

Oponentský posudek disertační práce

název disertační práce: Rozvoj databázového myšlení v kontextu školního vzdělávání

autor disertační práce: Ing. Mikoláš Panský

školitelka: doc. PaedDr. Radka Wildová, CSc.

studijní obor: pedagogika

zaměření: didaktika technické a informační výchovy

tento oponentský posudek vypracovala: PhDr. Mgr. Hana Voňková, Ph.D. et Ph.D.

Předložená disertační práce je příspěvkem k didaktice informačních a komunikačních technologií. Zaměřuje se na rozvoj databázového myšlení žáků, což je podstatné, avšak na našich školách opomíjené téma. Tématicky je tedy tato práce velmi zajímavá. Její cíle jsou velmi přehledně vymezeny na straně 9: stanovit cíl rozvoje databázového myšlení, strukturovat obsah rozvoje databázového myšlení, uvést metody výuky databázového myšlení a analyzovat charakteristiky rozvoje databázového myšlení studentů. Pro naplnění cílů své práce vychází autor z literatury především o informačních a komunikačních technologiích, didaktice informační výchovy a obecné didaktice. Jako hlavní cíl rozvoje databázového myšlení uvádí rozvoj kognitivní struktury znalostí databázového jazyka. Autor stanovuje šest základních úrovní rozvoje databázového myšlení: propedeutická, nácviková, expanzivní, stabilizační, integrační a aplikační úroveň. Mimo tohoto vertikálního členění udává i další členění: deklarativní, procedurální a strukturální oblast databázového myšlení. V praktické části se autor pokouší zavést kvantitativní charakteristiky rozvoje databázového myšlení. Tato kvantifikace je v daném kontextu inovativní. Využívá k tomu data sebraná v rámci kurzu databázových systémů, který je určen studentům učitelství ICT na Pedagogické fakultě Univerzity Karlovy.

Práce má sice jasně vymezenou strukturu cílů (viz zmíněná str. 9), kterých se snaží autor v jednotlivých kapitolách dosáhnout, avšak mnohdy není zřejmé, k čemu přesně autor směřuje, jak spolu souvisí jednotlivé kapitoly, odstavce a věty, jak spolu souvisí odkazy na literaturu, proč jsou použity určité statistické výpočty atd. V takto obsáhlé práci se obtížně vyhýbá překlepům. V této práci je jich však mnoho. Níže jsou konkrétně uvedeny příklady problematických částí práce. V souvislosti s tím je vyřčena celá řada dotazů, která by měla být během obhajoby zodpovězena.

Definice pojmů

Abstrakt a str. 7: databázové myšlení je v abstraktu definováno jako “soubor kompetencí a způsobu myšlení vedoucího k osvojení práce s databázovým systémem” a rozvoj databázového myšlení jako “pedagogické působení na žáka za účelem formování odpovídající kognitivní struktury a nácviku praktických dovedností”. Propojení těchto dvou definic není zřejmé. Pokud by platila definice databázového myšlení, pak by rozvoj databázového myšlení měl být rozvoj “souboru kompetencí a způsobu myšlení vedoucího k osvojení práce s databázovým systémem”, což nemusí odpovídat tomu, že se jedná o “pedagogické působení” - interakci mezi žákem a učitelem. To již spíše popisuje proces, při němž lze databázové myšlení rozvíjet.

Definice rozvoje databázového myšlení je uvedena i na str. 7: “Pod rozvojem databázového myšlení v disertační práci souhrnně označuji soubor dovedností a znalostí vedoucích ke zvládnutí komunikace s datovou bází, schopnost navrhovat a implementovat vnitřní uspořádání datové báze a dovednost prezentovat údaje získané z datové báze.” Tato definice významově odpovídá definici databázového myšlení uvedené v abstraktu. Neodpovídá definici rozvoje databázového myšlení uvedené v abstraktu (tam je uvedeno jako pedagogické působení, viz výše). Obecně je vhodné přistoupit k situaci takto: To, co chceme rozvíjet, nejprve definovat a následně uvést, jak tuto zdefinovanou věc rozvíjet.

Sporná tvrzení

str. 50: “Dále specifikovat strukturu jednotlivých polí tabulky ...”

Co je pole tabulky (u tabulek hovoříme o sloupcích)? Co je pak specifikace struktury pole (Struktura tabulky jako celku? Nebo snad struktura jednotlivých sloupců? Ty však žádnou vnitřní strukturu mít nemají.)

“... a schopnost definovat datový typ sloupců tabulky”

Jak se to liší od významu první části věty? Stačí napsat “nedefinovat tabulku, sloupce a datový typ sloupce”.

Jako další sporné tvrzení lze např. uvést: “Mimo to lze specifikovat výchozí hodnotu datového pole. Tím je zajištěna hodnota, která bude zobrazena v poli pro případ, že by tabulka obsahovala prázdnou hodnotu.” Specifikace výchozí hodnoty přece nahradí nezadanou hodnotu (hodnotu NULL), nejen změni způsob její prezentace, takže se informace o nezadané hodnotě ztratí.

“Aproximativní vyjadřování” se ve vědeckém textu nesmí objevovat (s výjimkou analýzy nepřesného vyjadřování respondentů apod.). Při definování pojmů používaných v pedagogice, databázovém jazyce a statistice je nutné být přesný.

Logická struktura textu, návaznost jednotlivých částí, odstavců a vět

Práci je náročné číst především díky nelogické struktuře samotného textu. Zde uvádím několik příkladů. V samotném textu se jich však vyskytuje mnoho.

a) první odstavec na str. 5, není zřejmý význam “z moderního hlediska” ve větě začínající “Z moderního hlediska je proto ...”. O pravdivosti tvrzení této věty by šlo uvažovat, ale nejasným způsobem navazuje na předchozí věty a navíc z předchozích vět nevyplývá, proč by to mělo být právě z moderního hlediska.

b) poslední odstavec na str. 5 zakončený na str. 6 (“Etablující se didaktika informační výchovy ...”) - nejasné propojení cílů informační výchovy a cílů didaktiky informační výchovy. Cíle informační výchovy jsou v textu vztaženy k ICT. Odstavec dále pokračuje definicí cílů didaktiky informační výchovy. Vzhledem k soudržnosti odstavce lze předpokládat, že se vyjde z cílů informační výchovy. Oproti tomu však autor pojednává o didaktickém myšlení obecně a cituje z knihy *Obecná didaktika* od prof. Skalkové (Skalková, 1999) bez rozumné návaznosti na předchozí tvrzení o cílech informační výchovy. Nelze pochybovat o pravdivosti ani jedné části odstavce, nicméně o jejich vzájemném propojení. Jedná se spíše o “slepení” citací z různých zdrojů bez hlubšího promyšlení.

c) str. 115 a 116 “Experiment by bylo vhodné realizovat kontinuálně a ve větším měřítku. S ohledem na velikost souboru dat nelze všechny výsledky zobecňovat. Mohou však poskytnout výstup pro hlavní

cíl práce a tím je stanovení možných charakteristik učení dotazovacího jazyka. Aby bylo možné z výsledků vyvodit konkrétní vzory chování studenta při učení dotazovacího jazyka, bylo by nutné provést analýzu ve větším měřítku. ... ”

První věta v odstavci je nejasná, ve čtvrté větě je část této věty zopakováno.

Jazyková vyjádření

a) str. 8: “Pouze 0,55 % z celkového počtu respondentů” - buď stručně uvést vyjádření relativní četnosti v procentech: Pouze 0,55 % respondentů, nebo absolutní četnosti: Pouze x z celkového počtu y respondentů, kde x a y jsou přirozená čísla.

b) str. 50 “Při návrhu tabulky se jedná konkrétně o schopnost tabulku navrhnout a pojmenovat.” Jedná se o takřka tautologickou větu.

c) Často se v práci vyskytuje spojení “domnívám se, že”. Je ve většině případů neopodstatněné, není jasný důvod, proč se autor danou věc “domnívá”.

Statistická analýza

V praktické části se autor pokouší zavést kvantitativní charakteristiky rozvoje databázového myšlení. Využívá k tomu data sebraná v rámci kurzu databázových systémů, který je určen studentům učitelství ICT na Pedagogické fakultě Univerzity Karlovy. Během výuky se zaznamenávaly SQL dotazy včetně informace o tom, kdo a kdy ji zadal, typ příkazu a hlášení serveru o výsledku dotazu. Hlavní charakteristikou, která byla analyzována, byl “počet dotazů vykonaných databází s ohledem na jejich správnost”. Struktura jednotlivých lekcí byla následující: výklad tématu, procvičování zadáváním dotazů do databázového systému – v první fázi s využitím textového rozhraní a v druhé fázi s využitím grafického rozhraní.

Jak autor uvádí, “v kontextu cíle praktické části bylo nezbytné analyzovat nejenom kvantitu a správnost činností studenta ve výuce, ale i určit statistické metody pro jejich zpracování” (str. 114). Na straně 118 autor uvádí výčet statistických metod, které hodlá využít k analýze dat. Není však zřejmé, proč volí právě tyto metody. Na straně 119 pokračuje: “Pro komparaci obou metod zadávání SQL dotazů byly použity testy normality statistického souboru s cílem ověřit. ” Co hodlá autor ověřit? Je možné, že hodlá testovat, zda-li je rozdělení nějaké statistické veličiny normální. V tabulkách 9 a 10 jsou pak prezentovány testy normality pro učení v textovém rozhraní a grafickém rozhraní. Jaká data do testu přesně vstupují? Za jaké lekce? Co vyjadřuje “p” v posledním sloupci těchto tabulek? Z jakého důvodu jsou provedeny K-S test i Lilliefors test?

Na str. 122 je uvedeno: “Testy variability průměrů ukázali statisticky významný rozdíl mezi průměrným počtem všech, správných i nesprávných dotazů zadaných v grafickém uživatelském rozhraní v porovnání s textovým uživatelským rozhraním. Výsledek lze interpretovat tak, že studenti se při používání textového rozhraní zadávají menší počet dotazů. ”

Zde je několik nepřesností. Testy variability průměrů pravděpodobně autor myslí testy rozdílu průměrů. Vyjde-li z toho, že autor testuje rozdíl mezi průměry a nikoli variabilitu a že dále v textu konstatuje, že shledává statisticky významný rozdíl mezi průměry, pak z toho nevyplývá jím uvedená interpretace, že studenti při používání textového rozhraní zadávají menší počet dotazů. Jsou-li průměry dvou veličin X a Y statisticky odlišné, pak z toho nevyplývá, že průměr veličiny X je menší než průměr veličiny Y. Může to být i naopak.

Na straně 123 následuje kapitola o Kolmogorově-Smirnově testu. Není zřejmé, proč autor opisuje učebnici statistiky, jedná se o standardní test. Proč je v textu uvedeno odvození tohoto testu? (Jakým způsobem dosazuje autor do jednotlivých vzorců svá data? Nebo hodlá nějak konkrétně využít tohoto odvození při výuce tohoto testu? Z předchozích stran by spíše vyplývalo, že autor použil při analýze dat nějaký standardní statistický software, kde byl již test k dispozici. K čemu tedy slouží to odvození? A proč právě na tomto místě? A co v odvození znamená věta: “Necht’ je $S_n(x)$ sledované kumulativní rozdělení náhodné proměnné, x_1, x_2, \dots, x_n také známé jako empirické rozdělení.”?)

Dále jsou uvedeny obrázky 16 a 17 o trendech chybovosti studentů. Osy grafů nejsou označeny, takže pro čtenáře není zřejmé, co je v grafech přesně uvedeno.

Na straně 128 jsou v tabulce 13 uvedeny popisné statistiky souboru. Popisné statistiky se většinou uvádějí na začátek statistické analýzy před prováděním testů inferenční statistiky. Proč jsou popisné statistiky uvedeny až zde? Tuto tabulku by však bylo vhodné dát do přílohy a stručně např. uvést, že směrodatná odchylka je u většiny studentů v rozmezí 0,01 a 0,02. Dále co znamená číslo 0,00 v řádku Celkem a sloupci Směrodatná chyba?

Např. na obrázcích 21 (str. 129), 23 (str. 133), 24 (str. 133), tabulce 26 (str. 137), obrázku 25 (str. 138) je zakreslena regresní přímk. Pokud jsou data rozložena do kruhu, je využití lineární regrese zcela nevhodné a nic nevypovídající. Proč je tedy využita tato metoda pro analýzu dat? Pokud to má být pouze příklad toho, jak lze analyzovat chybovost v kontextu práce s databází, pak je minimálně v tomto případě nevhodné tuto metodu využít. Jaká jiná metoda, vhodná pro daná data by šla využít?

Na straně 134 je uvedena definice korelace. Opět se naskytá podobný dotaz jako ke stránce 123, kde je definován K-S test. Proč je zde uváděna jeho definice? Navíc se opět nejedná o přesné vymezení. Např. ve větě: “Pro výpočet korelačního koeficientu ρ_{XY} byl použit následující vzorec: ...”, přičemž ve vzorci není uvedeno žádné ρ_{XY} . Navíc není zřejmé, jak autor značí náhodnou veličinu a její realizaci (co je označeno velkým a co malým písmenem). Označení by mělo být konzistentní v celém textu, což neodpovídá (jiné značení u K-S testu a jiné značení u korelace).

Anglické výrazy

a) Překlad abstraktu: “Thesis is focused on problematic of purposes ...” - problematic je přídavné jméno problematický, není to problematika, jedna z možností správného vyjádření: “Thesis is focused on purposes ...”

b) Překlad abstraktu: Překlad abstraktu nemusí být doslovný, ale neměl by vyjadřovat něco jiného. V českém abstraktu je: “Praktická část je orientována na metodiky a evaluační kritéria.” V anglickém abstraktu tomu má odpovídat věta: “Experimental part proposes quantitative metrics for database thinking development.”

c) další anglické výrazy: str. 1 (obsah práce) a 56 – “Standard Qualification and Curriculum Authority”. Je zde několik překlepů. Navíc není zřejmé, co tím autor myslí. Qualifications and Curriculum Authority, která stanovuje standardy (standards)?

Překlepy

V takto obsáhlé práci se velmi obtížně vyhýbá překlepům. V práci jich však je nepříjemně mnoho.

Uvádím příklady z úvodní části práci pro ilustraci jejich hustoty.

str. 3 “ ... s narůstající důležitosti informací ... ”

str. 3 “ dostává kromě samotné učení nových poznatků”,

str. 4 Nerozumím návaznosti první a třetí věty v souvětí: “Občan informační společnosti si často neuvědomuje, jaký prostředek ovládá, a databáze se staly transparentní součástí informační struktury.”

str. 6 “autoritativního klimatu 3školy”

str. 7 “Souhrnně pro databázové znalostí a dovedností”

str. 7 “V něm se věnuji otázkou vymezení kompetencí ... ”

str. 8 “... tezi integraci ... ”

Závěr: I přes uvedené připomínky doporučuji disertační práci k obhajobě. Během ní je však nutné, aby doktorand reagoval na uvedené dotazy a nejasnosti.

27. 7. 2011 v Praze

Hana Voňková, oponentka disertační práce