

Az ORBIT válasza:

PROG:

SS 4 PSTG (80)

Nyolcvan ílyet talált.

SS 5/C?

USER:

3 AND GER (LA)

(Válaszd ki a 3. utasításban találtak közül a német – GER – nyelven írottakat!)

Erre a válasz:

SS 5 PSTG (7)

hét ílyet talált a program. Végül a felhasználó a 3. utasítás szerint találtakból az 1973 után megjelent közleményeket kéri:

SS 6/C?

USER:

3 AND GREATER THAN 73

PROG:

SS 6 PSTG (30)

azaz 30 ilyen van. A keresés befejezése után a felhasználó a megtalált közlemények bibliográfiai adatait kéri. Erre szolgál a kinyomtatást kérő utasítás. A

„PRT 5, TI, AU, SO, PY”

jelsorozat beütésére például az ORBIT kiírja az 5. utasításban talált 7 db német nyelvű dokumentum címét (TI), szerzőjét (AK), megjelenési helyét (SO) és idejét (PY).

Az ORBIT rendszer nyújtotta lehetőségek közül még a NEIGHBOR (szomszéd) utasítást mutatjuk be. A felhasználó által beütött

„NB ACRYL 4”

utasítás hatására az adatbázisban fellelhető, az ábécé sorrendben az acryl szóhoz legközelebb eső négy szomszéd íródik ki, azaz

PROG:

POSTINGS

TERMS

1 ACRYDANO/(IT)

2 ACRYIC/(IT)

29 ACRYL/(IT)

2 ACRYLACETATE/(IT)

8 ACRYLALDEHYDE/(IT)

UP N OR DOWN N ?

Ahol a baloldali számok az egyes szavak előfordulásainak számát, a zárójelben lévő rövidítés (IT) pedig a minőségét jelenti, azaz az acrydano szó tehát egyszer fordul elő az adatbázisban, mégpedig deskriptorként (Index Term). Az utolsó sorban az ORBIT azt kérdezi, hogy folytassa-e a további szomszédok kiírását felfelé (up) vagy lefelé (down) a betűrendben.

Nagyobb mennyiségű bibliográfiai adatot nem érdemes a terminálon közvetlenül kiírni, mert ez drága és esetleg kényelmetlen is. Ezért az ORBIT rendszer lehetővé teszi, hogy a felhasználó levélcímét is beütve, az eredményt számítógépi nyomtatásban postafordultával megkapja.

Több hasonló on-line keresőrendszer működik már az USA-ban, sokféle szakterület adatbázisán. Bár jelenleg az on-line irodalomkutatás még viszonylag drága, a prognózisok szerint ez a módszer az USA-ban egy évtizeden belül kiszorítja a hagyományos eljárásokat.

/BUCKLEY, J. S. jr.: *Planning for effective use of on-line systems*; PREWITT, B. G.: *Searching the Chemical Abstracts Condensates data base via two on-line systems = Journal of Chemical Information and Computer Sciences*, 15. köt. 3. sz. 1975. p. 161-164; p. 177-183./

(Valkó Péter)



Automatikus referátumkészítés

Sok kísérlet történt már eredeti dokumentumok automatikus referáltatására, de eddig még ezzel a módszerrel a hagyományos referálás színvonalát nem sikerült elérni. Nem is valószínű, hogy a közeljövőben erre sor kerülne. Ennek ellenére az ilyen irányú kutatások nagy erővel folynak, mert

a hagyományos referálás költséges és időigényes;

a közeljövőben egyre több publikáció lesz, már géppel olvasható formában is hozzáférhető, ezzel viszonylag olcsó és kényelmes adatbázis hozható létre a kísérletekhez;

a referálandó szövegek gépi olvashatósága folytán az automatikus referálás olcsóbb és gyorsabb lesz a hagyományos referálásnál;

az automatikus referátum gépi továbbfeldolgozásra alkalmas formában áll rendelkezésre;

bár az automatikusan készült referátumok esetleg soha nem érik el a hagyományos módon készülték színvonalát, gyakorlatban kielégítik az igényeket, különösen megfelelő utószerkesztéssel.

A most ismertető rendszer, amely a *Chemical Abstract Service (CAS)* számára készült, két szempontból különbözik az eddig használatos rendszerektől. Egyrészt egy viszonylag szűkebb terület, a kémia – ezen belül pedig a gyógyszerkémia – tárgykörére specializált, másrészt előírták, hogy a program által készített referátumoknak meg kell felelniük a CAS szabványoknak.

Az ADAM elnevezésű referálási módszer az eredeti dokumentum egyes kiragadott részeit illeszti össze úgy, hogy azok a dokumentum tartalmát reprezentálják, de nem egyszerűen mondatokat ragad ki az eredeti szövegből, hanem kissé át is alakítja azokat. Mindazonáltal az így készült referátumban nem lesznek teljesen új, illetve összevont mondatok, ellentétben a hagyományos referátummal.

A korábbi algoritmusok előfordulási statisztika szerint választották ki a szövegből a fontos mondatokat. Az

ADAM csak *segédeszközként* *folyamodik a statisztikához*, és inkább mondatok elhagyásán, semmint kiválasztásán alapul.

Az így előállt referátum *főbb jellemzői*:

indikatív jellegű, azt szolgálja, hogy az olvasó eldöntesse, szükséges-e számára az eredeti dokumentum, tehát nem törekszik annak helyettesítésére;

az eredeti szöveg 10–20%-át tartalmazza;

a dokumentum terminológiáját használja;

a szövegből kiragadott karaktersorozatokról áll;

nem tartalmaz egyenleteket, táblázatokat, ábrákat stb.;

bevezetést, negatív eredményeket, az adatgyűjtés módszerét, magyarázatokat, véleményeket, példákat sem tartalmaz;

célok, eredményeket és következtetéseket közöl.

A rendszer működésének alapja egy *tárgyszójegyzék*, amelyikben a válogatást irányító szavak és kifejezések szerepelnek, mindegyikük szemantikai és szintaktikai kódokkal ellátva.

A *szemantikus kód* arra utal, hogy az illető kifejezést tartalmazó mondat várhatóan lényeges információt hordoz-e. A szemantikus kódok egy része negatív; negatív kód járul pl. „*az előző kutatások*”, „*jól ismert tény*”, „*lényegtelen*” stb. kifejezésekhez, úgyhogy ilyen tartalmú mondatok kimaradnak a referátumból. A „*jelen munkák során*” típusú kifejezések viszont pozitív szemantikai kódot kapnak. Bizonyos szavak, kifejezések mondatközi kapcsolatra utalnak, így a negatív kód hatása átterjed az előző, a következő stb. mondatra is.

A *szintaktikus kódok* a szavak és kifejezések nyelvtani szerepét jelzik és elsősorban a mondat-átszerkesztést segítik, segítségükkel az ADAM képes kihagyni pl. a vesszők közé tett megjegyzéseket.

Az előfordulási statisztika a tárgyszójegyzék módosításán keresztül befolyásolja a végleges eredményt. Egy

papírgyártással foglalkozó cikkben pl. a „*this paper*” kifejezés, amely az angolban „*ez a papír*” jelentést és „*ez a dolgozat*” jelentést egyaránt hordozhat, az adott esetben automatikusan „*leértékelődik*”, azaz a szokásosnál negatívabb szemantikai kódot kap, mivel sokszor fordul elő.

Végül a program *simító szerkesztést* végez, hogy a referátum megfeleljen a CAS szabványainak, így pl. rövidítéseket használ, amerikai helyesírásra ír át egyes angol helyesírású szavakat, egyszerűbb vegyületneveket képlettel helyettesít stb.

Az ADAM által készített referátumokat *összevetették a CAS referátumaival*. Az értékelés – a dolog természetéből adódóan – szubjektív elemeket is tartalmazott. Megállapítható, hogy az ADAM referátumok *tömörségben és egyéb szempontokból is elmaradnak a hagyományosak mögött, de lényegében megfelelnek a célnak*. Az eredmény nagymértékben függ a dokumentum szerkezetétől. Az automatikus referátum az ún. lineáris felépítésű cikkek – bevezetés, cél, módszer, eredmények és következtetés – esetében tükrözi leginkább az eredeti dokumentum lényegét. Mindig tartalmaznak azonban néhány olyan mondatot, amelyek első ránézésre a témakör, sőt az eredeti cikk alapos ismerete nélkül is kihagyhatók. Az *utólagos szerkesztéssel* – amely már sokkal kevesebb időt, hozzáértést igényel – *jelentősen javítható a referátum minősége*.

Figyelemre méltó tény, hogy egy referátum elkészítése *átlag 0,6 másodperc gépidőt vesz igénybe*.

/POLLACK, J. J. – ZAMORA, A.: *Automatic abstracting research at Chemical Abstracts Service. = Journal of Chemical Information and Computer Sciences*, 15. köt. 4. sz. 1975. p. 226–232./

(Valkó Péter)



SZABADALMI TÁJÉKOZTATÁS

Szabadalmi tájékoztatás Angliában

Az angol *Derwent* vállalat 1976. januárjától kezdve kibővítette gépesített tájékoztató szolgáltatását, amennyiben mind a teljes szabadalmi file-jét, mind a folyóiratcikkek file-jét bekapcsolta egy interaktív on-line keresőrendszerbe. A rendelkezésre álló szabadalmi adatbázis 1963-ig visszamenőleg 1200 ezer szabadalom, a gyógyszerészeti irodalom adatbázisa 1964-ig visszamenően 450 ezer folyóiratcikk adatait tartalmazza. Az előbbi havi 40 ezer szabadalommal, az utóbbit havi 3500 hivatkozással egészítik ki.

A vegyipar számára a visszamenőleges, gyors információkeresést az 1970 óta megjelenő *Central Patent Index (Központi Szabadalmi Index)* teszi lehetővé. A *Patent Abstracts Publications (Szabadalmi Referáló Kiadványok)* a nem vegyészeti tárgyú szabadalmi leírások ábrákkal ellátott referátumait közli. A *World Patent Index (Világ Szabadalmi Index)* heti több mint 12 ezer szabadalmi leírás legfontosabb adatainak gépi indexe.

További felvilágosítás: Derwent Publications Limited, Rochdale House, 128 Theobalds Road, London WC1X 8RP.

/FID News Bulletin, 26. köt. 2. sz. 1976. p. 22./