se moraju ispitati istim testovima. Istraživanja je moguće pronaći računalnim pretraživanjem baza podataka [Yaakub, 1998], ali za vojna istraživanja u SAD-u je potrebna komplicirana procedura.

Yaakub je kod svoje meta-analize istraživanja koja se bave usporedbom postignuća učenika uz pomoć CAI i klasičnog predavanja pronašao veličine učinka od  $-0.09$  do  $1.13$  [Yaakub, 1998]. Jedno istraživanje je imalo veličinu učinka  $2.92$  što je izvan tog intervala pa je eliminirano iz meta-analize. Ukupna veličina učinka koju je dobio Yaakub je  $0.35$  (CAI u odnosu na klasično poučavanje). Pri tome okolina u kojoj se provodilo poučavanje uz pomoć CAI (npr. srednja škola, fakultet, vojska) nije imala statistički značajnog utjecaja na postignuće.

### *2.1.1. Povećanje učinkovitosti poučavanja uz pomoć računala*

Tipičan način istraživanja učinkovitosti je usporedba konvencionalnog poučavanja s predavanjem, tekstualnim materijalima i pristupa uz pomoć CAI s interaktivnim multimedijским poučavanjem. Mnoga su se istraživanja time bavila, a ni jedno istraživanje samo ne može dokazati je li CAI općenito učinkovito. Potrebno je uzeti u obzir rezultate različitih istraživanja koja su provedena u različitim uvjetima i okolnostima, te ih usporediti meta-analizom [Yaakub, 1998]. U Tablici 2-1 su prikazane veličine učinka dobivene meta-analizom.

Tablica 2-1: Neke veličine učinka za CAI i ITS sustave  
[Fletcher, 1995; prema Regian, 1997] & [Shute, 1996; prema Regian, 1997]

Postavke poučavanja	Broj istraživanja	Veličina učinka	Centil rezultata
Osnovna škola CAI	28	0.47	68
Srednja škola CAI	42	0.42	66
Visoko obrazovanje CAI	101	0.26	60
Obrazovanje odraslih CAI	24	0.42	66
Vojna obuka CAI	38	0.40	66
<b>Ukupno CAI</b>	<b>233</b>	<b>0.39</b>	<b>65</b>
Visoko obrazovanje ITS	1	0.97	83
Vojna obuka ITS	1	1.02	84
Srednja škola ITS	1	1.00	84
<b>Ukupno ITS</b>	<b>3</b>	<b>1.00</b>	<b>84</b>

Cohen i Dacanay [Cohen & Dacanay, 1992; prema Yaakub, 1998] su analizom 37 istraživanja CBI dobili veličinu učinka 0.41, a Fletcher-Flinn i Gravatt [Fletcher-Flinn i Gravatt, 1995; prema Yaakub, 1998] su analizirali 120 istraživanja CAI i dobili veličinu učinka 0.24.

### 2.1.2. Smanjenje vremena potrebnog za postizanje ciljeva poučavanja

Provođena su istraživanja i o tome može li se smanjiti vrijeme koje je potrebno da bi učenici nešto usvojili. Podaci su u Tablici 2-2:

Tablica 2-2: Smanjenje vremena potrebnog za postizanje ciljeva poučavanja za CAI i ITS  
[Fletcher, 1995; prema Regian, 1997] & [Regian & Shute, 1990; prema Regian, 1997].

Postavke poučavanja	Broj istraživanja	% smanjenja vremena
Vojna obuka CAI	23	28
Vojna obuka CAI	--	30
Visoko obrazovanje CAI	17	34
Obrazovanje odraslih CAI	15	24
<b>Ukupno CAI</b>	<b>55+</b>	<b>29</b>
Visoko obrazovanje ITS	3	55

Prema istraživanjima se pokazalo da se u prosjeku uštedi oko 30% vremena. Ako se ta ušteda preračuna u dolare, npr. za Ministarstvo obrane SAD-a (eng. Department of Defense, kratica DoD), koji troše 4 milijarde dolara godišnje za obučavanje posebnih vještina (npr. tehničare i operatere na radaru, mehaničare za posebna vojna vozila), ušteda od 30% na samo 20% ljudi koje obučavaju iznosi preko 250 milijuna dolara godišnje. Prema tome, smanjenje vremena koje je potrebno za poučavanje je povezano sa smanjenjem troškova poučavanja.

### 2.1.3. Trošak poučavanja uz pomoć računala

Glavno pitanje koje zanima istraživače je često o tome je li novi pristup bolji u odnosu na postojeći tj. postižu li se s njim veća postignuća. Međutim, istraživači često nisu oni koji donose odluke o tome hoće li se to usvojiti, pa treba razmišljati i o tome što se mora uložiti da bi se poboljšalo poučavanje.

Isplativost se može procijeniti na dva načina:

- može se za konstantne troškove procjenjivati različite sustave koji povećavaju učinkovitost
- ili se može za konstantnu učinkovitost procjenjivati različite sustave kojima bi se mogli smanjiti troškovi

Tablica 2-3 pokazuje da su najisplativiji pristupi poučavanja uz pomoć računala i međusobno poučavanje učenika:

Tablica 2-3: Prosječni troškovi poboljšanja rezultata u matematici po jednu standardnu devijaciju [Fletcher, 2003]

<b>Način poučavanja:</b>	<b>Trošak \$</b>
Tutorsko (20 min dnevno)	
Učenici tutori	472
Odrasli tutori	2654
Smanjenje razreda od:	
35 do 30	1619
35 do 20	2251
Povećanje vremena poučavanja 30 min dnevno	4391
CAI 10 min dnevno:	
Mini-računala u laboratoriju (1976)	
3. razred	382
5. razred	785
Mikroračunala u razredima (1990)	
3. razred	316
5. razred	339

## 2.2. Metode vrednovanja inteligentnih tutorskih sustava

S porastom broja ITS-ova koji se razvijaju i uvode u obrazovanje, raste i potreba za njihovim vrednovanjem [Mark & Greer, 1993]. U literaturi postoji određeni broj metoda vrednovanja koje se mogu koristiti, a često se teško odlučiti koja bi metoda odgovarala za određenu namjenu. Većina istraživača ITS-ova se manje bavila vrednovanjem, a više potencijalima ITS-ova i problemima implementacije pri izradi komponenti. Iako vrednovanje nije lak posao, omogućuje nam određivanje utjecaja ITS-ova na osobe koje uče. Littman i Soloway (1988) [prema Iqbal *et al.*, 1999] ističu kako im je vrednovanje sustava PROUST pomoglo pri shvaćanju kako početnici uče programirati, kako poučavati programiranje i kako izgraditi ITS-ove koji zaista poučavaju.

*Formativno vrednovanje* (eng. formative evaluation) se provodi tijekom oblikovanja i ranih faza razvoja projekta kako bi se poboljšalo oblikovanje i ponašanje sustava. Preporučuje se u ranim fazama razvoja prije nego što se uloži više vremena i sredstava. Rezultati se zatim koriste kao osnova za daljnji razvoj sustava, a ne za neke tvrdnje o sustavu. Odnosi se na pitanje: "Koja je veza između arhitekture ITS-a i njegovog ponašanja?" [Littman & Soloway, 1988; prema Mark & Greer, 1993].

*Sumativno vrednovanje* (eng. summative evaluation) se bavi vrednovanjem cijelih sustava. Odgovara na drugo važno pitanje: "Koji je obrazovni utjecaj ITS-a na učenike?" [Littman & Soloway, 1988; prema Mark & Greer, 1993]. Pokušava se dokazati ili opovrgnuti neka formalna tvrdnja o sustavu ili tehnikama koje se u njemu koriste.

ITS-ovi su složeni sustavi i može ih se vrednovati kao cijele sustave, po komponentama i po određenim značajkama. Metode koje odgovaraju za vrednovanje cijelih sustava, možda neće biti dobre za vrednovanje komponenti (i obrnuto) [Mark & Greer, 1993]. Postoje metode koje su razvijene u drugim područjima, a mogu se primijeniti ili se primjenjuju na ITS-ove.

### 2.2.1. Dokazi ispravnosti (eng. proofs of correctness)

Ispituje se ispunjava li sustav zahtjeve i ciljeve ili postoji li odnos između strukture i ponašanja sustava i njegovih specifikacija [Iqbal *et al.*, 1999]. Ova metoda ne odgovara za programe iz umjetne inteligencije koji rade s analitički neuhvatljivim problemima, predstavljenim nepotpuno određenim funkcijama [Partridge, 1986; prema Mark & Greer, 1993], pa prema tome nije dobra ni za vrednovanje ITS-ova. Možda se može iskoristiti za komponente ili dijelove sustava koje nisu toliko vezane uz umjetnu inteligenciju.

### 2.2.2. Vrednovanje na osnovi kriterija (eng. criterion-based evaluation)

Sustav se smatra uspješnim ako ne pokazuje velike nedostatke unutar okoline koja je namijenjena za njegovu primjenu [Partridge, 1986; prema Mark & Greer, 1993]. Osoba koja vrednuje sustav mora odrediti koji su to "veliki nedostaci", prema tome kako se čini da sustav ispunjava svoje zahtjeve i specifikacije. Najbolje odgovara za ranije faze razvoja sustava, gdje se radi o općenitijim karakteristikama, a ne o preciznim detaljima ili točno određenim aspektima sustava. Preporučuje se odrediti neke općenite smjernice tj. napravi se popis zahtjeva i specifikacija koji se zatim provjeravaju kako bi se ispitali nedostaci. Prema [Mark

& Greer, 1993] te smjernice mogu biti korisne za usmjeravanje razvoja ITS-ova, ali bi se učinkovitost programa ili njegovih svojstava trebala potvrditi na neki drugi način. Iqbal [Iqbal *et al.*, 1999] smatra da se ta metoda može koristiti za vrednovanje komponenti ITS-a i cijelih sustava. Ford (1988) [prema Iqbal *et al.*, 1999] je koristio ovu metodu za vrednovanje ITS-a.

### 2.2.3. Znanje i ponašanje eksperta (*eng. expert knowledge and behaviour*)

Često se za vrednovanje sustava kao standard koristi znanje eksperta [Gasching *et al.*, 1983; prema Mark & Greer, 1993] kako bi se procijenilo da li program zadovoljava standardnu razinu performansi [Iqbal *et al.*, 1999]. Da bi ekspert mogao pregledati sustav, ponašanje sustava mora biti konzistentno i predvidljivo. Za ITS je primjena ove metode za vrednovanje ograničena jer je ponašanje ITS-a složeno i dinamično, pa nije moguće sve intuitivno razumjeti i pregledati. Ponekad se primjenjuje samo na komponente ITS-a.

### 2.2.4. Potvrđivanje (*eng. certification*)

Određivanje kompetentnih sustava za poučavanje bi se moglo zasnivati na metodama određivanja kompetentnih ljudskih učitelja. Možda bi ljudski učitelji mogli procijeniti ITS-ove, tj. njihove prednosti i nedostatke. Ovdje se radi o standardima za vrednovanje programa, kriterijima koji se mogu koristiti za vrednovanje sustava i komponenti, a također i točnosti s kojom ljudi mogu odrediti koji su učinkoviti programi za obrazovanje. Dok to ne bude potpuno jasno, neće se moći primijeniti na ITS-ove [Mark & Greer, 1993].

### 2.2.5. Analiza osjetljivosti (*eng. sensitivity analysis*)

Pregledava se sustav ili komponenta sustava prema tome na koji način ponašanje odgovara na različite informacije koje sustav prima. To je važno za ITS-ove koji bi se trebali prilagođavati različitim karakteristikama individualnih učenika. Osjetljivost sustava na te različite karakteristike bi mogla pokazati treba li sustav dodatno poboljšati. Sustav koji prikazuje slične odgovore na prilično različite informacije nije poželjan. Ako se mogu precizno odrediti mjere osjetljivosti i ako im se može dodijeliti značenje (koja su ponašanja poželjna, kada ih treba pozvati), onda se ova metoda može primijeniti za sumativno vrednovanje, a inače se primjenjuje za formativno vrednovanje [Mark & Greer, 1993]. Odgovara za vrednovanje i komponenti i cijelog sustava [Iqbal *et al.*, 1999].

### 2.2.6. 'Pilot' testiranje (*eng. pilot testing*)

Ako autori sustava mogu pretpostaviti da će korisnici biti slične stručnosti i iskustva kao i oni, onda mogu izraditi sustav prema svojim potrebama i biti prilično sigurni da će to odgovarati i korisnicima. Međutim, jasno je da se korisnici mogu razlikovati prema svom znanju o tutorskom poučavanju i tutorima, te osnovnim spoznajnim sposobnostima. Stvarni sustav bi se trebao ispitati na stvarnim korisnicima kako bi se otkrili eventualni problemi i neočekivani rezultati. Gagné *et al.* (1988) i Golas (1983) [prema Mark & Greer, 1993] određuju tri tipa 'pilot' testiranja: jedan-prema-jedan testiranje, testiranje malih grupa,



testiranje u stvarnim uvjetima (eng. field testing). Kod *jedan-prema-jedan testiranja*, promatraju se interakcije učenika s materijalom za poučavanje koji se razvija (otkrivaju se neodgovarajuća očekivanja; nejasni smjerovi, pitanja i informacije), a obično se provodi u ranim fazama razvoja. *Testiranje malih grupa* se obično provodi na malim grupama učenika koji predstavljaju ciljnu populaciju. Izvodi se u kasnijim fazama razvoja, kada se program i njegov sadržaj stabiliziraju. *Testiranje u stvarnim uvjetima* ispituje upotrebu sustava sa stvarnim predavačima i učenicima u stvarnom okruženju, kada je sustav gotovo završen. Nastoji se otkriti kakvi se problemi događaju kada se sustav uvodi u okolinu gdje će se zapravo koristiti [Hoecker & Elias, 1986; prema Mark & Greer, 1993].

### 2.2.7. Eksperimentalno istraživanje (eng. experimental research)

Eksperimentalno istraživanje je uobičajena metoda u psihologiji i obrazovanju. Prema [Mark & Greer, 1993] ova metoda odgovara za obrazovne sustave kao što su ITS-ovi jer se mogu ispitati veze između akcija koje se poduzimaju u poučavanju i odgovarajućih rezultata koje učenici postižu u odnosu na te akcije, a također se može odrediti i koliko su takve veze značajne. Najprije je potrebno odrediti što se želi istraživati, pa se zatim odrede hipoteze (npr. da će nakon nekog načina poučavanja ili tretmana postojati značajne razlike među grupama). Hipotezu treba biti moguće testirati, te potvrditi ili opovrgnuti na osnovi određenih uvjeta i rezultata istraživanja. Zatim se određuje oblik istraživanja kako bi istraživač mogao ispitati hipotezu. Istraživač nakon provedenog istraživanja analizira podatke. Ako rezultati ne potvrđuju hipotezu, bilo bi dobro da istraživači, ako je moguće, predlože moguća objašnjenja za svoje rezultate.

## 2.3. Kriteriji za vrednovanje utjecaja inteligentnih tutorskih sustava na proces učenja

Glavni cilj sustava za poučavanje je poučavanje učenika, a glavni način provjere sustava za poučavanje je ispitivanje da li učenici učinkovito uče uz pomoć tog sustava [Hasselbring, 1986; prema Mark & Greer, 1993]. Kriteriji koji se koriste za vrednovanje utjecaja ITS-ova na obrazovanje uključuju: postignuće i afekt. Postignuće (eng. achievement) se odnosi na stjecanje, razumijevanje, performanse, zadržavanje i prijenos učenikovog znanja i vještina [Haertel & Calfee, 1983; prema Mark & Greer, 1993]. Afektivne mjere se odnose na stavove i emocije koje mogu utjecati na način kako učenici koriste ITS-ove i uče iz njih [Malone, 1981; prema Mark & Greer, 1993].

### 2.3.1. Kriteriji za određivanje mjera postignuća

- Kategorije znanja

Znanje se može podijeliti na kategorije: referencijalno, činjenično i proceduralno. Znanje o simbolima i značenjima se naziva *referencijalno* znanje [Mark & Greer, 1993]. *Činjenično* znanje je znanje o objektima i vezama između objekata unutar svijeta. *Proceduralno* znanje je o tome kako nešto raditi. Dijeli se na eksplicitno i implicitno. Eksplicitno proceduralno znanje se može opisati u terminima algoritama, pravila ili procedura [Mayer, 1987; prema Mark & Greer, 1993]. Implicitno proceduralno znanje je znanje o tome kako nešto raditi što se ne može opisati riječima. Mayer to naziva vještinama [Mark & Greer,

1993]. Metaznanje [McGraw & Harbison-Briggs, 1989; prema Mark & Greer, 1993] je strateško znanje ili strategije učenja [O'Neil, 1978; prema Mark & Greer, 1993].

- Razine razumijevanja ili rezultati učenja

Bloom [Bloom, 1956; prema Mark & Greer, 1993] je napravio klasifikaciju ciljeva obrazovanja, općih kategorija rezultata učenja koje se mogu povezati s pokazateljima performansi. Određuje šest glavnih tipova ciljeva koje naziva znanje, razumijevanje, primjena, analiza, sinteza i vrednovanje. Prema [Mark & Greer, 1993] Bloomova 'razina znanja' se naziva *prisjećanje* pa će se dalje koristiti taj termin kako bi bilo jasnije o čemu se radi, a označava mogućnost učenika da se prisjeti nekog termina, procedure ili činjenice, ali se ne radi o razumijevanju i primjeni naučenog. *Razumijevanje* označava da učenik može do određenog stupnja koristiti naučeni materijal, što je dovoljno za definicije i izvlačenje izravnih zaključaka. Učenik počinje upotrebljavati znanje u konkretnim situacijama na razini *primjene*. *Analiza* označava da učenik može odrediti ideje, ali i ispitati njihove međusobne veze. *Sinteza* označava da učenik može organizirati znanje za generiranje novih ideja. *Vrednovanje* uključuje mogućnost procjene vrijednosti znanja.

Kategorije znanja i razine razumijevanja mogu služiti kao kriteriji pri određivanju mjera koje će se koristiti za vrednovanje postignuća.

- Ostali kriteriji u vezi postignuća

Drugi kriteriji koji se mogu koristiti pri vrednovanju obrazovnog utjecaja uključuju prijenos, zadržavanje, vrijeme učenja i brzinu završavanja. Prva tri kriterija se mogu procijeniti eksperimentalno, a brzine završavanja se promatraju tijekom isprobavanja sustava.

*Prijenos* je sposobnost primjene informacija ili vještina naučenih u jednom kontekstu u novi i obično nepoznati kontekst, gdje je to znanje također važno [Mayer, 1987; prema Mark & Greer, 1993]. Prijenos podrazumijeva generalizaciju ili apstraktno razumijevanje materijala. Važno je da kad učenik uči nešto npr. u simuliranoj okolini da to može primijeniti i u stvarnoj okolini. Učenika se testira u situaciji u kojoj uči i u novoj situaciji kada to znanje treba primijeniti [Johnson *et al.*, 1987; prema Mark & Greer, 1993].

*Zadržavanje* je sposobnost zadržavanja naučenog tijekom vremena [Mayer, 1987; prema Mark & Greer, 1993]. Smatra se da je učenik bolje naučio neko znanje što se duže može toga prisjetiti. To bi također bio dokaz prednosti nekog pristupa poučavanju. Ispituje se na način da se mjeri isto znanje tijekom vremena. Npr. dva testa iz matematike, koji naravno nisu isti, ali se mogu usporediti, se daju učenicima u dva različita vremenska razdoblja (možda na početku i kraju polugodišta), kako bi se vidjelo je li se njihovo znanje promijenilo tijekom vremena, neovisno o poučavanju.

*Vrijeme učenja* je vrijeme potrebno da bi se nešto naučilo [Mark & Greer, 1993]. Ponekad se vrijeme učenja na različitim sustavima uspoređuje kako bi se pokazalo je li jedan sustav bolji u odnosu na drugi. Vrijeme učenja se mora gledati i u kontekstu postignuća i motivacije.

*Brzina završavanja* (eng. completion rate) ili *odustajanje* (eng. drop-out) se promatra prilikom testiranja sustava u stvarnim uvjetima kada učenici mogu prestati ili nastaviti koristiti program za poučavanje [Hoecker & Elias, 1986; prema Mark & Greer, 1993]. Npr. ako je neki sustav težak za upotrebu onda može i obeshrabriti učenike, te ih tako navesti da odustanu. Čak i ako učenici koji koriste takav sustav postižu bolje rezultate, a većina učenika

ga smatra nepoželjnim i teškim za upotrebu, onda ipak nije dobar izbor za poučavanje. Također, može se umjesto pojma "odustajanje" reći da učenici osjećaju *otpor* (eng. attrition). [Phipps & Merisotis, 1999, Terry, 2001; prema Wisher & Olson, 2003].

### 2.3.2. Afektivne mjere (eng. affective measures)

Pri vrednovanju učinkovitosti ITS-ova treba uzeti u obzir i motivaciju ili želju učenika da budu aktivni i uključeni u poučavanje. Treba razmišljati o interesu učenika za uključivanjem u aktivnosti poučavanja, interesu za postizanje nekog uspjeha ili željenog cilja i o stavovima prema obrazovanju, računalima ili određenim programima [Thomas, 1979; prema Mark & Greer, 1993]. Uobičajeni način vrednovanja je da se zatraži od učenika da ocjene koliko im odgovaraju određene aktivnosti [Moore, 1985; prema Mark & Greer, 1993]. Također se uspoređuje i vrijeme provedeno na aktivnostima koje su vezane i koje nisu vezane uz zadatak. Učenici bi trebali imati želju za učenjem i uključivanjem u poučavanje, trebali bi biti aktivni i zaključivati o materijalu koji im se prikazuje [Woolf *et al.*, 2003].

Afektivne mjere se mogu koristiti da bi se vidjelo što učenici misle o sustavu, ali one nužno ne odražavaju postignuće. Mogu se koristiti za određivanje problema, ali nisu dovoljno pouzdane da bi mogli biti sigurni u određivanje postojećih problema. Prema [Mark & Greer, 1993] mogu se koristiti kao prijedlog o prihvaćanju ili neprihvaćanju ITS-a. Mogu se koristiti kao dodatne mjere prilikom vrednovanja postignuća, ali nisu uvjerljiv dokaz o učinkovitosti ako se koriste same.

## 2.4. Učinkovitost Web-orijentiranog poučavanja

Prema [Phipps *et al.*, 2000] istraživanje koje je provodio američki "Odjel nacionalnog obrazovnog centra za statistiku u obrazovanju" (eng. U.S. Department of Education's National Center for Education Statistics, kratica NCES) pokazuje da se od 1994-95 do 1997-98 broj programa za učenje na daljinu povećao za 72%. Dodatnih 20 % institucija je pokazalo namjeru uvođenja programa učenja na daljinu. Procjenjuje se da je više od 1.6 milijuna studenata bilo uključeno u neki oblik predavanja obrazovanja na daljinu 1997-98 godine u SAD-u. Dakle, raste zanimanje za Web-orijentirano poučavanje, a također raste i potreba za njegovim vrednovanjem. Prilikom procesa vrednovanja, odnosno provođenja istraživanja, potrebno je izbjegavati nedostatke koji mogu utjecati na konačne rezultate.

### 2.4.1. Nedostaci istraživanja Web-orijentiranog učenja i poučavanja

Obično se prilikom ispitivanja učinkovitosti Web-orijentiranog obrazovanja na daljinu u procesu obrazovanja na akademskoj razini vrednuju rezultati studenata (ocjene i rezultati testova), te stavovi studenata o učenju na sustavima za obrazovanje na daljinu [Phipps & Merisotis, 1999]. Često kod takvih istraživanja dolazi do propusta ili nedostataka. Phipps i Merisotis [Phipps & Merisotis, 1999] su pregledali 40 istraživanja koja su se bavila učenjem na daljinu. Ovdje se učenje na daljinu odnosi na pojedine kolegije ili cijele akademske programe koji su održani uz pomoć videa ili uz pomoć računala. Phipps i Merisotis su primijetili određene nedostatke i probleme koji se odnose na samu provedbu istraživanja učenja na daljinu, a nisu u bliskoj vezi s korištenom tehnologijom, pa iako se samo 28%

istraživanja odnosi na učenje uz pomoć računala tj. ono što nas ovdje najviše zanima, njihove zaključke možemo smatrati korisnima i za istraživanje Web-orijentiranog učenja i poučavanja na daljinu. Nedostaci uključuju sljedeće:

- Mnoga istraživanja ne kontroliraju vanjske varijable pa ne mogu pokazati uzrok i posljedice.

To se u prvom redu odnosi na eksperimentalna istraživanja kada istraživač nastoji usporediti rezultate eksperimentalne i kontrolne grupe. Želi se ispitati kako određena tehnologija (ovdje se odnosi na Web-orijentirane ITS-ove) tj. "uzrok" djeluje na rezultate učenja tj. "posljedicu". Da bi se dobio odgovarajući odnos, ne smije postojati neki drugi "uzrok" koji bi djelovao na rezultate.

- Mnoga istraživanja ne koriste slučajno odabrane subjekte.

Najbolji način kontroliranja vanjskih varijabli ili utjecaja je slučajni odabir pripadnika kontrolne i eksperimentalne grupe, što se kod nekih istraživanja ne radi.

- Valjanost i pouzdanost instrumenata koji se koriste za mjerenje rezultata i stavova studenata je upitna.

Dobro istraživanje bi trebalo uključivati valjanost i pouzdanost instrumenata (testova, kvizova, upitnika i sl.) kako bi čitatelj mogao vjerovati rezultatima. Gotovo nijedno istraživanje koja su pregledali Phipps i Merisotis [Phipps & Merisotis, 1999] nema te informacije.

- Mnoga istraživanja ne kontroliraju na odgovarajući način osjećaje i stavove studenata i predavača koji se nazivaju *reakcijski učinci* (eng. reactive effects).

Jedan od reakcijskih učinaka se naziva *Učinak novoga* (eng. novelty effect), a odnosi se na povećano zanimanje, motivaciju ili sudjelovanje sa strane studenata samo zbog toga što rade nešto novo ili različito, što ne mora biti nužno bolje. Drugi tip reakcijskog učinka koji se naziva "*John Henry*" – *učinak* odnosi se na činjenicu da se ponekad pripadnici kontrolne grupe osjećaju ugroženima ili možda smatraju suprotstavljenost novom pristupu izazovom, te daju sve od sebe kako bi bili što bolji, što se u normalnim okolnostima ne bi moglo očekivati. U mnogim istraživanjima nisu poduzete mjere protiv toga.

Samo jedno istraživanje od 40 koje su pregledali Phipps i Merisotis [Phipps & Merisotis, 1999] je pokušalo odrediti koje su karakteristike studenata povezane s njihovim uspjehom pri učenju uz pomoć računala. Npr. pokazalo se da su uspješniji:

- studenti koji su smatrali da postoje ozbiljne posljedice ako ne polože kolegij;
- studenti koji su bolje ocjenjivali svoje mogućnosti za uspjeh na studiju;
- studenti kojima nije bila potrebna pomoć kod izvršavanja zahtjevnijih zadataka i nisu smatrali da je važno razgovarati s kolegama u vezi kolegija;
- studenti s visokom razinom pismenosti;
- studenti koji su smatrali da su dobro organizirani što se tiče vremena tj. mogli su reći da imaju dovoljno vremena za obavljanje svega što im je potrebno;
- studenti koji su smatrali da se vrlo dobro pripremaju tj. uče za studij

Nadalje, Phipps i Merisotis [Phipps & Merisotis, 1999] ističu da je potrebno obratiti pozornost na sljedeće probleme:

- 1) Istraživanja naglašavaju rezultate studenata kod individualnih kolegija, a ne cijelog programa studija.

Većina istraživanja se odnosi na pojedinačne kolegije što povlači pitanje bi li Web-orijentirano učenje i poučavanje na daljinu bilo učinkovitije od klasičnog učenja i poučavanja u odnosu na cijeli akademski program. Potrebno je ispitati rezultate u odnosu na kritičko razmišljanje, stavove i vrijednosti itd.

- 2) Istraživanja ne uzimaju u obzir razlike među studentima.

Obično se ispituju razlike u postignućima između grupa koje koriste klasično predavanje u odnosu na novi pristup. Postoje mnoge razlike među studentima, kao što su spol, dob, motivacija i sl. Kada se skupi uzorak studenata koji se smatraju "prosječnima", ne primjećuju se njihove međusobne razlike. Istraživanja bi se trebala više baviti kako pojedinci uče, a ne kako grupe uče.

- 3) Istraživanja ne opisuju na odgovarajući način zašto su odustajanja studenata (eng. drop-out) veća za učenje na daljinu.

Prilikom pregleda istraživanja Phipps i Merisotis [Phipps & Merisotis, 1999] su utvrdili da kod nekih postoji viši postotak studenata koji odustaju (eng. drop-out) od on-line predavanja prije nego što je završeno, u odnosu na klasično predavanje. To se negativno odražava na istraživanje. Moguće je i da se iz istraživanja izbace rezultati studenata koji odustaju, ali time automatski u istraživanju ostaju oni "uspješniji". Wisher i Olson [Wisher & Olson, 2003] također ističu da je odustajanje veći problem za predavanja koja su potpuno on-line nego za predavanja koja su samo djelomično on-line. Međutim, i jedni i drugi se slažu da ne postoji dovoljno istraživanja koja se time bave, odnosno nema dovoljno dostupnih podataka.

- 4) Istraživanja ne uzimaju u obzir kako različiti stilovi učenja svakog studenta utječu na upotrebu određene tehnologije.

Informacije o tome na koji način neki student uči utječu na način oblikovanja predavanja i odabir tehnologije koja će se koristiti. Karakteristike osobe koja uči su važne za ostvarivanje postignuća. Razumijevanje veze između učenika, onoga što uči i tehnologije je ograničeno.

- 5) Istraživanja su većinom usmjerena na utjecaj individualnih tehnologija, a ne na međusobnu interakciju više tehnologija.

Većina istraživanja se usmjerava na individualne tehnologije, ali većina tehnologija često ima višestruke funkcionalnosti (eng. multifunctional) te se mogu prilagoditi prema više različitih rezultata učenja. Postoji mali broj istraživanja koja ispituju utjecaj više tehnologija, te spomenutih tehnologija s višestrukim funkcionalnostima na rezultate učenja, a na osnovi tih istraživanja se ne mogu donositi općeniti zaključci (ili je uzorak premalen, ili nije dovoljno objašnjen postupak).

- 6) Istraživanja ne uključuju teoretsku osnovu.

Lakše je provoditi istraživanje koje se zasniva na teoriji, jer nam teorija omogućuje izgradnju na radu drugih ljudi, povećava se vjerojatnost otkrivanja značajnih stvari, a također omogućuje ponavljanje istraživanja što pridonosi općenitosti.

---

7) Istraživanja ne uzimaju u obzir učinkovitost digitalnih "knjižnica" (eng. digital libraries)

Studentima je omogućen pristup digitalnim "knjižnicama" koje sadrže razne izvore podataka. Za neke od njih se ističe da pružaju veliki izbor podataka, te da se mogu usporediti s klasičnim knjižnicama, ali se postavlja pitanje koliku podršku zaista mogu pružiti. Postoje i slučajevi kada je bilo potrebno mijenjati ciljeve curriculum-a jer nije bilo dovoljno dostupne literature.

Fletcher [Fletcher, 2003] ističe veličinu učinka koju ostvaruju ITS-ovi od 0.84 (na 11 pregledanih istraživanja), a prema [Regian, 1997] može se reći da ITS-ovi rade dobro u različitim uvjetima, mogli bi biti znatno učinkovitiji od CAI sustava, ali nema dovoljno dostupnih istraživanja da bi mogli biti potpuno sigurni u to. Ipak, ohrabreni dobrim rezultatima koje ITS-ovi postižu i sa željom da ispitamo sustav DTEEx-Sys, odlučili smo provesti vlastito istraživanje koje je opisano u slijedećem poglavlju.

### 3. PRISTUP VLASTITOM ISTRAŽIVANJU

U ovom radu se predstavlja istraživanje učinkovitosti Web-orijentiranog inteligentnog tutorskog sustava DTEEx-Sys (Distributed Tutor Expert System), kako bi se pokušalo odgovoriti na pitanje: "Koliko DTEEx-Sys utječe na poučavanje?". Odabirom metode eksperimentalnog istraživanja pokušavamo ustvrditi hoće li biti poboljšanja u postignućima studenata u odnosu na klasično predavanje u razredu.

#### 3.1. Planiranje i realizacija eksperimenta

Cilj istraživanja je ispitati je li DTEEx-Sys utjecao na poučavanje studenata u određenom područnom znanju i koliki je taj utjecaj, izraženo u standardnoj mjeri *veličine učinka* (eng. effect size).

Za uzorak na kojoj će se izvoditi eksperiment je odabrana grupa od ukupno 31 studenta druge godine koja se sastoji od studenata triju studijskih grupa: matematika i informatika, matematika, te informatika i tehnička kultura, a svi skupa slušaju kolegij Računalni praktikum I. Područno znanje na kojem su studenti testirani je objektno-orijentirana metodologija. Studenti su podijeljeni na eksperimentalnu (20 studenata) i kontrolnu grupu (11 studenata) slučajnim odabirom. (Ispitana je situacija s dvije eksperimentalne grupe, ali su dobiveni isti rezultati). Svi zajedno su slušali predavanje o objektno-orijentiranoj metodologiji dva školska sata (svaki po 45 min), gdje se nastojalo klasičnom metodom predavanja dati pregled osnovnih pojmova iz područja objektno-orijentirane metodologije (sadržaj predavanja u Prilogu A). Omogućen im je pristup sadržaju predavanja u tekstualnom obliku.

Za sedam dana su svi pisali pismeni pre-test koji se sastojao od 20 pitanja, a ocjenjivan je bodovima od 0-100 (pitanja iz testa su u Prilogu B), da bi se vidjelo ima li kakvih razlika u predznanju između kontrolne i eksperimentalne grupe.

Eksperimentalnoj grupi je sljedeći tjedan u trajanju od dva školska sata omogućen pristup sustavu DTEEx-Sys na računalima u učionici u svrhu učenja, a također i nakon tog vremena je bilo moguće pristupiti sustavu i učiti. Testiranje u tom razdoblju nije bilo moguće. Kontrolna grupa nije imala pristup, pa su oni nastavili učiti iz ranije dobivenih tekstualnih materijala.

Kao što je već spomenuto, u sustavu DTEEx-Sys se baze znanja, najjednostavnije rečeno, sastoje od čvorova i veza. Čvorovi su pojmovi ili koncepti, a veze pokazuju u kakvom su međusobnom odnosu ti koncepti. Baza znanja na kojoj su se studenti učili se sastoji od 202 čvora i 189 veza, a gotovo svi čvorovi imaju objašnjenje u obliku hiperteksta. Neki od čvorova imaju i dodatna objašnjenja u obliku slika ili prezentacija (ukupno 45 prezentacija i 45 slika). Prikaz čvorova i veza baze znanja se nalazi u Prilogu D.

Nakon tjedan dana, studenti eksperimentalne grupe su se testirali na sustavu DTEEx-Sys, a zatim su nakon toga zajedno sa studentima kontrolne grupe pisali post-test koji je imao 16 pitanja (13 teorijskih pitanja i 3 zadatka, u Prilogu C), kako bi se vidjelo ima li značajne razlike između eksperimentalne i kontrolne grupe nakon korištenja sustava. Test je također ocjenjivan bodovima od 0-100.

### 3.2. Analiza rezultata

Nakon završetka testiranja, potrebno je analizirati podatke. Potrebno je utvrditi je li bilo statistički značajnih razlika između kontrolne i eksperimentalne grupe u početnim uvjetima tj. rezultatima iz pre-testa, a zatim i u rezultatima post-testa. Prvi korak je postavljanje nul-hipoteze da među grupama nema značajnih razlika. Zatim se odabirom nekog statističkog testa ispituje nul-hipoteza, a takvi statistički testovi se nazivaju *testovima razlike* (eng. tests of difference). Nul-hipotezu je moguće prihvatiti ili odbaciti. Prihvaćanje nul-hipoteze znači da među grupama koje se ispituju ne postoje značajne razlike, a odbacivanje znači da razlike postoje. Prilikom ispitivanja nul-hipoteze mogu se pojaviti dva tipa grešaka:

- Greška tipa I – Odbacuje se nul-hipoteza koja je zapravo točna. Kod testova razlike to znači da se tvrdi kako razlika među grupama postoji, a zapravo te razlike nema. Obično se postavlja vrijednost koja označava vjerojatnost (eng. probability) greške tipa I. Ona se označava slovom  $p$ , a ako je  $p < 0.05$ , to znači da postoji 5% vjerojatnosti greške tipa I. Moguće je smanjiti mogućnost greške tipa I tako da se smanji ograničenje za  $p$  (npr. na 0.01, tj. 1% vjerojatnosti greške), ali u tom slučaju se povećava mogućnost greške tipa II.
- Greška tipa II – Ovaj tip greške se javlja ako se nije uspjelo odbaciti nul-hipotezu koja nije točna. U tom slučaju kod testova razlike se tvrdi da među grupama nema razlike, ali razlika zapravo postoji.

Kako bi se odabrao odgovarajući test, potrebno je razmisliti o kakvom se tipu podataka radi, tj. jesu li podaci iz nekog intervala (eng. interval data) ili se odnose na poredak (eng. ordinal data) [Becker, 1999]. Prema tim vrstama podataka, postoje i dvije grupe testova:

1. Za podatke iz intervala:
  - a) nezavisni t-test (eng. independent t-test)
  - b) zavisni t-test (eng. dependent t-test)
2. Za podatke poretka:
  - c) Mann-Whitney U test
  - d) Wilcoxon test

Općenito su testovi koji se provode na podacima iz intervala moćniji. Pošto su se u ovom istraživanju koristili testovi koji su ocjenjivani bodovima iz intervala od 0 do 100, onda će se koristiti jedan od navedenih t-testova. Ako su sudionici eksperimenta slučajnim odabirom podijeljeni na dvije različite grupe, koje se zatim uspoređuju, onda se radi o nezavisnim grupama, pa se prema tome bira i nezavisni t-test. U našem istraživanju, kao što je već navedeno, radi se o dvije nezavisne grupe, kontrolnoj i eksperimentalnoj, pa je odabran i nezavisni t-test, kojeg ćemo dalje u tekstu zvati t-test. T-test se može izračunati primjenom nekog statističkog računalnog programa. Ovdje je korišten program StatSoft Inc. STATISTICA 6.

#### 3.2.1. T-test

T-test je statistički test koji pomaže pri provjeri postoji li stvarna razlika između tretmana koji se testiraju. U našem slučaju tretmani su načini poučavanja. Računanjem t-testa dobiju se vrijednosti "t" i "p" [StatSoft, 2004]. Vrijednost  $p$  određuje vjerojatnost pogreške pri prihvaćanju nul-hipoteze. Za razinu značajnosti (eng. significance level) vrijednosti  $p$  se obično uzima vrijednost 0.05 kao minimalna vrijednost za koju je razlika dovoljno velika da



bi bila "statistički značajna" (eng. statistically significant). Drugim riječima, ako se razlika može dogoditi potpuno slučajno u više od 5 slučajeva od 100 ( $p > 0.05$ ), onda tu razliku ne smatramo statistički značajnom razlikom. Što je  $p$  manja, razlika je značajnija. Ako je npr.  $p = 0.001$  to znači da postoji mogućnost od 1 prema 1000 da se razlika dogodila sasvim slučajno. Dakle, ako je  $p < 0.05$  onda se nul-hipoteza odbacuje i zaključuje da postoje značajne razlike među grupama, a ako je  $p > 0.05$ , onda se prihvaća nul-hipoteza i zaključuje da nema značajnih razlika među grupama.

Nul-hipoteza za rezultate iz pre-testa je H1: *Nema značajnih razlika između eksperimentalne i kontrolne grupe u rezultatima pre-testa.* Kao što se vidi iz Tablice 3-1. vrijednost  $p = 0.74$  je veća od 0.05 pa se hipoteza H1 prihvaća i zaključuje da nema značajnih razlika između eksperimentalne i kontrolne grupe u početnim uvjetima.

Tablica 3-1: Rezultati t-testa za pre-test

	Vrijednosti t-testa	Značajna razlika
<b>Eksperimentalna vs. Kontrolna</b>	$t = 0.34$ $p = 0.74$	Ne

Postavili smo nul-hipotezu za post-test H2: *Nema značajnih razlika između eksperimentalne i kontrolne grupe u rezultatima post-testa.* Rezultati u Tablici 3-2. pokazuju da je  $p > 0.05$  pa se prihvaća nul-hipoteza i zaključuje da nema statistički značajnih razlika između kontrolne i eksperimentalne grupe u rezultatima post-testa.

Tablica 3-2: Rezultati t-testa za post-test i postignuće

	t-test (post-test)	t-test (postignuće)	Značajna razlika
<b>Eksperimentalna vs. Kontrolna</b>	$t = 0.58$ $p = 0.56$	$t = 0.29$ $p = 0.77$	Ne

Dakle, prema rezultatima t-testa možemo zaključiti da tretman eksperimentalne grupe tj. učenje na sustavu DTEx-Sys nije imalo značajnog utjecaja na njihov uspjeh. Bilo kakva razlika između eksperimentalne i kontrolne grupe je mogla nastati kao posljedica greške.

Test kojim se utvrđuje značajnost razlika između dviju grupa nije potpun bez provjere veličine učinka. Npr. ako se t-testom pokaže da između kontrolne i eksperimentalne grupe postoji statistički značajna razlika, onda se odbacuje nul-hipoteza i zaključuje da je tretman koji je imala eksperimentalna grupa imao značajan utjecaj na rezultate. Međutim, ne možemo vidjeti snagu veze između tretmana i postignuća u populaciji iz koje smo odabrali uzorak ispitanika. Iz tog razloga se koristi *veličina učinka*. Iako kod našeg istraživanja t-test pokazuje da nema statistički značajnih razlika, ipak ćemo izračunati veličinu učinka.

### 3.2.2. Veličina učinka

Veličina učinka je općenita mjera veličine utjecaja nekog novog načina poučavanja na eksperimentalnoj grupi, u odnosu na kontrolnu koja je koristila standardni način poučavanja, a računa se formulom [Yaakub, 1998]:

$$ES = \frac{AS_E - AS_K}{SD_K} \quad (3-1)$$

$AS_E$  = Aritmetička sredina postignuća eksperimentalne grupe

$AS_K$  = Aritmetička sredina postignuća kontrolne grupe

$SD_K$  = Standardna devijacija kontrolne grupe

Aritmetička sredina  $n$  brojeva ili rezultata  $A_1, \dots, A_n$  se računa formulom:

$$AS = \frac{A_1 + A_2 + \dots + A_n}{n} = \frac{1}{n} \sum_{i=1}^n A_i \quad (3-2)$$

Veličina učinka se može izračunati i kada nije baš jasno koja je grupa kontrolna, a koja eksperimentalna. Tada se vidi razlika između dviju grupa, ali je važno naglasiti kojim redom se radilo računanje. Također, prema [Yaakub, 1998] postoji i više načina za računanje veličine učinka, a preporučuje se formula (3-1) [Kulik & Kulik, 1989; prema Yaakub, 1998].

Standardna devijacija je standard za mjerenje varijabiliteta rezultata, najčešće se označava s grčkim slovom  $\sigma$  (sigma) ili SD [Petz, 2002], a računa se formulom:

$$\sigma = SD = \sqrt{\sum_i \frac{(X_i - M)^2}{N - 1}} \quad (3-3)$$

$X_i$  = izmjerena vrijednost ili rezultat

$M$  = aritmetička sredina rezultata

$N$  = broj rezultata

Formule za izračunavanje standardne devijacije su ugrađene u razne statističke alate i proračunske tablice što olakšava izračunavanje.

U Tablici 3-3 su prikazane vrijednosti aritmetičkih sredina i standardnih devijacija za testove i postignuće.

Tablica 3-3: : Aritmetičke sredine i standardne devijacije

Grupa	Pre-test	Post-test	Postignuće (eng. gain)
<b>Eksperimentalna</b>	AS= 59.35 SD= 26.00	AS= 46.05 SD= 23.89	AS= -13.3 SD= 17.93
<b>Kontrolna</b>	AS= 56.18 SD= 22.61	AS= 41.00 SD= 21.61	AS= -15.18 SD= 15.12

AS = aritmetička sredina; SD = standardna devijacija

Postignuće (eng. gain) je izračunato na način da se od rezultata koje su studenti postigli na post-testu oduzmu rezultati iz pre-testa. Vrijednosti aritmetičkih sredina i standardnih devijacija se koriste za izračunavanje veličine učinka.

Izračunali smo veličinu učinka (*ES*) po formuli (3-1) za naše istraživanje:

$$ES = \frac{-13.3 - (-15.18)}{15.12} = \mathbf{0.12} \quad (3-4)$$

Dakle, veličina učinka sustava DTE<sub>x</sub>-Sys koji je dobiven istraživanjem je 0.12. Prema [Pedhazur & Schmelkin, 1991; prema Yaakub, 1998] veličina učinka manja od 0.2 se smatra malom veličinom (eng. small), veličina učinka jednaka 0.5 se smatra srednjom (eng. medium), a veličine učinka veće ili jednake 0.8 su velike (eng. large) veličine učinka. Dobivena veličina učinka je mala, a kako smo očekivali značajnije rezultate, bilo je potrebno ispitati što je moglo utjecati na istraživanje. Prvi korak je provjera mjernih instrumenata tj. testova koji su se koristili za ispitivanje znanja studenata.

### 3.2.3. Metričke karakteristike testova

Kod učenja i poučavanja najčešće se ispituju stavovi i postignuća učenika [Phipps & Merisotis, 1999]. Obično se za mjerenje stavova (eng. attitudes) koriste upitnici, a za mjerenje postignuća testovi. U našem istraživanju je bilo potrebno izmjeriti postignuće učenika, pa smo odabrali testove kao instrumente za mjerenje. Testovi bi trebali točno mjeriti postignuće učenika, kako bi istraživanje imalo smisla. Dakle, potrebno je ispitati metričke karakteristike testova. Prema [Mužić, 1977] postoje slijedeće metričke karakteristike testova:

1. valjanost
2. pouzdanost
3. objektivnost
4. osjetljivost
5. diskriminativna vrijednost zadataka

Ovdje ćemo opisati samo one karakteristike koje se na određeni način odnose na naše istraživanje.

- Valjanost

Valjanost je karakteristika testa koja nam pokazuje da li i koliko test mjeri ono što s njim želimo mjeriti. Ne postoji "opća" valjanost testa, tj. valjanost se može odrediti samo u vezi konkretne namjene za koju se test koristi.

Jedan od kriterija valjanosti je nastavni program. To znači da bi se sadržaj i ciljevi nastave trebali slagati sa sadržajem testa. Ako se s testom točno ispituje ono što bi učenik trebao naučiti, onda je test valjan. Uobičajen postupak je određivanje ciljeva nastave koji se zatim uspoređuju sa sadržajem testa. Javlja se problem ako još ne postoji način na koji bi se ispitalo slaganje ciljeva i sadržaja nastave sa sadržajem testa. Većinom se testom ne ispituje cijeli sadržaj nastave nego samo jedan uzorak koji na što bolji način treba predstavljati cjelinu.

Prilikom provedbe našeg istraživanja, nastojalo se da se sadržaji testova podudaraju sa sadržajem predavanja o objektno-orijentiranoj metodologiji koje je trajalo dva školska sata. Prema tome, može se reći da se sadržaji testova i predavanja podudaraju.

Međutim, kriterij valjanosti ima nedostatak jer se ne može kvantitativno izraziti tj. ne postoji broj kojim bi se moglo izraziti razina slaganja sadržaja testa i nastavnog programa.

Tako nije moguće od više testova odabrati onaj koji se najviše slaže sa sadržajem, odnosno onaj kojeg bi trebalo primijeniti.

- Pouzdanost

Ova karakteristika određuje možemo li se i kojoj mjeri osloniti na rezultat koji se dobije testom tj. da li test točno mjeri, bez obzira što se mjeri. Pri određivanju pouzdanosti postoji nekoliko načina, a spomenut ćemo dva s kojima je ispitana pouzdanost testova iz našeg istraživanja.

Jedan od načina je određivanje *koeficijenta homogenosti testa*. Potrebno je test podijeliti na dva jednaka dijela i zatim ispitati njihovo slaganje. Nakon primjene testa, zadaci se podijele, obično na parne i neparne zadatke. Izračunavanjem korelacije između parnih i neparnih zadataka dobije se pokazatelj pouzdanosti. Nedostatak ovog pristupa je što se polazi od pretpostavke da su parni i neparni zadaci međusobno slični, što ne mora biti točno, pa tako ne možemo biti dovoljno sigurni u dobiveni koeficijent pouzdanosti. Ako brzina rješavanja nema utjecaja na rezultate, tj. ako je vrijeme gotovo neograničeno i nema utjecaja na klasifikaciju rezultata, onda se može primjenjivati ovaj pristup, a u protivnom su potrebne dodatne analize. Kao najniža vrijednost koeficijenta korelacije Mužić navodi 0.80, a ponekad se traži i 0.85 ili čak 0.90, npr. ako se radi o matematičkim zadacima s računanjem gdje se očekuje veća pouzdanost.

Korelacija je termin koji se odnosi na povezanost među varijablama koju je moguće kvantitativno izraziti tj. računskim postupkom se dobije visina stupnja povezanosti ili koeficijent korelacije. Npr. ako za varijable odaberemo kvalitetu nastave i ukupan uspjeh učenika, onda porast u kvaliteti nastave prati i porast u uspjehu učenika, te se dobije pozitivni koeficijent korelacije. Koeficijent korelacije može imati vrijednosti između  $-1$  i  $1$ , gdje  $-1$  označava potpunu negativnu korelaciju,  $0$  označava da korelacija ne postoji, a  $1$  označava potpunu pozitivnu korelaciju. Ostale vrijednosti iz intervala se nazivaju djelomičnim pozitivnim ili negativnim korelacijama (ovisno o predznaku).

Za pouzdanost je važna jednoznačnost pitanja, jer učenik koji zna odgovor može odgovoriti netočno ako je pogrešno razumio pitanje. Ako postoje zadaci s dvočlanim izborom (npr. točno/netočno ili istina/laž) onda postoji i veća mogućnost da učenik slučajno pogodi odgovor što nema veze s njegovim znanjem, a smanjuje pouzdanost testa.

Faktor koji se povezuje s pouzdanosti je duljina testa tj. broj zadataka ili pitanja koje test sadrži. Povećanjem broja zadataka povećava se i pouzdanost testa, pod uvjetom da su novi zadaci iste kvalitete kao postojeći.

Odlučili smo provjeriti pouzdanost naših testova metodom izračunavanja koeficijenta homogenosti. Za svakog studenta posebno smo zbrojili bodove ostvarene na parnim zadacima (prva varijabla), te bodove ostvarene na neparnim zadacima (druga varijabla). Podatke smo upisali u tablicu (eng. spreadsheet) statističkog programa STATISTICA 6. Na izborniku je potrebno odabrati: *Statistics* → *Basic Statistics and Tables* → *Correlation matrices*. Zatim se odabere *Two lists*. Za prvi popis (eng. First variable list) se odabere uspjeh na parnim zadacima, a za drugi popis (eng. Second variable list), uspjeh na neparnim zadacima, te na kraju *Summary*. Dobiveni koeficijent za pre-test je 0.747, a na isti način je dobiven koeficijent za post test koji iznosi 0.753. Prema [Siegle, xxxx] pošto se radi o pola testa, potrebno je primijeniti Spearman-Brown-ovu formulu:

$$r_{cijeli} = \frac{2 \cdot r_{pola}}{1 + r_{pola}} \quad (3-5)$$

$r_{cijeli}$  = koeficijent pouzdanosti cijelog testa

$r_{pola}$  = koeficijent pouzdanosti pola testa

Dakle, iz formule (3-5) slijedi da je pouzdanost za pre-test i post-test približno jednaka i iznosi 0.85. To je dovoljno visok koeficijent, ali kao što je već navedeno, ne možemo biti sigurni da oba dijela testa sadrže slične zadatke.

Formula (3-5) je specijalni oblik Spearman-Brown-ove formule (3-6) koja se još naziva i *formula predviđanja* (eng. prophecy formula) zbog toga što se koristi za predviđanje pouzdanosti testa prilikom promjene duljine testa tj. promjene broja zadataka.

$$r_n = \frac{n \cdot r_t}{1 + (n-1) \cdot r_t} \quad (3-6)$$

$$n = \frac{\text{novi\_broj\_zadataka}}{\text{trenutni\_broj\_zadataka}} \quad (3-7)$$

$r_n$  = novi koeficijent pouzdanosti

$r_t$  = trenutni koeficijent pouzdanosti

Ako se želi odrediti koliko zadataka je potrebno dodati da bi se dobila željena pouzdanost, onda se koristi formula (3-8):

$$n' = n \cdot \frac{r' \cdot (1-r)}{r \cdot (1-r')} \quad (3-8)$$

$n'$  = broj zadataka koji je potreban za željenu pouzdanost

$n$  = trenutni broj zadataka

$r'$  = pouzdanost koja se želi postići

$r$  = trenutna pouzdanost

Drugi način određivanja pouzdanosti je pomoću *Cronbach-ovog  $\alpha$  koeficijenta* (eng. Cronbach's Coefficient Alpha) [Hempel, 2003]. Ovaj način se često koristi za procjenu pouzdanosti jer se s njim može ocjenjivati pouzdanost testova čiji se zadaci ocjenjuju samo s 0 ili 1 bod (odnosno točno/netočno), te testovi čiji se zadaci ocjenjuju s više od jednog boda [Siegle, xxxx]. Koeficijent  $\alpha$  daje najnižu procjenu pouzdanosti koja se može očekivati. Ako je koeficijent visok, znači da je pouzdanost visoka. Međutim, ako je koeficijent nizak, ne može se ništa zaključiti o pouzdanosti, te je treba ispitati na neki drugi način. Za razliku od prethodnog pristupa gdje se test dijeli na parne i neparne zadatke, Cronbach-ov koeficijent se računa kao aritmetička sredina svih mogućih podjela testa. Ovisi o broju zadataka, a moguće ga je izračunati pomoću statističkih računalnih programa.

Kod izračunavanja Cronbach-ovog  $\alpha$  koeficijenta za naše testove koristili smo program STATISTICA 6. Potrebno je u tablicu (eng. spreadsheet) za svakog studenta upisati ostvarene bodove na pojedinim zadacima (to su varijable). Na izborniku se odabere *Statistics* → *Multivariate Exploratory Techniques* → *Reliability/Item Analysis*. Zatim se odaberu varijable tj. svi zadaci. U dijalogu *Review Descriptive Statistics* se odabere OK, a zatim se u dijalogu *Reliability Results* odabere *Summary*. Prema [Hempel, 2003] smatra se da je test pouzdan ako je  $\alpha > 0.70$ , ali preveliko približavanje 1 može značiti redundanciju tj. nepotrebno ponavljanje istih stvari kroz različite zadatke.

Kao što se vidi iz Tablice 3-4, koeficijent  $\alpha$  za pre-test je 0.881647, ali dovoljne su nam dvije decimale tj.  $\alpha=0.88$ . Pri analizi rezultata obično se detaljnije promatra stupac tablice koji sadrži  $\alpha$  koeficijente. To su koeficijenti koje bi test mogao imati ako se određeni zadatak izbriše. Npr. ako bi se izbrisao zadatak 10. onda bi se  $\alpha$  povećao tj. bio bi jednak 0.882940. To je neznatno povećanje pouzdanosti pa nema smisla brisati taj zadatak (i dalje je to približno jednako 0.88), a smanjivanjem broja zadataka ugrožava se osjetljivost testa. Isti komentar se odnosi na zadatke 6. i 15. Ako bi se npr. izbrisao zadatak 20., onda bi se pouzdanost testa smanjila.

Tablica 3-4: Cronbach-ov koeficijent za pre-test

Ako se zadatak izbriše:	Aritmetička sredina	Standardna devijacija	$\alpha$
Zadatak 1	21,29	9,06	0,881447
Zadatak 2	21,35	9,01	0,880211
...	...	...	...
Zadatak 6	21,23	9,10	<b>0,882534</b>
...	...	...	...
Zadatak 10	21,29	9,10	<b>0,882940</b>
...	...	...	...
Zadatak 15	20,90	8,48	<b>0,882306</b>
...	...	...	...
Zadatak 20	21,87	8,79	0,875062

Cronbach:	0,881647
-----------	----------

U Tablici 3-5 je dio rezultata za post-test, a koeficijent  $\alpha$  je 0.80. Ovdje također ne možemo izbaciti zadatke 2. i 4. na osnovi neznatnog povećanja pouzdanosti koje bi dobili. S druge strane, izbacivanjem zadatka 16. pouzdanost se smanjuje na 0.78. To znači da je taj zadatak važan za pouzdanost testa.

Tablica 3-5: Cronbach-ov koeficijent za post-test

Ako se zadatak izbriše:	Aritmetička sredina	Standardna devijacija	Alpha
Zadatak 1	15,58	7,95	0,801576
Zadatak 2	15,13	8,02	<b>0,804960</b>
Zadatak 3	15,35	7,90	0,798413
Zadatak 4	14,97	7,97	<b>0,804431</b>
...	...	...	...
Zadatak 16	12,90	6,33	0,784187

Cronbach:	0,803077
-----------	----------

Dakle, primjenom dva načina ispitivanja pouzdanosti dobili smo pozitivne rezultate (tj. koeficijent homogenosti  $>0.80$  i  $\alpha >0.70$ ), ali je uzorak od 31 studenta prilično mali pa bi te rezultate trebalo oprezno interpretirati, tj. ne možemo biti dovoljno sigurni u njihovu općenitu pouzdanost da bi ih bez ikakve sumnje ili dodatnog ispitivanja mogli primijeniti na nekom

drugom uzorku studenata. Ipak, u slučaju negativnih rezultata tj. kad bi se pokazalo da testovi na ovom uzorku nisu pouzdani, onda se ne bi uopće mogli osloniti na rezultate testova.

- Objektivnost

Ovdje se radi o razini objektivnosti prilikom ocjenjivanja odgovora ispitanika. Ako zadatak ima samo jedan mogući točan odgovor (npr. zadaci s više ponuđenih odgovora od kojih se bira jedan) onda ocjenjivač ne može biti u nedoumici: odgovor je ili točan ili netočan. Problemi nastaju kod pitanja na koje nije moguće odgovoriti na jedinstven način, ispitanici odgovaraju "svojim riječima", a ocjenjivač mora subjektivno ocijeniti je li odgovor točan. Događa se i da je pitanje postavljeno na nespretan način ili je previše složeno da bi se na njega mogao dati jednostavan odgovor.

Za određivanje objektivnosti je potrebno usporediti ocjene raznih ocjenjivača. Pri tome se očekuje stupanj korelacije veći od 0.98, osim ako su pitanja prilično složena pa se tada dozvoljavaju i niže vrijednosti, a potrebno je dati i preciznije upute za ocjenjivanje.

Kod našeg istraživanja nije ispitana objektivnost testova iz praktičnih razloga. Naime, ovi su se testovi trebali upotrijebiti samo jedan put te se ne bi isplatilo okupiti veći broj ocjenjivača koji bi morali poznavati specifično područno znanje i biti upoznati s istraživanjem. Ova karakteristika je primjerenija ako se testove planira koristiti i za neku buduću namjenu. Dakle, ne može se zaključivati o tome je li i kojoj mjeri objektivnost utjecala na rezultate naših testova jer nije ispitana, a postoje pitanja u testovima na koja nije moguće dati jednostavan odgovor pa bi objektivnost možda mogla biti upitna.

- Osjetljivost

Test je osjetljiv ako se s njim mogu razlikovati i male veličine onoga što se mjeri. Ako su svi zadaci laki, onda se ne mogu međusobno razlikovati oni najbolji, a ako su svi teški onda postoji mogućnost da slabiji učenici ne riješe gotovo ništa, što naravno ne znači da ništa ne znaju. Pri tome ono što istraživaču izgleda lako (ili teško) ne mora biti lako (ili teško) za učenika. Dakle, subjektivna procjena istraživača nije dovoljna. Test se smatra osjetljivijim što je duži [Mužić, 1977]. Test koji se sastoji od samo deset zadataka ne može razlikovati više od 11 kategorija ispitanika (jer se rezultati kreću od 0-10), bez obzira koliko ima ispitanika (može ih biti i nekoliko tisuća). Naravno, u ovom slučaju se zadaci označavaju s točan/netočan. U slučaju kada se zadaci ocjenjuju s više bodova, onda se mjere i manje razlike. Međutim, prema [Mužić, 1977] osjetljivost nema smisla ako su pouzdanost i objektivnost testa niske. Konačan kriterij osjetljivosti je distribucija rezultata. Ako je mnogo istih rezultata, onda test ne može biti dovoljno osjetljiv.

Rezultati studenata kod našeg istraživanja su u Prilogu E. U pre-testu je bilo 20 zadataka, a u post-testu 16. Moglo se očekivati da će post-test biti manje osjetljiv. To se pokazalo i na rezultatima jer kod pre-testa postoje 24 različita rezultata, a kod post-testa 19 različitih rezultata. Međutim, ne možemo donositi općenite zaključke o osjetljivosti jer se radi o samo jednoj primjeni testova na malom uzorku, a već je ranije spomenuto da nije ispitana objektivnost naših testova.

- Diskriminativna vrijednost zadataka

Pojedini zadaci u testu su više valjani što se više slaže uspjeh učenika na tom zadatku s uspjehom na čitavom testu. Prema tome, takvi zadaci (ako su valjani) omogućuju razlikovanje uspješnih i neuspješnih učenika. Ako uspješni učenici dobro rješavaju neki zadatak, a neuspješni slabo, onda će stupanj slaganja između uspjeha u tom zadatku i ukupnog uspjeha u

testu biti visok, a diskriminativna vrijednost zadatka visoka i pozitivna. Može se dogoditi i da neki zadatak učenici sa slabijim uspjehom rješavaju, a uspješniji učenici ne. U tom slučaju je diskriminativna vrijednost zadatka negativna, a takav zadatak ne bi smio biti dio testa.

Diskriminativna vrijednost zadataka se određuje računanjem "point-biserijalnog" koeficijenta korelacije koji se obično označava s  $r_{pb}$  [Petz, 2002]. Obično se provodi sljedeći postupak. Najprije se "numerira" dihotomna varijabla tj. jedna karakteristika se označi s jednim brojem, a druga s drugim, npr. 0 i 1. Tako se dobije varijabla koja sadrži samo dvije različite vrijednosti tj. dihotomna varijabla. Npr. s 1 možemo označiti uspjeh studenta koji je "prošao" tj. ostvario uspjeh veći ili jednak 50 bodova, a s 0 uspjeh studenta koji je "pao" tj. ostvario manje od 50 bodova.

Sljedeći korak je računanje Pearson-ovog koeficijenta korelacije, a najlakše ga je izračunati uz pomoć računala. Ako to nije moguće, može se koristiti formula (3-9).

$$r_{pb} = \frac{n \sum Y_1 - n_1 \sum Y}{\sqrt{n_1 n_0 [n \sum Y^2 - (\sum Y)^2]}} \quad (3-9)$$

X = dihotomizirana varijabla (označena s 0 i 1)

Y = kontinuirana varijabla (npr. bodovi)

$\sum Y_1$  = suma vrijednosti Y koje su vezane uz vrijednost varijable X označene s 1

$\sum Y$  = suma svih Y vrijednosti

$\sum Y^2$  = suma svih kvadriranih Y vrijednosti

$N_1$  = broj X varijabli za koje vrijedi X=1

$N_0$  = broj X varijabli za koje vrijedi X=0

N = ukupan broj slučajeva tj.  $N = N_0 + N_1$

Prema [Mužić, 1977] potrebno je rangirati ispitanike prema uspjehu ostvarenom na testu, te se izdvoji 27% najboljih studenata i 27% najslabijih, a srednji ne ulaze u razmatranje. Za naše istraživanje bi dakle od 31 studenta trebalo odvojiti 9 najboljih i 9 najslabijih, ali pošto Mužić [Mužić, 1977] naglašava da ukupna veličina uzorka ne bi trebala biti manja od 370, kako bi u obje grupe bilo bar po 100 ispitanika, onda ovu podjelu nismo radili jer je ionako uzorak malen. Dakle, koeficijent korelacije je izračunat na cijelom uzorku.

Dio dobivenih rezultata je u Tablici 3-6. Zadaci su poredani prema veličini koeficijenta  $r_{pb}$ . Statistički značajni koeficijenti su označeni podebljano. Ostali rezultati pokazuju određenu korelaciju, ali ona nije statistički značajna tj. mogla je nastati slučajno.

Tablica 3-6: Diskriminativna vrijednost zadataka za testove

Pre-test	$r_{pb}$	Post-test	$r_{pb}$
Zadatak 10	0,01	Zadatak 2	0,12
Zadatak 3	0,19	Zadatak 3	0,25
Zadatak 2	0,20	Zadatak 1	0,28
Zadatak 18	0,27	Zadatak 4	0,30
Zadatak 4	0,29	Zadatak 5	0,35
Zadatak 20	0,31	Zadatak 15	<b>0,41</b>
Zadatak 9	0,32	Zadatak 14	<b>0,43</b>
Zadatak 5	<b>0,36</b>	Zadatak 11	<b>0,51</b>
Zadatak 1	<b>0,37</b>	Zadatak 6	<b>0,57</b>
...	...	...	...



Dakle, prema rezultatima u Tablici 3-6, zadaci 3. i 10. u pre-testu, te zadatak 2. u post-testu imaju niski stupanj korelacije pa smo pokušali analizirati rezultate bez tih zadataka. Mora se opet naglasiti da je uzorak premalen da bi mogli biti dovoljno sigurni i izbaciti te zadatke. Ipak, kao komentar, dobivena veličina učinka bez tih zadataka je 0.18 (za cijeli test je bio 0.12). Možda bi se izbacivanjem još kojeg zadatka dobili bolji rezultati, ali se to ne smije napraviti jer iako nisu u statistički značajnoj korelaciji, neka korelacija ipak postoji, a također smanjivanjem broja zadataka smanjuje se pouzdanost i osjetljivost testa.

U Tablici 3-7 je dan pregled ispitanih metričkih karakteristika naših testova.

Tablica 3-7: Metričke karakteristike naših testova

Karakteristika	Opis
Valjanost	Sadržaji testova i predavanja se podudaraju, pa zaključujemo da su u ovom slučaju testovi valjani.
Pouzdanost	Prema dobivenim koeficijentima pouzdanosti, testovi su u ovom slučaju pouzdani, ali je uzorak ispitanika malen.
Objektivnost	Nije ispitana (potrebno je više ocjenjivača testova).
Osjetljivost	Testovi razlikuju veći broj kategorija, ali pošto je uzorak ispitanika premali i testovi su primijenjeni samo jedan put, ne možemo donositi općenite zaključke. Pouzdanost i objektivnost utječu na ovu karakteristiku, a objektivnost nije ispitana.
Diskriminativna vrijednost zadataka	Možda je moguće izbaciti neke zadatke, ali uzorak je opet premali da bi bili potpuno sigurni u ispravnost takvog postupka, a također pouzdanost i osjetljivost postaju upitne.

Što možemo zaključiti o testovima, jesu li u redu? Općenito, ne možemo biti sigurni jer je problem prilikom ispitivanja karakteristika testova bio premali uzorak ispitanika, a također i činjenica da su testovi primijenjeni samo jedan put. Za ovu primjenu, ako zanemarimo te probleme i ograničimo se samo na dobivene rezultate, onda možemo reći da su testovi u redu (podrazumijevajući i da je ocjenjivanje testova bilo objektivno).

#### 3.2.4. Komentar rezultata

Istraživanje pokazuje pozitivnu veličinu učinka od 0.12 standardnih devijacija. To je manje od očekivanog rezultata, a t-test je pokazao da nema statistički značajnih razlika između kontrolne i eksperimentalne grupe nakon učenja na sustavu DTEEx-Sys. Prema tome, iako je veličina učinka pozitivna i u korist eksperimentalne grupe, t-test pokazuje da je ta razlika mogla nastati i slučajno. Veličina učinka od 0.12 standardnih devijacija označava poboljšanje od 50-tog do 54-og centila koje pokazuju studenti eksperimentalne grupe u odnosu na studente kontrolne grupe. Npr. ako se rezultat studenta nalazi između 54-og i 55-og centila, onda možemo reći da je 54% studenata ostvarilo lošiji rezultat od njega, a 45% bolji rezultat.

Prema [Mužić, 1977] postoji više tipova pogrešaka koje mogu utjecati na rezultate eksperimenta. Jedan tip pogreške leži u subjektima (u našem slučaju studentima) na kojima se vrši eksperiment. Podjelom na grupe može doći do toga da se razlikuju po predznanju, općoj mentalnoj sposobnosti, marljivosti, itd. Od navedenog je samo predznanje moglo biti

provjereno pre-testom. Tijekom provjere metričkih karakteristika testova, često se javljao problem malog uzorka studenata na kojem se provodilo istraživanje. Wisher i Olson [Wisher & Olson, 2003] navode da veličina uzorka utječe na statističku moć testa. Oni su analizirali istraživanja koja su provedena u svrhu vrednovanja Web-orijentiranog poučavanja. Veličina učinka je bila veća kod većih uzoraka, ali nije bilo dovoljno dostupnih podataka da bi tu tvrdnju mogli generalizirati.

Mogući razlog slabijih rezultata je mogla biti i nezainteresiranost studenata da uče za eksperiment i da aktivno sudjeluju u eksperimentu. Neka od pitanja na koje su studenti slabije odgovorili u pre-testu su preformulirana i ponovljena u post-testu kako bi se vidjelo jesu li učili na pogreškama. Pretpostavljalo se da će studenti na ta "ponovljena" pitanja odgovoriti puno bolje. Dakle, za 5. pitanje iz pre-testa i 10. pitanje iz post-testa se tražio odgovor da se postojanost ostvaruje objektivno-orijentiranom bazom podataka, a za 20. i 4. da je objektivno-orijentirano prirodan izbor za oblikovanje distribuiranih sustava zato što i objekti i distribuirani sustavi komuniciraju slanjem poruka. Kao što se vidi iz Tablice 3-8. studenti ne pokazuju veliki napredak u uspjehu kod ta dva pitanja što možda pokazuje da svi studenti nisu dovoljno učili.

Tablica 3-8: Uspjeh na sličnim pitanjima

<b>Kontrolna grupa</b>			
Pre-test	Post-test	Pre-test	Post-test
5.	10.	18.18 %	54.55 %
20.	4.	0 %	59.09 %
<b>Eksperimentalna grupa</b>			
Pre-test	Post-test	Pre-test	Post-test
5.	10.	35.00 %	30.00 %
20.	4.	20.00 %	42.50 %

Iz tih činjenica i promatranja ponašanja studenata tijekom trajanja cijelog procesa istraživanja može se pretpostaviti da dio studenata nije bio dovoljno motiviran, a pošto se radi o malom uzorku, to može imati veći utjecaj na rezultate. Jedan od problema prilikom istraživanja učenja na daljinu koje navode Phipps i Merisotis [Phipps & Merisotis, 1999] (opisan u Poglavlju 2.4.1.) je u tome što se dovoljno ne ispituju razlozi zbog kojih studenti odustaju od takvog načina učenja. Problem odustajanja se ovdje spominje jer smatramo da je vezan uz nedovoljnu motiviranost studenata, ali pošto ovo istraživanje nije trajalo tijekom cijelog kolegija, nije ga bilo moguće istražiti.

#### 4. ZAKLJUČAK I PRAVCI DALJNJEG ISTRAŽIVANJA

Inteligentni tutorski sustavi nude mnoge prednosti kod učenja i poučavanja. Jedna od najčešće isticanih prednosti je poučavanje tipa "jedan učitelj – jedan učenik". ITS-ovi se mogu prilagođavati posebnim potrebama svakog učenika. Za razliku od ljudskih tutora, ITS-ovi su stalno motivirani za rad, ne osjećaju umor, ne zaboravljaju, ne ljute se, jednako se odnose prema svim učenicima i sl. ITS-ovi također smanjuju troškove poučavanja.

Učenici i učitelji postaju sve više fizički udaljeni. Uz Web-orijentirane inteligentne tutorske sustave učenici se više ne moraju okupljati u isto vrijeme na određenom mjestu radi učenja i poučavanja. Ono postaje dostupno u njihovom domu. Razvoj raznih područja i tehnologija zahtijeva od ljudi obrazovanje koje traje cijeli život. Web-orijentirani ITS-ovi obećavaju dostupnost kvalitetnog obrazovanja svim osobama kojima je potrebno, pod uvjetom da imaju odgovarajuću tehnologiju za pristup. Ta tehnologija postaje sve bolja i jeftinija, a na taj način i lakše dostupna.

Iako se čini da Web-orijentirani ITS-ovi nude mnoge prednosti, potrebno je provjeriti je li to zaista tako, odnosno provesti istraživanje kojim će se ispitati učinkovitost konkretnog Web-orijentiranog ITS-a. Svako provedeno istraživanje pridonosi općenitoj ocjeni učinkovitosti Web-orijentiranih ITS-ova.

Istraživanje je potrebno planirati. Kada se odredi što će se konkretno istraživati, potrebno je odabrati uzorak ispitanika. To se najčešće radi slučajnim odabirom, a ako to nije moguće onda se posebno ispituju karakteristike svakog ispitanika. Važan korak je priprema testova i provjera njihovih metričkih karakteristika, zato što o njima ovisi ispravnost rezultata istraživanja. Dobiveni rezultati istraživanja se analiziraju odabranim statističkim metodama.

Prilikom provedbe našeg istraživanja sustava DTEEx-Sys nismo dobili visoku veličinu učinka, ali smo stekli dragocjeno iskustvo koje nam može pomoći pri budućim istraživanjima. Smatramo da je nemotiviranost studenata koji su sudjelovali u istraživanju mogla imati negativnog utjecaja na provedbu i rezultate istraživanja. Samo istraživanje je trajalo relativno kratko na manjem broju studenata, a vjerojatno bi se kroz dulje vrijeme studenti bolje priviknuli na rad sa sustavom DTEEx-Sys. Potrebno je prilikom analize istraživanja uzeti u obzir i koliko vremena su studenti učili on-line na sustavu DTEEx-Sys, te ispitati imaju li svi odgovarajuću tehnologiju za pristup (pristup Internetu, računalo i odgovarajuću programsku podršku).

---

**Literatura**

[Albacete & VanLehn, 2000]

Albacete, P.L. & VanLehn, K.A. "Evaluating the Effectiveness of a Cognitive Tutor for Fundamental Physics Concepts" in *Proceedings of the 22nd Annual Meeting of the Cognitive Science Society*, 2000.

[Amižić *et al.*, 2002]

Amižić, A. Stankov, S., Rosić, M. "Model traganja – dijagnostička tehnika inteligentnih tutorskih sustava", (URL [http://www.pmfst.hr/~ani/znanstveni\\_radovi/Mipro\\_2002.pdf](http://www.pmfst.hr/~ani/znanstveni_radovi/Mipro_2002.pdf)), Fakultet prirodoslovno-matematičkih znanosti i odgojnih područja, Split, 2002.

[Beck *et al.*, 1996]

Beck, J., Stern, M. & Haugsjaa, E. "Applications of AI in Education", (URL <http://www.acm.org/crossroads/xrds3-1/aeid.html>), 1996.

[Becker, 1999]

Becker, L. A. "Testing for Differences Between Two Groups: t test", (URL <http://web.uccs.edu/lbecker/SPSS/ttest.htm>), 1999.

[Cotton, 1989]

Cotton, K. "Educational Time Factors", 1989, Northwest Regional Educational Laboratory: (URL <http://www.nwrel.org/scpd/sirs/4/cu8.html>)

[Cotton, 1991]

Cotton, K. "Computer-Assisted Instruction", Northwest Regional Educational Laboratory, (URL <http://www.nwrel.org/scpd/sirs/5/cu10.html>), 1991.

[Debak, 2002]

Debak, D. "Inteligentni tutorski sustavi: Prošlost, sadašnjost, budućnost", diplomski rad, Fakultet prirodoslovno-matematičkih znanosti i odgojnih područja, Split, 2002.

[Fletcher, 2003]

Fletcher, J. D. "Evidence for Learning From Technology-Assisted Instruction", in H. F. O'Neal, R. S. Perez, *Technology applications in education: a learning view*, Mahwah, NJ: Lawrence Erlbaum Associates, 2003, str. 79-99.

[Goodkovsky, 1997]

Goodkovsky, V.A. "Intelligent Tutor: Top-down Approach to Intelligent Tutoring System Design", Learning Technology Standards Committee (P1484) --Developing Technical Standards for Learning Technology, (URL <http://www.goodkovsky.org/publication.html>), 1997.

[Graham, 1994]

Graham, I. *Object-oriented methods*, Addison-Wesley Publishing Company, 1994.

[Hempel, 2003]

Hempel, S. "Reliability", University of Derby, (URL <http://ibs.derby.ac.uk/~susanne/PTT/lectures/PTTrelability2003.pdf>), 2003.

[Iqbal *et al.*, 1999]

Iqbal, A., Oppermann, R., Patel, A. and Kinshuk "A Classification of Evaluation Methods for Intelligent Tutoring Systems", *Software Ergonomie '99 – Design von Informationswelten* (Eds. U. Arend, E. Eberleh & K.Pitschke), B. G. Teubner Stuttgart, Leipzig, 1999, str.169-181.

[Lewis *et al.*, 1999]

Lewis, L., Snow, K., Farris, E. & Levin, D. "Distance Education at Postsecondary Education Institutions: 1997-98", U.S. Department of Education, National Center for Education Statistics (URL <http://nces.ed.gov>), Washington, DC: 1999.

[Mark & Greer, 1993]

Mark, M. A. & Greer, J. "Evaluation Methodologies for Intelligent Tutoring Systems", *Journal of Artificial Intelligence and education*, 4 (2/3), 1993, str. 129-153.

[Mužić, 1977]

Mužić, V. *Metodologija pedagoškog istraživanja*, Svjetlost, Sarajevo, 1977.

[Park & Lee, 2003]

Park, O. & Lee, J. "Adaptive Instructional Systems", *Encyclopedia of Educational Technology* (URL <http://coe.sdsu.edu/eet/articles/cmi/>), 2003.

[Petz, 2002]

Petz, B. *Osnovne statističke metode za nematematičare*, Naklada Slap, Jastrebarsko, 2002.

[Phipps & Merisotis,1999]

Phipps, R. & Merisotis, J. "What's the Difference? - A Review of Contemporary Research on the Effectiveness of Distance Learning in Higher Education", The Institute for Higher Education Policy, Washington, (URL <http://www.ihep.com/Pubs/PDF/Difference.pdf>), 1999.

[Phipps *et al.*, 2000]

Phipps, R., Merisotis, J., Harvey, M. "Quality on the line: Benchmarks for success in internet-based distance education", The Institute for Higher Education Policy, (URL <http://www.ihep.com/Pubs/PDF/Quality.pdf>), Washington, 2000.

[Regian, 1997]

Regian, J. W. "Functional area analysis of intelligent computer-assisted instruction", (Report of TAPSTEM ICAI-FAA Committee), Brooks AFB, TX, 1997.

[Rosić, 2000]

Rosić, M. "Zasnivanje sustava obrazovanja na daljinu unutar informacijske infrastrukture", magistarski rad, Fakultet elektrotehnike i računarstva, Sveučilište u Zagrebu, Zagreb, 2000.

[Rosić *et al.*, 2001]

Rosić, M., Stankov, S., Glavinić, V. "DTEEx-Sys – A Web Oriented Intelligent Tutoring System", in *Proceedings of Intelligent Conference On Trends in Communication – EUROCON 2001*, Vol 2/2, 2001, str. 255-258.

[Sadr, 1998]

Sadr, B. *Unified Objects: Object-Oriented Programming Using C++*, IEEE Computer Society, Los Alamitos, California, 1998, str. 1-56.

[Shareef & Kinshuk, 2003]

Shareef, A. F. & Kinshuk "Student Model for Distance Education System in Maldives", in A. Rossett (Ed.) *Proceedings of E-Learn 2003*, Phoenix, (Arizona, USA), 2003, (URL <http://infosys.massey.ac.nz/~kinshuk/papers/elearn2003.pdf>).

[Siegle, xxxx]

Siegle, D. "Reliability", Neag School of Education - University of Connecticut, (URL [http://www.gifted.uconn.edu/siegle/research/Instrument Reliability and Validity/Reliability.htm](http://www.gifted.uconn.edu/siegle/research/Instrument%20Reliability%20and%20Validity/Reliability.htm))

[Smith, 1997]

Smith, S. P. "Developing an authoring environment for procedural task tutoring systems", PhdThesis, Massey University, Palmerston North, New Zealand, 1997 (URL <http://www-users.cs.york.ac.uk/~shamus/thesis/glossary.pdf>).

[Stankov, 1996]

Stankov, S. "Student Model Developing for Intelligent Tutoring Systems" (URL [http://www.pmfst.hr/~stankov/zn\\_radovi/1996/coins/COINS\\_rad.pdf](http://www.pmfst.hr/~stankov/zn_radovi/1996/coins/COINS_rad.pdf)), Fakultet prirodoslovno-matematičkih znanosti i odgojnih područja, Sveučilište u Splitu, 1996.

[Stankov, 1997]

Stankov, S. "Izomorfni model sustava kao osnova računalom poduprtog poučavanja načela vođenja", doktorat, Fakultet elektrotehnike, strojarstva i brodogradnje, Sveučilište u Splitu, Split, 1997.

[Stankov *et al.*, 2001]

Stankov, S., Rosić, M. & Glavinić V. "New Generation of Intelligent Tutoring Shell Designed through Unified Modeling Language", (URL [http://www.pmfst.hr/~stankov/zn\\_radovi/2001/INES\\_2001.pdf](http://www.pmfst.hr/~stankov/zn_radovi/2001/INES_2001.pdf)), Fakultet prirodoslovno-matematičkih znanosti i odgojnih područja, Sveučilište u Splitu, Split, 2001.

[Stankov, 2003]

Stankov, S. "Struktura sustava i prikaz znanja u sustavu TEx-Sys", (URL [http://www.pmfst.hr/~stankov/texsys/texsys\\_razvoj/dio\\_1234.pdf](http://www.pmfst.hr/~stankov/texsys/texsys_razvoj/dio_1234.pdf)), Fakultet prirodoslovno-matematičkih znanosti i odgojnih područja, Sveučilište u Splitu, Split, 2003.

[Stankov *et al.*, 2004]

Stankov, S., Rosić, M., Granić, A., Maleš, L., Grubišić, A. & Žitko, B. "Paradigma e-učenja & Inteligentni tutorski sustavi", Fakultet prirodoslovno-matematičkih znanosti i odgojnih područja, Sveučilište u Splitu, MIPRO 2004, Opatija.

[StatSoft, 2004]

StatSoft, Inc. "Electronic Statistics Textbook", (URL <http://www.statsoft.com/textbook/stathome.html>), 2004.

[Wisher & Olson, 2003]

Wisher, R.A. & Olson, T. M. "The Effectiveness of Web-Based Instruction", U.S. Army Research Institute for the Behavioral and Social Sciences, Virginia, Research Report 1802, 2003.

[Woolf *et al.*, 2003]

Woolf, B. P., Beck, J., Eliot, C. & Stern, M. "Growth and Maturity of Intelligent Tutoring Systems: A Status Report", Department of Computer Science, University of Massachusetts, Amherst, MA.

[Yaakub, 1998]

Yaakub, M. N. "Meta-Analysis of the Effectiveness of Computer-Assisted Instruction in Technical Education and Training", doctoral dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 1998.

[Yang *et al.*, 2002]

Yang, A., Kinshuk & Patel, A. "A Plug-able Web-based Intelligent Tutoring System", in S. Wrycza (Ed.) *Proceedings of the Xth European Conference on Information Systems*, Gdansk, Poland: Wydawnictwo Uniwersytetu Gdanskiego, (URL [http://fims-www.massey.ac.nz/~kinshuk/papers/ecis2002\\_ang\\_yang.pdf](http://fims-www.massey.ac.nz/~kinshuk/papers/ecis2002_ang_yang.pdf)), 2002.

[Zhou & Evens, 1999]

Zhou, Y. & Evens, M. W. "A Practical Student Model in an Intelligent Tutoring System", iz *Proceedings of the 11th IEEE International Conference on Tools with Artificial Intelligence*, Chicago, IL, str. 13-18, (URL <http://www.cs.iit.edu/~cirsim/documents/yzitai99.pdf>), 1999.

**PRILOG**

- Prilog A – Sadržaj predavanja
- Prilog B – Pre-test
- Prilog C – Post-test
- Prilog D – Izgled baze znanja
- Prilog E – Tablica uspjeha studenata



***Prilog A - Sadržaj predavanja***

***Prilog B - Pre-test***

***Prilog C - Post-test***

***Prilog D - Izgled baze znanja***

Pri izradi baze znanja korištena je slijedeća literatura: [Sadr, 1998] i [Graham, 1994].

*Prilog E - Tablica uspjeha studenata*

	<b>Pre-test bodovi</b>	<b>Post-test bodovi</b>	<b>DTEEx-Sys kategorije</b>	<b>DTEEx-Sys bodovi</b>	<b>DTEEx-Sys ocjena</b>
<b>Student 1.</b>	45	28	0-5-5	25 / 50	2
<b>Student 2.</b>	63	64	-	-	-
<b>Student 3.</b>	97	44	0-1-9	46 / 58	4
<b>Student 4.</b>	24	39	1-5-4	27 / 46	3
<b>Student 5.</b>	34	22	2-3-1	7 / 22	1
<b>Student 6.</b>	74	61	0-3-7	50 / 54	5
<b>Student 7.</b>	50	25	-	-	-
<b>Student 8.</b>	55	17	2-1-0	0 / 8	1
<b>Student 9.</b>	37	25	-	-	-
<b>Student 10.</b>	42	36	6-4-0	13 / 28	2
<b>Student 11.</b>	95	61	0-1-9	46 / 58	4
<b>Student 12.</b>	71	61	-	-	-
<b>Student 13.</b>	42	53	-	-	-
<b>Student 14.</b>	61	19	-	-	-
<b>Student 15.</b>	87	81	0-1-9	52 / 58	4
<b>Student 16.</b>	29	28	-	-	-
<b>Student 17.</b>	50	69	0-1-9	46 / 58	4
<b>Student 18.</b>	92	64	-	-	-
<b>Student 19.</b>	100	92	0-3-7	42.5 / 54	4
<b>Student 20.</b>	63	50	0-2-8	42 / 56	4
<b>Student 21.</b>	42	19	0-1-9	49 / 58	4
<b>Student 22.</b>	13	19	1-8-1	21 / 40	3
<b>Student 23.</b>	92	67	-	-	-
<b>Student 24.</b>	68	64	0-3-7	31 / 54	3
<b>Student 25.</b>	42	22	0-4-6	25 / 52	2
<b>Student 26.</b>	53	33	1-2-7	33 / 52	3
<b>Student 27.</b>	26	6	-	-	-
<b>Student 28.</b>	55	39	-	-	-
<b>Student 29.</b>	82	86	0-1-9	43 / 58	4
<b>Student 30.</b>	32	25	2-5-3	21.5 / 42	3
<b>Student 31.</b>	89	53	0-1-9	56.5 / 58	5