



Garai batean idazmakina mekanikoak (Underwood zaharrak, adibidez) urte askotan erabili ziren. Gero, 70.eko hamarkadan, IBM etxeak bolazko makina elektrikoak atera zituen, eta azkenaldian mikroordenadoreak ugaltu dira erabat testuak idazteko. Ordenadorean izan ere, testu-prozesaketaren bidez gauza asko egin daiteke: zati bat ebaki, beste leku batean itsatsi, letra-motak aldatu, lerro-tartea handiagotu, etab. Dena den, bai makina zaharretan

dez pantailan zuzenean idatz daiteke (orrian idatziko balitz bezalaxe). *Newton* izena duelarik, Apple etxeak atera du honelako aparatua. PDA (*Personal Digital Assistant*) da oro har aparatua hauek izendatzeko erabiltzen den hitza, eta ordenadorea, agenda eta oharra idazteko "orria" (pantaila, alegia) ditu funtsean.

Ordenadore hauek erabiltzeko ez dago gauza handirik jakiteko premiarik; idazten jakitea aski baita. Baina ordenadoreak bere barnean informatikako sistema garatuenak ditu: pantaila sentikorra, RISC prozesadorea, sare neuronalak, etab. Horiek guztiak behar dira edozein umek egin dezakeen lana burutzeko; idatzitako letrak ezagutzeko, alegia. Eragiketa hori guk egunero automatikoki egiten dugu irakurtzen dugunean, eta intuizioz letrak baino gehiago hitzak identifikatzen ditugu. Hitz ezezaguna topatzen dugunean, lehenbizi dezifratu eta esanahiaz jabetu behar izaten dugu, eta gero identifikatzen dugu erraz. Testuingurutik ezagutzen ditugu karaktereak. Baina horrelakorik ezean, makinak nola jakingo du I (i maiuskula) ala I (ele minuskula) den? 0 (zero) ala o (bokala) den? Ordenadorea, egia esan, adimenez ez da ongi osatua egoten eta idazkera ezagutzeko arazo latzak izaten dira.

Ordenadoreek horretarako OCR (*Optical Character Recognition*) izeneko programak erabiltzen dituzte. 80.eko hamarkadan hasi ziren programa hauek egiten, etorkizun oparoa ikusi ziotelako. Hainbat banketxe, administrazio-sail eta bulegotan milaka tona paperetan zeuden dokumentuak informatiza zitezkeen, eta bide batez milioika dolarreko merkatua zegoen programagileentzat.

Karaktereak ezagutzeko erabiltzen den edozein teknologia, lehenbizi numerizazio-etapa izaten du. Dokumentua eskaner batetik pasatzen da, hortik irudi

Teklaturik gabeko ordenadoreak

Jon Otaolaurretzi*

Ordenadoreak orain arte beti ikusi izan ditugu beren teklatu eta guzti, baina azkenaldian teklaturik gabe funtzionatzen dutenak ere azaldu dira. Norberaren idazkera ezagutzen dute eta lapitz batez zuzenean pantailan idatzita jasotzen dituzte datuak.

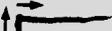
eta bai ordenadorean, beti dago teklataua, inprimatu aurretik testu idazteko.

Gaur egun ordea, nonahi ohiturak berritzen ari zaizkigun honetan, teklaturik gabeko ordenadoreak azaldu dira. Miniaturizazioan egindako aurrerapenei esker, ordenadore-fabrikatzaileek lehengoak bezain makina ahaltuak orrierdiko tamainan eta 2 zentimetroko lodieran sartzea lortu dute. Lantegian, busean, kalean eta edozein tokitan erabil daitezke. Izan ere, 500 gramoko pisura iritsi gabe aparatua hauek pantaila eta prozesurako unitate zentrala baitituzte. Teklatuaren ordez lapitza dute eta horren bi-



TEKNOLOGIA

Mikroordenadoretan xaguari esker batzuetan teklatura alde batera utz dezakegu. Lapitzak xagua baztertuko ote du? Ez da harritzekoa izango. Izan ere, programagintzan ari diren etxeak pantailan lapitzez idazteko oinarriko eragiketak agintzeko ikurrak sortzen hasiak baitira. Ondoan erakusten dira horietako batzuk:

Testua onartu	Ebaki	Kopiatu	Erantsi	Anulatu
				
Ezabatu	Hitz osoa ezabatu	Tartea	Tabulazioa	Lerrora itzuli
				

informatikoa ateratzeko. Eragiketa horretarako eskanerrak argi-diodoko barra du, dokumentuaren zati guztiak banan-banan argitzen dituelarik. 0,1etik 0,5 zabaleraino zatiak hartzen ditu, eta paperean isladatutako argia eskanerrak analizatu egiten du, zuria ala beltza den erabakitzen duelarik.

Eragiketa hau amaitutakoan, puntu beltz eta zuriz osatutako irudia izaten da, baina puntu horiek ez dute inolako antolaketarik. Ezagutzeko programak erabakiko du gero puntu-multzo horiek zein karaktere diren. Karaktereak identifikatu ondoren, sistema bitarrean kodetuko dira, gero ordenadorean erabili ahal izateko.

Karaktereak identifikatzeko bi sistema daude: matriziala eta topologikoa. Sistema matrizialean, programak eredu zerranda edo bilduma du eta dokumentuaren aurkitzen duen karaktere bakoitza konparatu egiten du bere bilduman dituenekin. Arazo asko egon ohi da, ordea. Dokumentuko karaktereak elkarrengandik hurbilegi badaude, adibidez, zaila da sistemak banaka trata ditzan. Kasu horretan puntu-multzo handia topatzen du eta bere bilduman duen ezeren antzik ez du. Gainera jatorrizko dokumentu

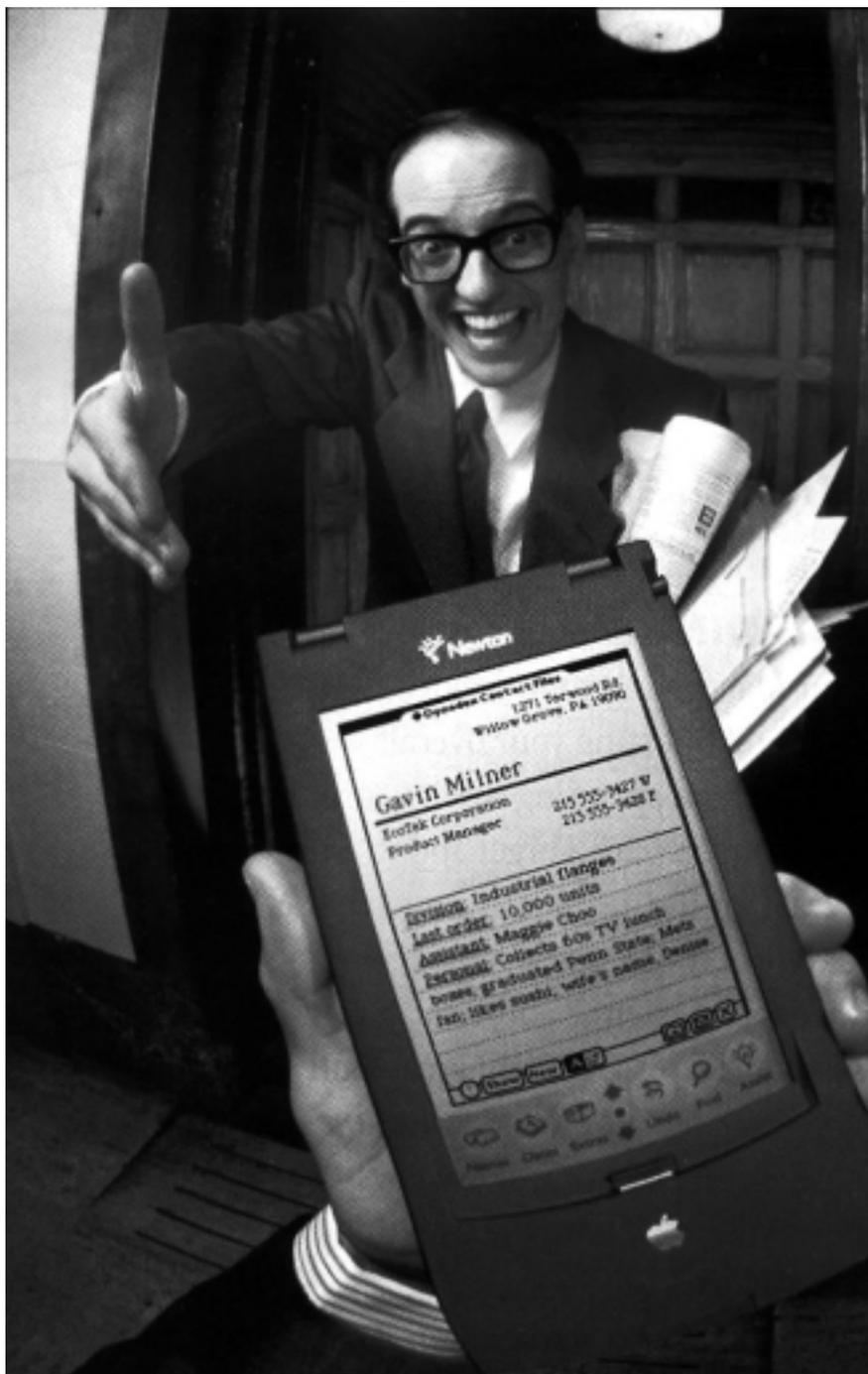
edo papera kalitate onekoa ez bada (paper txarra, horizatua, etab. baldin bada), karaktereak identifikatzea zaila da. Horrez gain batzuetan gerta liteke jatorrizko dokumentuan letra-mota berezia erabiltzea ere, eta programak bere bilduman horrelakorik ez edukitzea. Kasu horretan programak karaktere horiek "ikasi" egin behar ditu. Programak ezagutzen ez dituen karaktereak banan-banan erakusten ditu, eta erabiltzaileak teklaturik identifikatu egin behar ditu. Horrela identifikatutako karaktereak bildumara eranstean dira, hurrengo batean baliagarri izan daitezten.

Metodo topologikoa, karakterearen definizio logikoan oinarritzen da eta lerro bertikalak, horizontalak, kurbak, marrak, etab. hartzen ditu kontutan. Programak oinarriko forma hauek ezagutzen ditu eta gero karaktere bera zein den identifikatzen du. Hurbilketak eta finketak behin eta berriz eginez, programa hauek karaktereen % 96 baino gehiago zuzen identifikatzen dituzte. Horrek idatzitako orrialde normalean batezbeste 180 karaktere ezagutu gabe geratzen direla esan nahi du. Irakurle arrunta huts horietaz erraz konturatzen da eta testua zuzen nola den berehala jakiten du. Baina

hutsak dituen testu hori gero informatizatuta erabili behar bada, lortzen den emaitza kasu askotan ezin da onartu. Horregatik, karaktereak hobeto identifikatzeko sare neuronaletara jotzeko joera dago. Sare hauen bidez, esate baterako, "o" karakterea zirkulu perfektua bada, "o" bokala dela identifikatzen da, eta "0" zirkulua perfektua ez bada, zero zifra dela identifikatzen du. Sare neuronalek, neuronak erabiltzen dituzte oinarriko elementutzat eta elkarrekin sinapsien bidez konektaturik daude (Ikus *Elhuyar. Zientzia eta Teknika*. 79. alea. 1994.eko urtarrila. "Sare neuronal artifiziala". A. Tapia & J. Florez). Sistema honen oztopo nagusia, ahalmen informatiko handia behar izatea da, eta gaur egun ditugun mikroordenadoreak ez dira horretara iristen. Gaur egun aplikatzen den irtenbide bat, neurona-sareak eta

Eskuz pantailan orri zurian bezalaxe idatzita, teklaturik gabeko ordenadoreak idazkera ezagutu eta digitalizatuta gorde egiten du. Testu hori gero, informatikoki prozesa daiteke.





ipintzen denez gero, banaketa automatikoki egiten da. Banaketan ere ganbio-letrak identifikatzeko sistema automatikoak erabiltzen dira, baina, jakina, sistema horiek akatsik ezin dute eduki, dirua tartean delako.

Badira ordea, eskuz idatzita sartutako datuak ezagutzen dituzten sistemak ere. Datuak pantaila batean idazten direnez, hauek eskuzko idazkeran izaten diren esfortzu mekanikoak jasateko gai izan behar dute. Pantaila, taula edo PDA hauen bidez informatikaren industriari aro berria zabaldu zaio idazkera ezagutzeko programak medio. Apple etxearen Newton erredua aurrerapauso handia izan da horretarako. 1.000 dolar balio ditu teklaturik gabeko ordenadore honek, baina EO erredua ateratzekotan da eta 2.000-3.400 dolar inguru balioko du. Idazkera ezagutzeko ahalmenaz gain, telefonoa eta faxa integratuak izango ditu, eta horrek guztiak, teklatu eta kablerik gabe funtzionatuko du.

* ZETIAZ - Elhuyar

metodo tradizionalak nahastea da. Programak lehen bueltan metodo matrizial edo topologikoz identifikatzen du, eta ondoren, problemak dituzten karaktereak baino ez dira sare neuronalaz tratatzen.

Mikroordenadoreen MS DOS sistema eragilea egiten duen Microsoft etxeak, bere sistema garatu du. Programak karaktereak banan-banan aztertzen ditu,

eta dudazkoren bat topatzen duenean, "metakarakterez" baliatzen da. Adibidez, "ha(nr)i" (hau da, "n" ala "r" den) zalantza duenean, hiztegi ortografiko batez erabakitzen du.

Idazkera ezagutzeko ordenadore-programek aplikazio anitz dituzte jadanik edozein bulegotan. Aplikazio horietako bat, posta-bulegoetakoa da. Izan ere, gutunetan azalean posta-kodea

