



KOMPYUTER LINGVISTIKASIDA ALGORITMLIK NAZARIYASI VA ULARNING OMMAVIYLIK VA NATIJAVIYLIK XUSUSIYATLARI

Sodiqova Baxtigul Ibodullayevna

Xorijiy til va adabiyoti kafedrası o'qituvchisi

Denov tadbirkorlik va pedagogika instituti, Uzbekiston

sodiqova.baxtigul@mail.ru

(90)981-89-96

Annotatsiya:

Ushbu maqolada kompyuter lingvistikasiga kirib kelgan algoritmlik nazariyasi, uning diskretlik xususiyati, qoidalarining elementarlik xususiyati, algoritmning ommaviylik va natijaviylik xususiyatlari, sonli algoritmlar, mantiqiy algoritmlar haqida fikr yuritiladi hamda mavzuning murakkablik va dolzarblik muammolari ko'rib chiqiladi.

Kalit so'zlar :

Algoritm, algoritmlar nazariyasi, algoritmlarning ommaviylik va natijaviylik xususiyatlari, til birliklarini modellashtirish tamoyillari, lingvistik model, lisoniy model, sonli algoritmlar, mantiqiy algoritmlar.

Kirish. Tilshunoslik fani XIX asrda mustaqil fan sifatida shakllandi. Shundan boshlab u turli aspektlarda, yo'nalishlarda rivojlanib kelmoqda. Barcha fanlardagi kabi tilshunoslikda ham ikki fanning kesishuvida yuzaga kelgan fanlar jadal rivojlanmoqda, jumladan, ana shunday fanlar sirasiga sotsiolingvistika (sotsiologiya va tilshunoslik), psixolingvistika (psixologiya va tilshunoslik), etnolingvistika (etnografiya va lingvistika), neyrolingvistika (nevrologiya va tilshunoslik), matematik lingvistika va kompyuter lingvistikasi fanlarini kiritish mumkin. Bunday holatni boshqa fanlar doirasida ham kuzatish mumkin: bioximiyaga, astrofizika, matematik fizika, matematik logika kabilari. Buni fanlar tizimida bir necha fanlarning o'zaro integratsiyasi deb baholash mumkin. Mashina tarjimasining asoschilari kibernetika va matematika sohasi vakillari bo'lib, keyinchalik bu ishda tilshunoslar ham faol qatnasha boshlagan. Shu tariqa mashina tarjimasiga g'oyalari butun dunyoda nazariy va amaliy tilshunoslikning rivojlanishida katta ahamiyat kasb eta boshladi. Bu yo'nalish bilan parallel ravishda matematik lingvistika fanida ishlab chiqilib, bu o'z navbatida, kompyuter lingvistikasi fanining yuzaga kelishi uchun poydevor bo'la oldi. Shu asosda tilshunoslikning yangi yo'nalishi bo'lgan kompyuter lingvistikasi va tilshunoslikning bir qator nazariy va amaliy yo'nalishlari vujudga keldi. Matematik lingvistika fani XX asrning 50-yillarida tilshunoslikning alohida yo'nalishi sifatida yuzaga keldi. Bu fanning shakllanishida Kopengagen struktural tilshunoslik maktabi (glossematika)ning asoschisi Lui Yelmslevning g'oyalari o'ziga xos „turtki“ vazifasini o'tagan. U hatto til hosidalarini matematik bayonda tushuntiradigan fanning nomini ham taklif etgan. Olimning fikricha, bu fan „Til algebrasi“ („Lingvistik algebra“) deb atalishi lozim edi. Amerikalik tilshunos Noam Chomskiyning formal grammatika, transformatsion grammatika haqidagi qarashlari bevosita matematik lingvistikaning alohida yo'nalish sifatida yuzaga kelishiga sabab bo'lgan. Mana shunday qarashlar ta'sirida matematik lingvistika fani shakllandi. Matematik lingvistika bu – tabiiy tillarning matematik modellarini ishlab chiqish, xususan, sun'iy tillarni yaratish algoritmini tuzish bilan shug'ullanuvchi fandır. Matematik lingvistika oldida turuvchi eng muhim masalalar quyidagilar deb belgilash mumkin:

- tilning aksiomatik nazariyasini ishlab chiqish;
- formal grammatika yaratish;

- tillarning matematik modellarini ishlab chiqish.

- til hodisalarini matematik metodlar yordamida tahlil qilish (ehtimollar nazariyasi, statistika va kvantitativ metodlarni tatbiq etish). Kompyuter lingvistikasi matematik lingvistikaning mantiqiy davomi bo'lib, u amaliy tilshunoslikning eng muhim qismini tashkil etadi.

Asosiy qism. Kompyuter lingvistikasi 1954 yil AQSHda Jorjtaun universitetida mashina tarjimasiga bo'yicha dunyoda o'tkazilgan birinchi tajriba asnosida yo'nalish sifatida shakllana boshladi, 1960 yilga kelib mustaqil fan sifatida shakllandi. Kompyuter lingvistikasi inglizcha "computational linguistics" so'zidan olingan. XX asrning 80-yillariga qadar bu fan turlicha nomlar bilan atalgan: hisoblash lingvistikasi, matematik lingvistika, kvantitativ lingvistika, injener lingvistikasi kabi. Bu fanning asosiy maqsadi lingvistik masalalarni yechishning kompyuter dasturlarini ishlab chiqish, inson va mashina (kompyuter) muloqotini optimallashtirish, tabiiy tilni qayta ishlashdir. Kompyuter lingvistikasida tabiiy tillarning kompyuter analizi va sintezini o'z ichiga oladi. Bunda analiz tabiiy tilning kompyuterda morfologik, sintaktik va semantik tahlil yordamida tushunilishiga nisbatan ishlatiladi, sintez esa kompyuterda matnning grammatik shakllantirilishi va generatsiyasi (hosil qilinishi) demakdir. NLP bo'yicha yaratilgan dasturiy ta'minotlar quyidagilar: AlchemyAPI, Expert System S.p.A., General Architecture for Text Engineering (GATE), Modular Audio Recognition Framework, MontyLingua, Natural Language Toolkit (NLTK).

Kompyuter lingvistikasida qo'llaniladigan va yaratiladigan lingvistik vositalarni shartli ravishda ikkiga qismga bo'lish mumkin: deklarativ hamda protsedura qismlari. Deklarativ qismga til va nutq birliklari lug'ati, grammatik ma'lumotnomalar, matnlar korpusi kabilarni kiritish mumkin. Protседura qismi esa yuqoridagi lingvistik ta'minot bazasini boshqarish vositalarini (algoritmlar tuzish, dasturlar yaratish, kompyuter analizi va sintezi kabilari) o'z ichiga oladi. O'zbek tili materiallari bo'yicha kompyuter lingvistikasiga oid tadqiqot olib borgan olimlar sifatida H.Arziqulov, S.Rizayev, S.Muhamedov, A.Po'latov, S.Muhamedova, N.Jo'rayevalarni ko'rsatish mumkin. Olimlar asosan statistik tahlil, algoritmlash, o'zbek tilining aksiomatik nazariyasi, fe'llarning kompyuter analizi va sintezi kabi yo'nalishlar bo'yicha izlanishlarni amalga oshirganlar. Kompyuter lingvistikasi va nazariy tilshunoslik bir-birini to'ldiradi. Kompyuter lingvistikasida til insondan tashqarida mavhum sistema sifatida, L.Yelmslev ta'rifi bilan aytganda "sof munosabatlar tizimi" sifatida tavsiflanadi. Kompyuter lingvistikasi va klassik (mumtoz) tilshunoslik orasidagi farq quyidagi nuqtalarda ko'rinadi: - Mumtoz tilshunoslikda til inson bilan mutanosiblikda, juftlikda ko'rib chiqiladi. Ya'ni mumtoz tilshunoslik insonga yo'naltirilgan bo'ladi va uning faol ishtirokida qabul qilinadi. Kompyuter lingvistikasi esa tavsiflash jarayonida insonni istisno qiladi va u ko'proq kompyuterga moslashtiriladi. - Mumtoz tilshunoslik ko'proq tavsifiy (deskriptiv) xarakterga ega hisoblanadi. Kompyuter lingvistikasi esa masalani miqdoriy xarakteristikalar va aniq parametrlar asosida hal qiladi. Demak, mumtoz tilshunoslik ko'proq tavsifiy bayonga asoslansa, kompyuter lingvistikasi miqdoriy (kvantitativ) tavsifga, algoritmlash, modellashtirish, statistik tahlilga asoslanadi. - Mumtoz tilshunoslik ko'proq nazariy xarakterga ega bo'lib, tilshunoslikning nazariy masalalari bilan shug'ullanadi. Mumtoz tilshunoslikning tahlil obyekti tabiiy til hisoblanadi va tahlilda uning mavjud barcha imkoniyatlari e'tiborga olinadi. Kompyuter lingvistikasida esa ko'proq sun'iy tillarga (programmallashtirish tillari, algoritmik tillar)ga tayaniladi, tabiiy tillarning mavjud imkoniyatlari cheklanadi, bunda tabiiy tilga ishlov berilib (NLP), kompyuterga moslashtiriladi. Kompyuter lingvistikasida "Algoritm" tushunchasi matematikaning asosiy tushunchalaridan biridir. Bu so'z IX asrda yashagan matematik olim Abu Abdulloh Muhammad al -Xorazmiy nomining lotincha talaffuzidan kelib chiqqan. **Algoritm nazariyasi** deb berilgan ommaviy muammodagi barcha masalalarni umumiy bir shaklda aniq, ma'lum bo'lgan usul bilan echish jarayoniga aytiladi. Demak, quyidagi echuvlarning barchasi algoritmlar qatoriga kiradi:

- a) sonlar ustidagi arifmetik amallarni bajarish qoidalari;
- b) kvadrat ildiz chiqarish qoidasi;
- d) eng katta umumiy bo'luvchini topish qoidasi;
- e) kvadrat tenglamalarning echimini topish qoidasi va boshqalar.

Algoritmning xarakterli xususiyatlari quyidagilar bilan belgilanadi:

a) algoritmlar diskret xarakterga ega, ya'ni boshlang'ich holatdagi miqdorlar keyingi matematik amallar sistemasida hosil qilinadi. Til birliklari ham xuddi shunday diskretlik xususiyari bilan ajralib turadi:

bog'+bon=bog'bon;

bilim+li=bilimli; bosh+la=boshla kabilar (2 bo'linuvchan qismga o'zak va qo'shimcha qismlariga bo'linadi).

b) algoritm qoidalari elementarligi bilan ajralib turadi. Chunki ilgarigi miqdorlar sistemasidan keyingisini hosil qilish qonuni sodda jarayonlar bilan farqlanadi.

d) algoritmlar diterminatsiyalanuvchanligi bilan ajralib turadi, ya'ni boshlang'ich holatda bo'lgan miqdorlar keyingi holatdagi miqdorlar orqali aniqlanadi:

top/il/may/di/gan/lar/dan/siz;

tomosha/bin/lar/imiz/ga kabilar.

e) algoritmlar ommaviyligi va natijaviyligi bilan xarakterlanadi, chunki ularda boshlang'ich miqdorlar sistemasini ayrim potensial cheksiz to'plamlardan aniqlash mumkin.

Til birliklarini qismdan butunga, turdan jinsga, elementdan sistemaga, xususiyan umumiyga yo'nalishlarda tadqiq etish o'z navbatida tilshunoslikka "algoritm", "model" (qolip), "modellashtirish" (qoliplashtirish) tushunchalarining mustahkam o'rin olishiga sabab bo'ldi. Til birliklari, sathlari, sathlararo bog'lanishlarning induksiya va deduksiya asosida izchil tadqiq etilishi narijasida, bir tomondan, lisoniy birliklar, ikkinchi tomondan, shu lisoniy birliklardan tashkil topgan til umumiy mexanizmining model (qolip)lari haqidagi ta'limot vujudga keldi. Til sistemasi va model hodisalarini *element, munosabat, bog'lanish, butunlik* kabi muhim belgilar birlashtirib turadi. O'z ob'ektlarining holatiga ko'ra modellar statik (turg'un) va dinamik (funktional), tavsif maqsadiga ko'ra analitik va sintetik, ifoda belgilariga ko'ra grafik va metatil (mazmuniy, tavsifiy) kabi turlarga bo'lib o'rganish mumkin. Tilshunoslikda modellar nazariy bilish va amaliy faoliyat maqsadlari uchun xizmat qiladi.

Lisoniy model bilan lingvistik model bir xil hodisalar emas. Chunki lisoniy model lingvistik (ilmiy) modelning asli, originali yoxud real ob'ektidir. Masalan, muayyan tildagi fonemalarning modellari turli davrlarda va hozirgi davrda lingvistlar tomonidan har xil belgilanishi mumkin. Bundan ko'rinadiki, real til mutlaq ob'ektiv hodisa va haqiqat bo'lib, modellar uni nisbiy ravishda aks ettiradi. Til mohiyatan lisoniy modellar haqidagi ta'limotdir, deyishimiz mumkin.

Xulosa. Umuman olganda, tilshunoslikda vujudga kelayotgan modellashtirish jarayoni nutq asosi bo'lgan tilning system-strukturasini chuqurroq o'rganish uchun yangi imkoniyatlar yaratib bermoqda. Matematik amallar asosiy rol o'ynaydigan algoritmlarga sonli algoritmlar deyiladi. Bundan tashqari mantiqiy algoritmlar ham bor. Bunday algoritmlar barcha fanlarda va kasblarda, ayniqsa huquqshunoslikda, tibbiyotda, mantiqda, falsafada keng qo'llaniladi. Algoritmning yaratilishi o'z navbatida relekontaktli sxemalar, parallel va ketma-ket sxemalarning ishlab chiqarilishi va kibernetikada qo'llanilishiga olib keladi. O'zbek tilining barcha sath birliklarini modellashtirish bilan bog'liq kompyuter dasturlari ishlab chiqish va uni hayotga jalb etish kompyuter tilshunosligi oldida turgan dolzarb masalalardan biridir.

Adabiyotlar:

1. A. Po'latov, S. Muxamedova "Kompyuter tilshunosligida matnni avtomatik tahrir qilish uchun yaratilgan dasturning ba'zi lisoniy asoslari" // O'zbek tilshunosligi masalari (ilmiy maqolalar to'plami). – Toshkent: TDPU nashri. 2003. – B. 30-38.
2. Sh. To'rayev, Y. Haydarov "O'zbek tili affikslarining ingliz tili affikslari bilan solishtirma elektron ma'lumotlar bazasi // Tilshunoslikning dolzarb masalari (ilmiy maqolalar to'plami)" . II. – Toshkent: Universitet. 2004. – B. 58
3. B. Yo'ldoshev " Matematik va kompyuter lingvistikasi (uslubiy qo'llanma)" . – Samarqand: SamDU nashri. 2007. – B. 31-32.
4. A. Po'latov, S. Muxamedova "Kompyuter lingvistikasi". – Toshkent, 2007. – B. 4-5.

-
5. А. Нурмонов “Структур тилшунослик: илдизлари ва йўналишлари”. – Тошкент, 2008. – Б. 145.
 6. A.Rahimov “Kompyuter lingvistikasi asoslari” O’quv qo’llanma.-Toshkent ,2011.- B. 4-10