

Denbora-lerroak

Olatz Arregi Uriarte
Donostiako Informatika Fakultateko irakasle eta ikertzailea



Denbora-lerroak gertaerak denboran grafikoki adierazteko modu bat dira, hau da, gertaerak ordena kronologikoan kokatzeko era. Mota askotakoak izan daitezke denbora-lerroak, eta eskala eta modu ezberdinetan adieraz daitezke. Baina, nola atera automatikoki testu batetik zein jazoera gertatu den zeinen aurretik?

Artikulu honen xedea testuetatik denbora-lerroak automatikoki nola sor daitezkeen azaltzea da.

Demagun honako esaldi hau irakurtzen dugula egunkarian: «Atzo arratsaldean, ekaitzaren ondoren, bandak kontzertua eman zuen». Esaldia irakurrita gai gara kontzertua ematearen jazoera noiz gertatu den jakiteko, baita denboran zehar kokatzeko ere. Are gehiago, egunkariaren data ezagututa, kontzertuaren, eta bide batez, ekaitzaren data ere ezaguna izango zaigu.



Egunkariaren data, eta oro har dokumentuen argitaratze-data (Document Creation Time, DCT), oso baliagarria da denbora-lerroa zehazteko orduan, sistema automatikoetan gertaerak eta adierazpenak horren arabera ordenatzen direlako eta testuetan erreferentzia asko erlatiboak direlako: bihar, herenegun...

DENBORA-INFORMAZIOA

Denbora-lerroak testuetatik automatikoki ateratzeko, lehendabizi, denbora-informazioa etiketatu behar da, eta ondoren, etiketak baliatuz, kronologikoki irudikatu.

Denbora-informazioa etiketatzeko orduan, TimeML izeneko eskema erabili ohi da [<http://www.timeml.org/>], eta lau etiketa bereizten dira: gertaerak, denbora-adierazpenak, seinaleak eta erlazioak.

Gertaerak denboran eta espazioan kokatuta dauden eta kausa baten ondorioz eragin jakin bat sortzen duten jazoerak dira. Gehienetan aditzak edota izenak dira: “Bertsolari Txapelketa Nagusia”, “kontzertua eman”, “ekaitza”... Denbora-adierazpenak denbora adierazteko erabiltzen diren espresio testualak dira, adibidez, “2017-07-07”, “hamarrak eta hoge”, “urrian”... Seinaleek

denbora-erlazioak adierazten dituzte, baina ez dute nahitaez agertu behar: “ondoren”, “eta gero”, “aurretik”... Erlazioak, berriz, mota ezberdinetakoak izan daitezke, eta, gehien erabiltzen direnak bi gertaeraren arteko edo gertaera baten eta denbora-adierazpen baten arteko erlazio tenporala adierazten dutenak dira.

Aurreko adibidearekin jarraituz, esaldia honela etiketatuko litzateke: «**Atzo arratsaldean** (denbora-adierazpena)/ **ekaitzaren** (gertaera) **ondoren** (seinalea)/ bandak **kontzertua eman** (gertaera) zuen». Ondoren seinaleari esker, jakin badakigu “kontzertua eman” gertaera “ekaitza” gertaera izan eta gero jazo dela. Gainera, beste erlazio bat ere aurki daiteke esaldian “kontzertua eman” gertaeraren eta “atzo arratsaldean” denbora-adierazpenaren artean. Erlazio horrek kontzertua atzo arratsaldean izan zela zehazten du. Bestalde, DCTa kontuan izaten bada, data zehatzak ere

normalean, erregelatan oinarritutako teknikak erabiliz etiketatuko dira. Bada oso erabilia den tresna HeidelTime izenekoa [<https://github.com/HeidelTime/heideltime>], hainbat hizkuntzatarako prestatua, denbora-adierazpenak testuetan etiketatzeko. Azken finean, denbora-adierazpenek antzeko patroiak errepikatzen dituzte (urteak, hilak, aste-ko egunak, urtaroak...) eta ez da zaila horretarako erregelak definitzea. Denbora-adierazpenak identifikatu ondoren, normalizatzea komeni da, eta, une horretan, DCTa erabiltzen da adierazpenari balio normalizatua emateko. Adibidez, testuan “aurten” hitza topatzen badugu, eta testua 2015. urtekoa bada, “aurten” adierazpenari emango zaion balio normalizatua “2015” izango da. Testuko informazioaren arabera, balio normalizatuak eskala bat edo beste izango du: urtea, eguna, ordua...

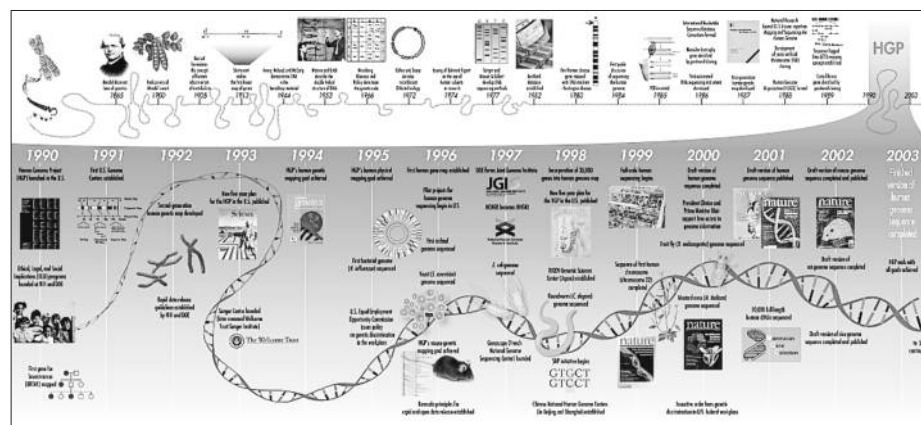
Hirugarren urratsean, erlazioak identifikatzen eta sailkatzen dira. Hau da, bi elementuren arteko lotura egiteaz gain, erlazio hori zer-nolakoa den esaten da (aurretik gertatutakoa, une berekoa...). Urrats honetarako ere, ikasketa automatikoko teknikak erabiltzen dira. Gertaeraren eta adierazpenen edota bi gertaeraren arteko bikote posible guztiak egiten dira, eta sistemak, horietako bakoitzeko, haien artean erlazioa dagoen ala ez esango du, eta, egotekotan, zer-nolakoa den.

Testuetan denbora-informazio guztia etiketatuta eta normalizatuta dagoenean, informazio hori grafikoki adieraztea besterik ez da gelditzen. Software ugari dago hori egiteko, eta eman nahi zaion itxuraren arabera bat edo beste aukeratu dugu. •

[<https://timeline.knightlab.com/>]

[<http://time.graphics/>]

[<http://www.ispringsolutions.com/blog/top-10-free-and-paid-interactive-time-line-makers/>]



osasuna / geltokiak / teknologia

3 BEGIRADA: