

УДК 004.934

**ЖАРНАМА КОМПАНИЯСЫНЫҢ АҚПАРАТТЫҚ ЖҮЙЕСІНДЕ СӨЙЛЕУДІ ТАҢУ
ТЕХНОЛОГИЯСЫН ҚОЛДАНУ**

Махмұт Ернар

yernar1511@gmail.com

Л.Н. Гумилев атындағы Еуразия ұлттық университеті, ақпараттық технологиялар

факультетінің 2-курс магистранты

Ғылыми жетекші - Ә.Омарбекова

Нұр-Сұлтан, Қазақстан

Заманауи веб-жүйелер - күрделі программалық кешендер болып табылады және жасанды интеллект жетістіктері оларды да айналып өтпеуі қисынды. Табиғи тілде сөйлеуді автоматты тану жасанды интеллект дамуының ең өзекті бағыты болып табылады. Бұл бағыттағы нәтижелер адам мен компьютер арасындағы тиімді қарым-қатынас орнату жабдықтарын жасау мәселесін шешуге мүмкіндік береді.

Сөйлеу технологияларының өндірістің кез-келген саласына енгізілуі зерттеу орталықтары мен өндіруші фирмалардың сөйлеуді танумен байланысты мәселелерге қызығушылығын арттырды. Ол дауыспен басқарылатын кез-келген жүйелерде кеңінен қолданыс табуда. Дыбыстық енгізу табиғилық, жеделдік, енгізудің мағыналық дәлдігі, қоланушының қолы мен көзін босату, экстремалды жағдайда басқару мен өңдеу мүмкіндігі тәрізді артықшылықтарға ие. Ал осы сөйлеуді тану технологиясын веб-жүйелерде қолдану -

көру мүмкіндігі шектеулі жандарға септігін тигізеді.

Сөйлеуді тану жүйелерін келесі белгілері бойынша классификациялауға болады:

- дикторға тәуелділігі;
- сөздік көлемі;
- танылатын сөйлеу ағынының сипаты.

Дикторға тәуелді жүйелер жекелеген қолданушыларға жасалады. Бұл жүйелер құрылымы бойынша қарапайымырақ, арзанырақ бірақ дикторлар тобына бейімдеуге немесе белгісіз диктормен жұмыс жасауға жарамсыз.

Дикторға тәуелсіз жүйелер кез-келген диктормен жұмыс істеу үшін жасалады. Бұл жүйелер әлі де қиындықпен дамып келеді, қымбатырақ болып табылады және дикторға тәуелді жүйелермен салыстырғанда тану сенімділігі төменірек болып келеді.

Дикторға бейімделуші жүйелер нақты бір дикторға немес дикторлар тобына баптау жасауға мүмкіндік береді. Мұндай жүйелер танудың жеткілікті тиімділігін қамтамасыз етеді, бірақ оны диктор немесе дикторлар тобына баптау уақыты ұзақ болуы мүмкін.

Сөйлеуді тану жүйесінің көлемі оның күрделілігімен тікелей байланысты және тану дәлдігі сипаттамасына елеулі әсер етеді.

Сөздік көлемі қолданбалы жүйенің нақты талаптарына сәйкес анықталады. Кейбір қолданбалы программалар тек бірнеше сөзді талап етеді (мысалы сандар), ал келесілері өте үлкен сөздіктерді қажет етеді (мысалы, мәтінді автоматты теру жүйелері).

Әдетті сөздік көлемінің келесі градациялары қарастырылады:

- кішкентай сөздік–ондаған сөздер;
- орташа сөздік–жүздеген сөздер;
- үлкен сөздік – мыңдаған сөздер;
- өте үлкен сөздік – он мыңдаған сөздер.

Сөйлеу ағынының сипатына байланысты сөйлеуді тану жүйелері жеке сөздерді тану және үздіксіз сөйлеуді тану жүйелері болып бөлінеді.

Жеке сөздерді тану жүйелері жеке сөздерді бөлек-бөлек кідіріс жасап айту арқылы тануға арналған. Бұл әдіс танудың ең қарапайым түрі болып табылады, себебі мұнда сөздің бастапқы және соңғы нүктелерін табу оңай, және көршілес сөздер бір-біріне әсер етпейді. Бұл жоғары сапалы тануға қол жеткізуді қамтамасыз етеді.

Үздіксіз сөйлеуді тану жүйелері сөздер бірге, кідіріссіз табиғи түрде айтылатын сөйлеуді тануға арналған. Үздіксіз сөйлеуді өңдеу қиынырақ, оған себеп ең алдымен сөздердің бастапқы және соңғы нүктелерін белгілеудің қиындығы. Сонымен қатар «коартикуляция» мәселесі бар. Коартикуляция дегеніміз көршілес дыбыстардың өзара әсер етуі. Сәйкесінше, сөздің басы мен соңындағы дыбыстарға көршілес сөздердің басы мен соңындағы дыбыстар әсер етеді. Үздіксіз сөйлеуді тануға сөйлеу темпі де әсер етеді. Оның үстіне үздіксіз сөйлеуді тану жүйелерінің сөздік көлемі үлкен болуы шарт (он мыңдаған сөздер), себебі олар негізінен мәтінді автоматты теруге арналған. [1]

Сөйлеу технологиясын веб-жүйелерде қолданудың қарапайым жолы - навигацияны дауыспен басқару және ондағы материалды синтездеу арқылы дыбыстау болып табылады. Сөйлеуді тану жүйелерінің жоғарыдағы классификациясын сараптай келе, навигацияны дауыспен басқару үшін кішкентай сөздікті, дикторға тәуелді, және жеке сөздерді танитын жүйе жеткілікті деген шешім шығады. Неге? Себебі біріншіден, веб-жүйе навигациясында сөздер саны аз, мысалы, біз жасап отырған жүйе мәзірінде келесі сөздер бар: "БАСТЫ БЕТ", "БІЗ ТУРАЛЫ", "ҚЫЗМЕТ ТҮРЛЕРІ", "БІЗДІҢ ЖҰМЫСТАР", "ПІКІРЛЕР", "БАЙЛАНЫС", "КЕРІ ҚОҢЫРАУ", "ҚАЗАҚ ТІЛІ", "ОРЫС ТІЛІ", "СЫРТҚЫ ЖАРНАМА", "ИНТЕРЬЕР ЖАРНАМА", "БАННЕР ШЫҒАРУ".

ӨЗІҢ ЖАЙЛЫ ЖАРИЯЛА!

БІЗ ТУРАЛЫ

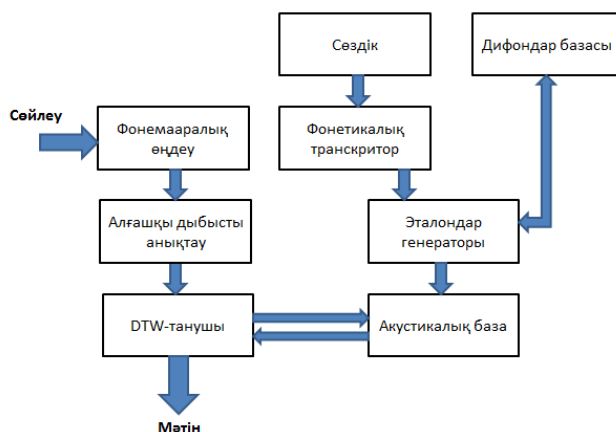
ТАПСЫРЫС БЕРУ

ПОЧЕМУ СТОИТ ВЫБРАТЬ НАС

Сөз тіркестерін бір сөз деп алып қарастырғандықтан, бар-жоғы 12 сөз шығып тұр. Екіншіден, сөздік қоры аз болғандықтан, жүйені әр қолданушының өзіне лайықтап үйретіп алуға мүмкіндігі бар, сондықтан жүйені дикторға тәуелді қылып жасауға болады. Оның үстіне дикторға тәуелсіз сөйлеуді тану жүйесі әлдеқайда қымбатқа түсетінін еске сала кетейін. Үшіншіден, басқару командаларына үздіксіз сөйлеу емес, керісінше, бөліп-бөліп, нақты дыбыстау тән болғандықтан, жеке сөздерді тану алгоритмдерінің бірін жүзеге асыру жеткілікті. Бұл да біздің міндетті жеңілдетеді және арзандата түспек.

Сонымен айтылған талапқа сай келетін дайын сөйлеуді тану алгоритмдерін қарастырайық. Сөйлеуді тану дегенде, нақты қазақ тілі туралы айтып отырғанымыз түсінікті, себебі орыс, ағылшын тілін танитын дайын құралдар жеткілікті.

Қазақ тілінде сөйлеуді автоматты тану бағытындағы зерттеулер жүргізу салыстырмалы түрде кейінірек басталды. Алғашқы жұмыстар небәрі 2006-2007 жылдарда пайда болды, және негізінен жеке сөздерді тануға арналған [2-5]. Қарабалаеваның жұмысында автор жеке сөздерді фонемалық тану алгоритмдерін сипаттаған. Алайда сөйлеуді тану дәлдігі бойынша ешқандай ақпарат келтірмеген. Бекманованың жұмысында жұмыста қазақ сөздерін DTW-алгоритмімен тану алгоритмі сипатталған, дегенмен жұмыста негізінен жазбаша мәтінді морфологиялық талдау мен синтездеуге көбірек назар аударылған. Ал Есенбаевтың жұмыстарында сыртқы шу жағдайында қазақша үздіксіз сөйлеуді тану бойынша үлкен жұмыс жасалған. Одан кейінгі Ә.К. Бөрібаева жұмыстарында нақты осы жеке сөздерді дифондар базасы негізінде танитын дикторға тәуелді жүйе ұсынылған [6]. Әрине, олар үлкен көлемдегі сөздіктер тануға бағдарланған жүйе, дегенмен ол біздің міндетті орындауға кедергі келтірмейді. Сонымен, ол ұсынған жүйенің негізгі құрылымы келесідей:

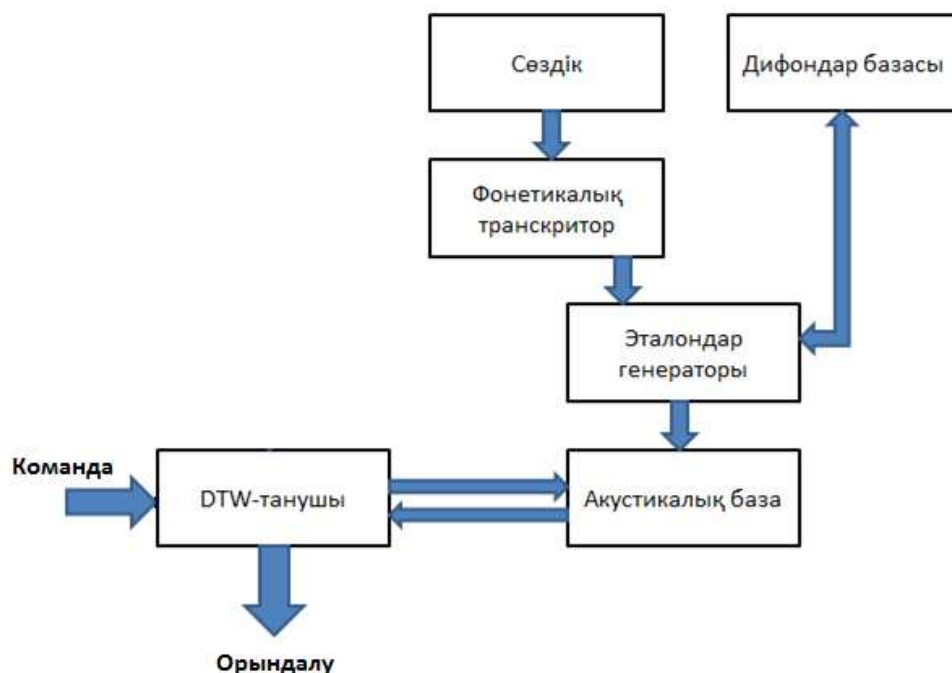


1-сурет. Жүйенің жалпы құрылымы

Программа іске қосылған кезде, сөздік автоматты түрде жүктеліп, транскриптор сөздерді жазылымнан айтылымға түрлендіріп, дифондар тізбегіне айналдырады. Эталондар генераторы сол дифондар тізбегіне сәйкес базадағы дифондар эталондарынан сөздер эталонын синтездейді. Осыдан соң программа сөздерді тануға дайын болады.

Жүйе жұмысының реті келесідей: танылушы сөз сигналы алдымен фонемааралық өңдеушіге келіп түседі де, сегменттеліп, стационар бөліктерінен арылады. Содан соң алғашқы дыбысты танушы іске қосылып, алдымен алғашқы дыбысты анықтайды. Содан соң сигналды сөздер эталоны базасындағы эталондармен салыстырып, ең аз DTW – қашықтықтағы сөзді экранға шығарады.

Жарнама компаниясының веб-жүйесіне ендірілгенде құрылымы келесідей болмақ:



2-сурет. Бейімделген жүйе құрылымы

Яғни, сөздік көлемі кішкентай болғандықтан, түскен сигнал фонемааралық өңдеуді және алғашқы дыбысты тануды қажет етпейді. Бірден базадағы эталондармен салыстырылып, ең аз DTW – қашықтықтағы команда орындалады.

Қолданылған әдебиет тізімі

1. <https://codeo.kz/blog/intelekt/204.html>
2. Карабалаева М.Х. Система распознавания целых слов с использованием динамического выравнивания времени // Қазақстан Республикасы Ұлттық ғылым академиясының хабаршысы. – 2009. – №6. – С. 7-11.
3. Бекманова Г.Т. Некоторые подходы к проблемам автоматического словоизменения и морфологического анализа в казахском языке // Вестник Вост.-Каз. гос. тех. университета им. Д. Серикбаева. - Усть-Каменогорск, 2009. – №4. – С. 192-197.
4. Yessenbayev Zh. The behavior of acoustic parameters in the noise // Вестник Евразийского нац. университета им. Л.Н. Гумилева. – 2010. – №6(79). – С. 101-106.
5. Yessenbayev Z., Saparkhojayev N., Tibeyev T. Implementation of the intelligent voice system for Kazakh // World Applied Sciences Journal. – 2013. – Vol. 28 (6). – P. 814-816.
6. Бурибаева А.К., Дорохина Г.В., Ниценко А.В., Шелепов В.Ю. Сегментация и дифонное распознавание речевых сигналов // Труды СПИИРАН. – СПб, 2013. – Вып. 31. – С. 20-42.