

ՊԱՇՏՈՆԱԿԱՆ ԸՆԴԴԻՄԱԽՈՍԻ ԿԱՐԾԻՔԸ

Գևորգ Արմենի Պետրոսյանի «Տեքստի ինքնատիպության աստիճանի որոշման բազմալեզու համակարգի մշակումը» թեմայով Ե.13.02 - «Ավտոմատացման համակարգեր» մասնագիտությամբ տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի հայցման ատենախոսության վերաբերյալ

Թեմայի արդիականությունը

Ամբենայական թարգմանության գործիքների արագ զարգացումը և բազմալեզու էլեկտրոնային ռեսուրսներին հասանելիության ընդլայնումը զգալիորեն պարզեցրել են թարգմանված գրագողության պրակտիկան: Նման դեպքերը դժվար է հայտնաբերել գրագողության հայտնաբերման ավանդական համակարգերի միջոցով, քանի որ դրանք հիմնականում կենտրոնացած են միալեզու տեքստերի համեմատության վրա:

Այս պայմաններում տեքստի ինքնատիպության աստիճանի որոշումը դարձել է արդիական խնդիր, որը պահանջում է հաշվի առնել հնարավոր բազմալեզու աղբյուրները: Հայերենի համար ևս այս խնդիրը արդիական է՝ հաշվի առնելով մասնագիտացված գործիքների սակավությունը կամ լայն գիտական շրջանակներին դրանց հասանելիության բացակայությունը: Հետևաբար, բազմալեզու միջավայրում հայերեն տեքստի ինքնատիպության աստիճանի ավտոմատ որոշումը կարևոր գործնական խնդիր է:

Աշխատանքի համառոտ բովանդակությունը

Ներածությունում հիմնավորված է թեմայի արդիականությունը, սահմանված են հետազոտության նպատակը, խնդիրները, օբյեկտը և առարկան, ներկայացված են հետազոտության մեթոդները, գիտական նորույթը, գործնական արժեքը և պաշտպանության հիմնական դրույթները:

Առաջին գլխում կատարվել է տեքստային փոխառությունների հայտնաբերման ոլորտի վերլուծություն: Սահմանվել են հիմնական հասկացությունները, կատարվել է գրագողության տարբեր տեսակների և գրագողության հայտնաբերման գոյություն ունեցող առևտրային համակարգերի վերլուծություն: Ներկայացվել է տեքստային փոխառությունների հայտնաբերման մեթոդների համակարգված վերլուծություն:

Երկրորդ գլխում ներկայացվել են առաջարկվող ավտոմատացված համակարգի հայեցակարգային և կառուցվածքային լուծումները: Մասնավորապես, ներկայացվել են համակարգի ճարտարապետությունը և ֆունկցիոնալ մոդելը՝ կիրառելով ֆունկցիոնալ մոդելավորման IDEFO ստանդարտը:

Երրորդ գլխում ներկայացված են միալեզու և միջլեզվական փոխառությունների հայտնաբերման մշակված մեթոդները: Մանրամասն նկարագրված են տեքստային փոխառությունների սկզբնաղբյուրների որոնման, մանրամասն վերլուծության, ինչպես նաև միալեզու փոխառությունների որոնման կառուցվածքային մեթոդի մաթեմատիկական

հիմնավորումները: Ներկայացված են նաև խոր ուսուցման կիրառված մոդելի վերապատրաստման համար զուգահեռ և վերաձևակերպված նախադասությունների զույգերի հավաքածուների ստեղծման և մոդելի վերապատրաստման քայլերի հաջորդականությունը, ինչպես նաև վերապատրաստման արդյունքները հայերեն-անգլերեն և հայերեն-ռուսերեն լեզվական զույգերի դեպքում:

Չորրորդ գլխում ներկայացված են միջլեզվական փոխառությունների որոնման մշակված մեթոդների փորձարարական հետազոտությունների արդյունքները: Կատարվել է նաև համեմատական վերլուծություն այլ աշխատանքներում ներկայացված արդյունքների հետ, ինչը հաստատում է մշակված մեթոդների արդյունավետությունը:

Հինգերորդ գլխում ներկայացված է նախագծված ավտոմատացված համակարգի ծրագրային իրականացումը: Մասնավորապես, նկարագրված են հիմնական ծրագրավորման լեզվի, վեբ-շրջանակի, հիմնական բաղադրիչների իրականացման մասնագիտացված գործիքների ընտրությունը, ինչպես նաև տվյալների բազայի սխեմայի մանրամասն նկարագրությունը:

Հավելվածում ներկայացված է ատենախոսության հիմնական դրույթների՝ Վանաձորի Հ. Թումանյանի անվան պետական համալսարանի «Մաթեմատիկայի և ինֆորմատիկայի» ամբիոնի «Արհեստական բանականության հիմունքներ» դասընթացի տեսական ու գործնական պարապմունքներում ներդրման ակտը:

Եզրահանգման մեջ ընդհանրացված են ստացված արդյունքները, ձևակերպված են գիտական եզրակացությունները, բերված են առաջարկություններ:

Ատենախոսության գիտական նորույթը

- Մշակվել է հայերեն տեքստում պարունակվող միջլեզվական փոխառությունների սկզբնաղբյուրների որոնման մեթոդ՝ հիմնված խոսքիմասային պիտակավորման վրա:
- Մշակվել է հայերեն միալեզու փոխառությունների որոնման մեթոդ՝ հիմնված մարկովյան շղթաների վրա:
- Մշակվել է հայերեն տեքստում պարունակվող միջլեզվական փոխառությունների մանրամասն վերլուծության մեթոդ՝ հիմնված բազմալեզու տրանսֆորմերային մոդելի վրա:
- Մշակվել է հայերեն նախադասությունների վերաձևակերպման մոտեցում՝ օգտագործելով բազմամակարդակ ֆիլտրման համակարգ, որն ապահովում է զույգերի իմաստային նմանությունը և բառապաշարային բազմազանությունը:
- Ստեղծվել են հայերեն-անգլերեն և հայերեն-ռուսերեն լեզվական զույգերի համար նախատեսված զուգահեռ և վերաձևակերպված նախադասությունների զույգերի հավաքածուներ, որոնք ներառում են նաև «դժվար բացասական» (hard negative) օրինակներ:
- Կատարվել է բազմալեզու տրանսֆորմերային LaBSE մոդելի վերապատրաստում՝ հայերեն-անգլերեն և հայերեն-ռուսերեն լեզվական զույգերի համար մշակված

զուգահեռ և վերաձևակերպված նախադասությունների զույգերի հավաքածուների հիման վրա:

- Հայերեն-անգլերեն և հայերեն-ռուսերեն լեզվական զույգերի համար հարմարեցվել է թեստային տվյալների հավաքածու՝ գոյություն ունեցող տվյալների հավաքածուի թարգմանության միջոցով:

Գիտական դրույթների, եզրակացությունների հիմնավորման և հավաստիության աստիճանը հաստատված են մշակված մեթոդների մաթեմատիկական հիմնավորումներով, փորձնական հետազոտությունների և համեմատական վերլուծությունների արդյունքներով:

Աշխատանքի գործնական նշանակությունը

Մշակված համակարգը երկմակարդակ ճարտարապետության, տվյալների կենտրոնացված պահպանման և մեծ ռեսուրսներ պահանջող առաջադրանքների ասինխրոն մշակման շնորհիվ ապահովում է արագագործություն և բարձր ճշգրտություն բառացի փոխառությունների, վերաձևակերպումների և թարգմանությունների հայտնաբերման համար: Այն կարող է օգտագործվել Հայաստանի Հանրապետության գիտահետազոտական և կրթական կազմակերպություններում՝ գիտական աշխատանքների ինքնատիպության աստիճանի որոշումն ավտոմատացնելու համար:

Ընդհանուր գնահատականներ

Ատենախոսության հիմնական արդյունքները հրատարակված են 9 գիտական աշխատանքներում, որոնցից երկուսը՝ առանց համահեղինակների, իսկ երկուսը՝ Scopus գիտատեղեկատվական շտեմարանում:

Դիտողություններ և առաջարկություններ

- Աշխատանքում նշված չէ համակարգի հաշվողական ծախսերի ժամանակային գնահատականը, մասնավորապես՝ փաստաթղթերի տվյալների բազայի ծավալը մեծացնելիս, ինչը կարևոր է անընդհատ աճող տվյալների բազայի պայմաններում մշակված համակարգի մասշտաբայնությունը գնահատելու համար: Ցանկալի կլիներ նաև գնահատել մշակված մեթոդների աշխատանքի ժամանակը և կատարել համեմատական վերլուծություն այլ աշխատանքների հետ՝ ըստ ժամանակային գնահատման:
- Վերաձևակերպված նախադասությունների տվյալների բազմություն ստեղծելիս հայերեն վերաձևակերպված նախադասությունների որակը գնահատվել է միայն իմաստային նմանության և բառապաշարային բազմազանության ավտոմատ չափանիշներով: Ցանկալի կլիներ ներգրավել որակի փորձագիտական գնահատում, ինչը կբարձրացներ ստեղծված տվյալների բազմության որակի հուսալիությունը:
- Հստակ ներկայացված չեն համակարգում առաջին և երկրորդ մակարդակներում որոնումների պարամետրերը, մասնավորապես՝ առաջին մակարդակում բառերի հատվող հաջորդականությունների երկարությունը, երկրորդ մակարդակում իմաստային նմանության շեմը և այլն:

- Ցանկալի է ընդլայնել մշակված համակարգի տվյալների բազան, որպեսզի այն ընդգրկի համացանցի հայալեզու տիրույթում հրապարակված գիտական աշխատանքների մեծ մասը:

Եզրակացություն

Չնայած նշված դիտողություններին, Գևորգ Արմենի Պետրոսյանի «Տեքստի ինքնատիպության աստիճանի որոշման բազմալեզու համակարգի մշակումը» թեմայով ատենախոսությունը կատարված է բարձր գիտական ու գործնական մակարդակով: Ատենախոսության շրջանակներում ստացված արդյունքները հավաստի են, փորձարարական ճանապարհով հաստատվել են: Ատենախոսության սեղմագիրը լիովին արտացոլում է ատենախոսության հիմնական դրույթները: Կարծում եմ, որ ատենախոսությունը բավարարում է ՀՀ-ում գիտական աստիճանների շնորհման կարգի թեկնածուական ատենախոսություններին ներկայացվող բոլոր պահանջներին, իսկ հեղինակը՝ Գևորգ Արմենի Պետրոսյանը, արժանի է տեխնիկական գիտությունների թեկնածուի գիտական աստիճանի շնորհմանը Ե.13.02 - «Ավտոմատացման համակարգեր» մասնագիտությամբ:

Պաշտոնական ընդդիմախոս՝

Ֆ.-մ.գ.թ., դոցենտ,

L. S. L.

Հ.Յ. Հակոբյան

Հ.Յ. Հակոբյանի ստորագրության իսկությունը հաստատում եմ՝

ԵՊՀ գիտական քարտուղար՝

Մ. Վ. Հովհաննիսյան
2

Մ.Վ. Հովհաննիսյան



“ 19 ” 06 2026թ.