

Ikasteko gai al dira robotak?

Ana Zelaia Jauregi*

Zergatik dute robotek ikasteko beharra?

Eguneroko bizitzan makina bat iharduera burutzen dugu ia pentsatu beharrik izan gabe. Erabat arrunta deritzogu esaterako, ezagutzen ez dugun gela batera sartu eta gela horretako objektuak jo gabe ibiltzeari.

Ingurunea edozein izanik, robotak ere bere helburua lortzeko gai izan behar du lan erabilgarriren bat egitea nahi badugu. Gaur egungo kon-

Robotak, bizitza errealean lagungarri izan dakigun, lan egin behar duen ingurunera egokitze gauza izan behar du, hau da, ingurune horretan lan egiten ikasi egin behar du.

putagailuen konputagarritasun-ahalmena handia dela kontutan hartuz, zergatik ez eman robotari ingurunearen deskribapen zehatza? Deskribapen hori erabiliz izan ere, robota bere betebeharra burutzeko gai izango da, ingurunea nolakoa den eta bertan nola jokatu behar duen ikasteko beharrik izan gabe.

Hori guztia jostailuzko mundu artifizialen batean posible izan zitekeen, baina ez bizitza errealeko inguruneetan. Robotari ezin izango diogu ingurunearen deskribapen zehatzik eman, konplexuegia izan daitekeelako edo ingurunea bera aldeztetik ez dugulako ezagutzen. Beharbada, robot hori ingurune batean baino gehiagotan lan egiteko gai izan dadila nahiko dugu. Eta, ingurunea dinamikoa edota aldakorra balitz? Bertan jendea

atzera eta aurrera ibiliko balitz? Nola programatu robota ingurune horietan lan egiteko?

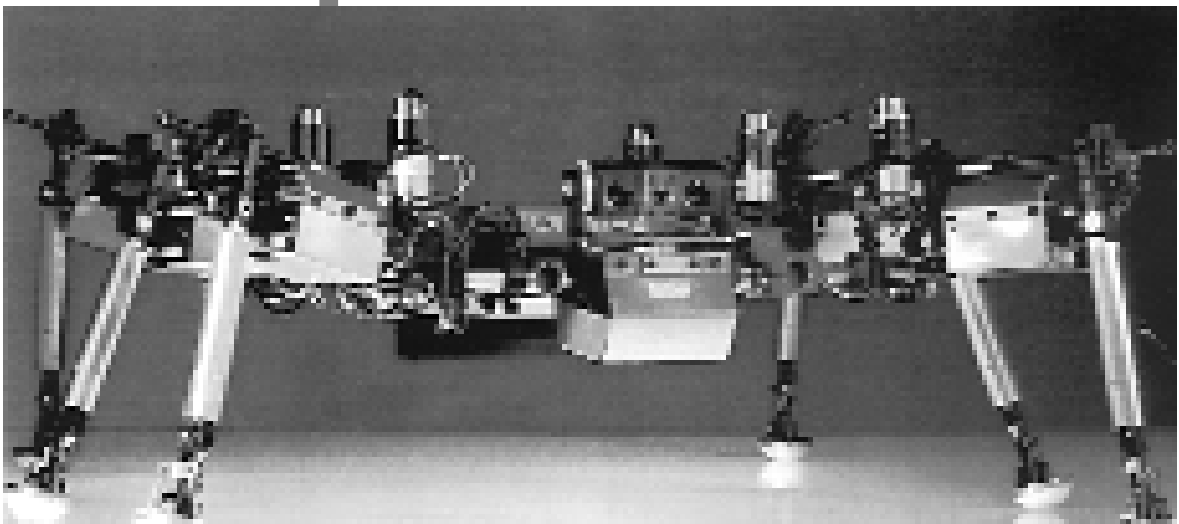
Garbi dago beraz, robotak, bizitza errealean lagungarri izan dakigun, lan egin behar duen ingurunera egokitze gauza izan behar duela, hau da, ingurune horretan lan egiten ikasi egin behar duela. *Ikasi* diogunean, zera esan nahi dugu: robotak bere jokaera hobetzeko gai izan behar duela, jokaera hori unean uneko ingurunearekiko egokiagoa eginez.

Indarkari bidezko ikasketa

Adimen artifizialean asko dira ikasteko erabiltzen diren metodoak. Hemen aztertu nahi den ikasketa metodoa *indarkari bidezko ikasketa* da, izena indarkari deritzon balioetik datorkiolarik. Indarkaria zer den ulertzeko azter dezagun ikasketa-prozesua pausoz pauso.

Robotak bere portaera ingurunean egokitze, inguruneari buruzko informazioa jaso beharko du. Hori sentsoreen bidez egiten du: argi-sentsoreak, ukipen-sentsoreak, ikusmen-sistema artifiziala, etab.

Robotak ingurunetik informazioa jaso eta gero, burutu ditzakeen ekintzen artetik bat aukeratu eta exekutatu du (Adibidez: pauso bat eman aurrerantz). Horren ondorioz, ekintza zenbaterainoko egokia izan den adierazten duen datu bat jasoko du robotak. Balio horri *indarkaria* esaten zaio.



Positiboa izango da (saria) burutu-tako ekintzak egoera hobegora eraman badu robota (hobea beti ere, lortu behar duen helburuarekiko) eta negatiboa (zigorra) ekintza burutu aurreko egoera baino okerragoan utzi badu. Robotak esperientzia guzti horiek gorde egin beharko ditu, hurrengo egoeratan ekintzarik komenigarriena erabaki ahal izateko.

Testuinguru horretan "ikastea" robotak jasoko duen balio indarkaria maximizatzeko une bakoitzean exekutatu behar duen ekintza zein den erabakitzen jakitea da.

Demagun bi gurpilen bidez higitzen den robot sinplea dugula eta ingurunean badela argi-iturri bat, robota argi-iturrirantz higitu dadila nahi dugularik. Ingurunetik informazioa jasotzeko bi argi-sentsore ditu: ezkerreko zuzenduta bata eta eskuinaldera bestea. Robotak bost iharduera-mota buru ditzake: gelditu, aurrera joan, atzera etorri, ezkerreko biratu eta eskuinaldera biratu. Sentsoreen bidez, hiru egoera desberdin bereizteko gai da: ezkerreko sentsoreak neurtutako argitasun orokorra eskuinaldekoak neurtutakoa baino handiagoa dela, txikiagoa dela eta berdina dela.

Egoera horietan egonik eta iharduera horiek exekutatzeko ahalmena izanik, robota ingurunean higitzen hasiko da. Jasoko duen indarkaria positibo altua izango da egoera berrian neurtutako argitasun orokorra aurreko pausoa neurtutakoa baino handiagoa bada.

Gauzak horrela, robotak ikasi behar duen egoera eta ekintza arteko egokitzapena ondorengoa izango da:

- Ezkerreko argitasuna eskuinekoa baino handiagoa bada, ezkerreko biratu.
- Ezkerreko argitasuna eskuinekoa baino txikiagoa bada, eskuinaldera biratu.
- Ezkerreko argitasuna eta eskuinekoa berdinak badira, aurrera jarraitu.

Algoritmo sinple bat: tartearen estimazioa

Algoritmo hau estatistikan oinarritzen da, Bernoulli-ren probabilitate-banaketan, hain zuzen ere.

Robotak ingurunean eragiten duen neurrian, jasotako informazioa gorde egingo du. Horretarako bi dimentsioko bi taula erabiltzen ditu, taulako laukitxo bakoitza egoera/ekintza bikote bati dagokiolarik.

- Lehen taulan, egoera/ekintza bikotea zenbat aldiz esperimentatua izan den gordetzen du.
- Bigarrenetan, egoera/ekintza bikotearen komenientzia edo egokitasun-maila gordetzen du.

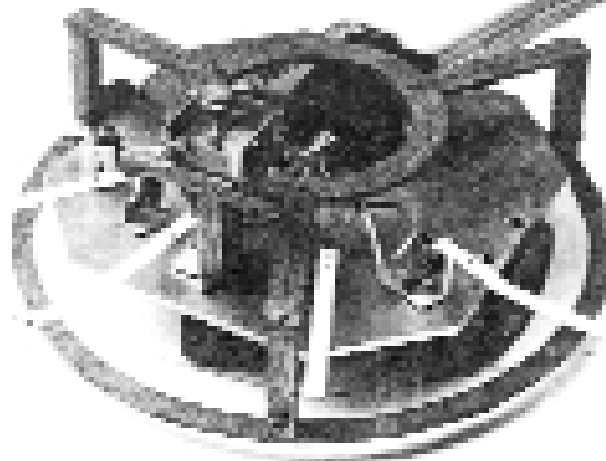
jakingo du, robotak erakusten duen jokabidea optimoa izan dadin.

Spanky robota

Tarte-estimazioaren algoritmoa darabil robot honek. Bi gurpilen bidez higitzen da. Argitasun-mailari buruzko informazioa jasotzeko lau argi-sentsore eta inguruko objektuak

Robotak ingurunetik informazioa jaso eta gero, burutu ditzakeen ekintzen artean bat aukeratu eta exekutatu du. Horren ondorioz, ekintza zenbaterainoko egokia izan den adierazten duen datu bat jasoko du.

Spanky robota



Ingurunean ekintzaren bat exekutatzeko duen bakoitzean bi taula horiek eguneratu egingo ditu: lehengoan egoera eta ekintza horiei dagokien laukitxoko balioa 1-ek gehituko du eta jasotzen duen indarkaria (positiboa edo negatiboa) bigarren taulan dagokion laukitxoan gehituko zaio.

Bi taula horiek izanda eta kalkulatu estatistikoak erabiliz zera lor daiteke: robota egoera batean egonik, ekintza bat exekutatu gero 1 balio indarkaria lortzeko probabilitaterako estimatzen den konfidantza-tartea. Tarte hori izanda, algoritmoak zein ekintza exekutatu

detektatzeko beste bost sentsore ditu. Horrez gain, robota sentsore biribil batez inguraturik dago talka detektatzeko.

Gailu guzti horiek direla medio, robotak inguruneari buruzko informazioa 5 bitetan jasotzen du.

- 0,1 bitak: Argitasun handiena nondik datorre robotari.
 - 00: aurrealdetik
 - 01: ezkerrealdetik
 - 10: eskuinaldetik
 - 11: atzealdetik
- 2 bita: Eskuinaldeko talka-sentsoreak objektuen batekin topo egin al du? Bai ala ez.

- 3 bita: Ezkerraldeko talka-sentsoreak objekturen batekin topo egin al du? Bai ala ez.
- 4 bita: Erdiko hiru talka-sentsoreetatik gutxienez batek objekturen bat jo al du? Bai ala ez.

Robotaren jokabidearen arabera, ondorengoak dira jasoko dituen indarkariak:

- Baldin robotak zerbaiten aurka jo badu: **indarkari negatiboa**
 - Banaketaren batezbestekoa: -2
 - Desbidazio estandarra: 0.5
- Robotak ez badu bere burua ezeren aurka jo: **indarkari positiboa**
 - Baldin argitasuna(orain) > argitasuna(lehen) orduan
 - * Batezbestekoa: 1
 - * Desbidazioa: 0.2
 - Baldin argitasuna(orain) ≤ argitasuna(lehen) orduan
 - * Batezbestekoa: 0
 - * Desbidazioa: 0.2

Guzti hori horrela izanik, Spanky robota objektuen aurka ez jotzen eta argi-iturrirantz abiatzen saiatzen da. Spanky 20 aldiz baino gehiagotan jarri zuten martxan. Guztietan estrategiarik hoberena ikasteko gai izan zen 2-10 minutuko tartean.

Gerora jasotako indarkaria

Ikusitako algoritmoak ekintza burutu ondoren jasotako indarkariaren arabera aldatzen du robotaren jokabidea.

Alabaina, robotak jasotako indarkari guztien batura maximizatzea da guri interesatzen zaiguna eta horretarako, ekintza bakoitza burutu ondoren jasotako indarkariak gain, gerora jasoko dituen indarkariak ere kontutan hartu behar dira. Beharbada, beste ekintza bat burutzeagatik ez litzateke indarkariarik handiena jasoko, baina luzarora begira, jasoko liratekeen indarkari baturak kontutan hartuz gero, handiagoa izango da.

Bada horretarako algoritmorik. Q ikasketa izena duen algoritmoa horietako bat da, eta nahikoa ona gainera. Algoritmoak Q ikasketa izena du, Q balioak gordetzen dituelako. Zer adierazten dute Q balio hauek? t unean dago robota eta i sarrera-egoera jasotzen du. Nola edo hala, a ekintza exekutatzea erabakitzen du une horretan eta r indarkaria jasotzen du. Gauzak horrela, aurrerantzean ere estrategiarik hoberenari jarraituz jokatuko duela suposatzen du robotak eta $r(1), r(2)...$ indarkariak jasotzea espero du. Geroago jasotzea espero dituen $r(j)$ hauek, $j = 1...∞$, kontutan izango dira, baina zenbat eta denboran urrutia-egoa egon, orduan eta eragina txikiagoa izango da. Hau γ deskontu-faktoreari esker lortzen da. Hasieran γ balioa 1-aren gertu egongo da eta gero eta txikiagoa izango da 0-ra hurbilduz.

$$Q(i(t), a(t)) = \sum_{j=0}^{\infty} \gamma^j r(t+j)$$

Q balioa: a ekintzaren erabilgarritasuna i egoeran

non

t : denbora

γ : deskontu-faktorea, $0 < \gamma < 1$

$r(t)$: t unean jasotako indarkaria

Egoera batean ekintza bat exekutatzeak zer nolako erabilgarritasuna duen adierazten du Q balioak.

Horrelako balio bat gordeko du egoera eta ekintza bikote bakoitzetarako. Taula bat osatzen joango da robotaren esperientzia gordez.

Robotak ingurunean saioak egingo ditu eta jasotzen dituen emaitza guztiak bere taulan gordeko ditu, aurrerantzean erabili ahal izateko.

Robota halako batean i egoeran aurkituko da eta egoera horren erabilgarritasuna jakin nahiko du. Zein da egoera horretan egonik exekutatu beharko lukeen ekintza gerora jasoko duen indarkaria maximoa izateko? Taulan begiratu eta egoera horretarako Q baliorik handiena duen ekintzak duen balioa izango da erabilgarritasun hori eta ekintza hori, ez besterik, burutzea erabakiko du.

$$U(i) = \max_a Q(i, a)$$

i egoeraren erabilgarritasuna.

Daukagun informazioarekin nahikoa daukagu ikasketa-algoritmo eraginkorra eraikitzeko. Algoritmo horrek taulako Q balioak eguneratu egiten ditu etengabe. Bere funtsezko lana horixe da, hortik eratorriko baitu robotak bere jokabidea. Horrela aldatuko ditu algoritmoak taulako balioak:

$$Q(i(t), a(t)) = r(t) + \gamma U(i(t+1))$$

Q balioen taula.

	a_1	a_2	...	a_j	...	a_n
i_1	$Q(i_1, a_1)$	$Q(i_1, a_2)$...	$Q(i_1, a_j)$...	$Q(i_1, a_n)$
i_2	$Q(i_2, a_1)$	$Q(i_2, a_2)$...	$Q(i_2, a_j)$...	$Q(i_2, a_n)$
.						
.						
i_j	$Q(i_j, a_1)$	$Q(i_j, a_2)$...	$Q(i_j, a_j)$...	$Q(i_j, a_n)$
.						
.						
i_m	$Q(i_m, a_1)$	$Q(i_m, a_2)$...	$Q(i_m, a_j)$...	$Q(i_m, a_n)$



Demagun t unean gaudela, robotak $i(t)$ egoeran ikusten duela bere burua eta $a(t)$ ekintza exekutatu duela. Horren ondorioz, $i(t+1)$ egoeran aurkitzen da eta r indarkaria jaso du. Orduan, algoritmoak Q balioen taulara joko du eta $i(t+1)$ egoera berriaren $U(i(t+1))$ erabilgarritasuna kalkulatu du, lehen adierazi bezala. Erabilgarritasun hau γ parametroaren bidez deskontatu du, jaso berri duen $r(t)$ indarkariari gehitzeko. Guztiarekin, exekutatu berri duen $a(t)$ ekintzaren erabilgarritasuna dauka, eta taulan dagokion lekuan gordeko du, aurrerantzean behar izanez gero erabiltzeko. Badaki zer nola joan zaion a ekintza i egoeran; zenbateraino ona izan den berau exekutatzea.

Robota munduan eragiten hasi berria denean, taulan gordetzen dituen Q balioak ez dira oso esanguratsuak, baina algoritmo horren egile den Watkins-ek frogatu zuenez, algoritmoak laster jotzen du benetako Q balioetara. Behin konbergitu duenean, robotak ikasi egin du ingurune horretan mugitzen eta estrategia hoberenari jarraituz jokatuko du beti. Algoritmoa dinamikoa da eta etengabe taulako Q balioak aldatuz doa aldakorra izan litekeen ingurunera moldatu asmoz.

Zertan ari zara pentsatzen? Konputagarritasun-komplexutasuna izugarria izan daitekeela? Ba, bai! Hala da. Ezin dena eduki...

Arazo izugarria da konputagarritasun-komplexutasunarena. Aurreko Spanky robotaren kasuan, sarrera-egoera 5 biten bidez adierazita datorkio, hau da, $2^5 = 32$ egoera desberdin bereizteko gai da. Jakina, Spanky jostailuzko robota besterik ez da. Errealitatean robotak bit gehiagoren bitartez jasoko du sarrera-egoerari buruzko informazioa eta ekintza ezberdin gehiago exekutatzeko gai izango da. Zein tamainatako taula erabili beharko du algoritmoak? Bai espazio eta denbora aldetik, konbinazio-lehertzapena gertatzen da.

Algoritmo honek egoera desberdin guztiak zehaztuak izatea es-

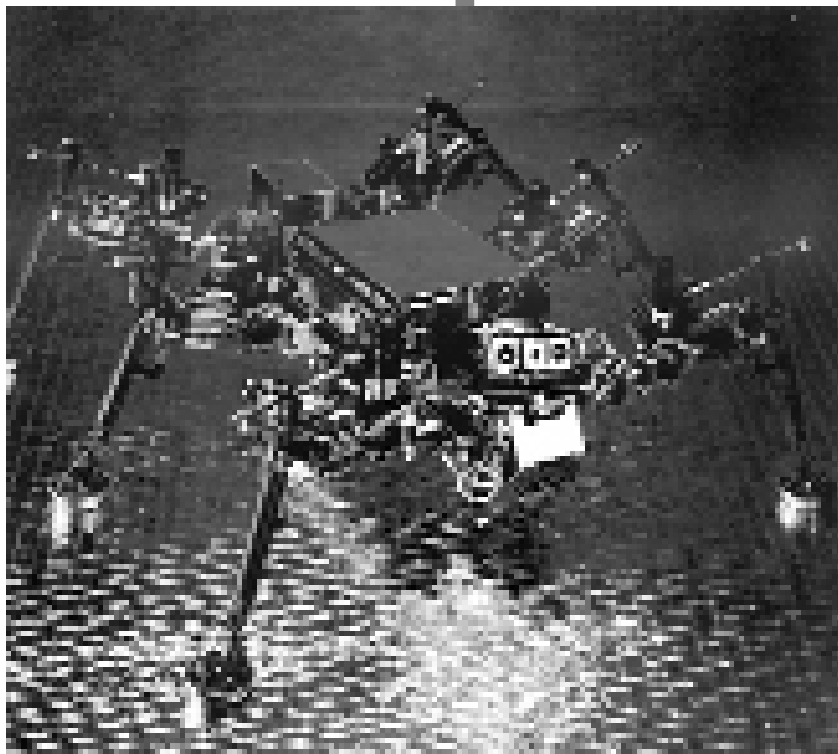
katzen du eta ikusi dugunez, eskatzen duen espazioa izugarria da. Tamalez, ez da hori arazo bakarra.

Sarrera-egoeren orokortzearen arazoa

Izaki bizidunaren jokaera aztertuta, egoera berrietan nola jokatu behar duen ez badaki ere, antzeko egoeratan izandako esperientzia erabiltzen duela ikus dezakegu. Horrek adierazten digu izakia orokortzeko gaitasuna dela. Normalean egoeren artean nolabaiteko erlazioa dago eta zenbait egoera oso antzekoak dira. Egoera baten aurrean nola jokatu

informazioa bit-segiden bidez datorkio. Bit-segida horien arteko Hamming distantzia kalkula daiteke; bi bit-segidek duten bit desberdinen kopurua adibidez. Primeran! Orokortzearen arazoa gaindi daiteke. Hala ere, egoera desberdin guztiak gorde behar al

Bide onetik al goaz? Asko dira arazoak. Zer esanik ez. Baina orain arteko saiakuntzak oso positiboak izan dira eta robotak ikasle on bihurtzeko gaitasuna dugu. Lan pixka bat gehiago egitea besterik ez da behar.



baldin badakigu egoera horri buruzko esperientzia badaukagulako, antzeko egoera desberdin eta berri batean egonez gero, antzera jokatuko genuke.

Nola jakin dezake robotak egoera bat beste baten antzekoa dela? Hori jakiteko modurik balu, egoera berrietan nola jokatu jakiteko antzeko egoeren esperientzia erabiliko luke, antzeko ekintzak sortzeko.

Arazo hau konpontzeko eratako bat Hamming distantzia erabiltzea da. Robotari egoerari buruzko

dira? Konbinazio-lehertzapenaren arazoa hor dago oraindik!

Hori gainditzeko ere izan dira saialkerak, talde estatistikoak erabiliz. Dena den, lan asko dago egiteko oraindik.

Bide onetik al goaz?

Zer deritzozue? Gauzak badabiltzala ematen du, ezta? Oraindik aipatu ez badugu ere, bada guzti honetan arazotxo bat.

Azaldutako algoritmo guztietan robota egoera jakin batean dagoela emanik, eta ekintza bat burutu ondoren, beste ezer kontutan hartu gabe hurrengo unean zein egoeratan aurkituko den badakigula suposatu dugu, egoeren arteko trantsizioa Markoviarra dela, alegia. Aitzitik, ez dugu inoiz oraingo egoeran egon aurretik zein egoeratan egon garen kontutan hartu eta hau garrantzitsua izan daiteke.

Jo dezagun robota azkarrak dela eta aurrerantz doala ziztu bizian. Halako batean, aurrealdean oztoporen bat dakusalako edo, eskuinaldera biratzeko erabakia hartuko du. Biratu ondoren zein egoeratan aurkituko da? Gure algoritmoaren arabera, robotak angelu zuzena deskribatzen duela diogu, baina hori hala ez dela denok dakigu. Orain arteko esperimenduetan erabilitako robotak nahikoa mantsoak dira eta arazoa ez da oso nabaria izan. Horrela jarraituz gero ordea, gaizki goazela konturatu gara, robotak egiten duena eta egiten duela uste duena oso desberdinak izatea gerta daiteke eta.

Asko dira arazoak. Zer esanik ez. Baina orain arteko saiakuntzak oso positiboak izan dira eta robotak ikasle on bihurtu daitezkeela ikusi dugu. Lan pixka bat gehiago egitea besterik ez da behar.



* CAF-Elhuyar sarietara aurkeztutako artikulua.

Ezustekoetan lagun

Zure aurpegia ezaguna egiten zait, baina ez naiz zutaz oroitzen... esaldia makina bat aldiz entzun eta erabili behar izan dugun leloa da, memoria-ahalmena eskaxegia gertatzen zaigula aitortzeko beldurrez. Arazoa ez da berria. Bai ordea Tokyoko **Casio Computer** konpainiak aurki eskainiko duen irtenbidea. Enpresak aurrikusitako epeak betetzekoan, urtea amaitu baino lehen merkaturatu izango da aurpegiak eta izenak erlazionatuz gure memoria-ahalmena erraz gaitutako duen sistema berria.

Ordenadore eramangarrien itxura izango du eta erabilerraza izango dela agindu dute. Funtzionamendurako aukeratu duten ereduak aurpegiaren ezaugarriak sailkatzean datza. Pertsona bakoitzaren irudia gordetzeak, izan ere, ikaragarriko memoria-gaitasuna eskatzen du. Aurkitu nahi den pertsonaren izena bila dezan, aurpegiaren ezaugarri nagusiak azaldu beharko dizkiogu: ile horia, begi urdin, etab. Datuak sartu orduko pantailan azalduko da aurpegi horri dagokion izena. Kale-erdian lotsagarri ez geratzeko sistema aparta. Oztopo bakarra, jakina, bere prezioa da.

3000 urte pasa eta gero, atsedetik ez

Medikuntzan arunt ezaguna den TC edo tomografia konputarizatuari esker, duela 3.000 urte Tebasen hil zen emakumezkoaren gorpua zehatz-mehatz aztertzea lortu da. Ez da garai bereko gorpuren bat aztertzen den lehen aldia, baina teknika horri esker, lehen aldiz ez da sarkofagoaren balio artistikoa galtetu. Tomografia erabiliz izan ere, sarkofagoaren kanpoaldeko arte-lana ez da ukitu ere egin behar eta pantaila batean beritzen da estalkiak ezkutatzen duen gorpua. Torontoko Ospitaleko ikerlariak tomografoaren bidez lortutako X izpizko irudiak eskaneatzean digitalizatu dituzte eta hiru dimentsiotako informatikako pro-

graman integratu. Era horretara, sarkofago emakumezkoaren gorpua "ikus" eta "azter" daiteke balio artistiko itzela duen kanpoaldea manipulatu gabe. Erradiazioaren erabilera neurririk gabekoa izan daiteke, hilik dagoenez gero bestelako kalterik ezin baitzaio egin. Horrek zehaztasun handiko irudiak lortzea ahalbidetu du eta horri esker dentsitate desberdineko atalak bereiztea lortu da.

Tomografoa erabiliz besteak beste, antzinako Egipton erabilitako medikuntza-teknikei buruzko informazio asko lortuko omen da. Orain arte argitu dutenez amaitzeko, ikerlan horretan aztertutako emakumezkoaren aginetako infekzio batek jota hil omen zen.

Medikuntzan arunt ezaguna den TC edo tomografia konputarizatuari esker, duela 3.000 urte Tebasen hil zen emakumezkoaren gorpua zehatz-mehatz aztertzea lortu da sarkofagoa ukitu ere egin gabe.

