

## Nyelvi elemek érzelmi töltetének vizsgálata és felhasználása természetes nyelvi dialógusrendszerben

Tatai Gábor<sup>1</sup>, Laufer László<sup>2</sup>

<sup>1</sup> Department of Computer Science, University College London  
Gower Street, WC1E 6BT, London, UK  
g.tatai@cs.ucl.ac.uk

<sup>2</sup> AITIA Informatikai Rt., 1117 Budapest, Infopark sétány 1.  
l.laufer@aitia.ai

### Absztrakt

Cikkünkben bemutatjuk az általunk fejlesztett BotCom beszélgető-rendszerbe integrált GALA érzelmi elemző és generáló alrendszert. Ennek funkciója az, hogy egy felhasználóval történő természetes nyelvi beszélgetés folyamatában a beszélgetés érzelmi rétegeit is figyelembe vegye a válaszadás során. A GALA többretegű architektúrával rendelkezik, és az R. Plutchik féle érzelmi modellt használja. A rendszer jól használható rövid chat-dialógusok érzelmi rétegének kezelésére, melyből kinyert információval animált beszélgető partnert tudunk vezérelni.

**Kulcsszavak:** Affective computing, érzelmi modellezés, beszélgetőrendszer

## 1 Bevezetés

Kutatásaink során – mely részben természetes nyelvi dialógus-keretrendszer fejlesztésére fókuszál – szembesültünk a szövegek érzelmi töltetének meghatározásának és felhasználásának fontosságával. Ismeretes, hogy a számítógépes nyelvi alkalmazások elsősorban a nyelv, a mondatok szerkezeti, strukturális problémáira koncentrálnak. A nyelv formális leírásában, praktikus nyelvi alkalmazások készítésében számos magyar kutatónak és fejlesztőnek jelentős tapasztalata van. Az ember-számítógép interfészekben felhasználható számítógépes dialógusrendszerek esetében azonban felmerül az igény, hogy a szöveg által hordozott szintaktikai információ és szemantikai tartalom mellett az emberi kommunikációban jelentős szerepet játszó érzelmi réteg is detektálható, leírható, felhasználható és generálható legyen.

Az alkalmazás fontosságát több oldalról is megvilágíthatjuk. Az emberek sokkal könnyebben ismerkedhetnek meg a számítógép nyújtotta lehetőségekkel, ha kevesebbet kell tanulniuk azok használatát. A természetes nyelvi dialógusrendszerek olyan virtuális „beszélgetőpartnerként” használhatók, melyek a bevitt természetes nyelvi szöveges információt fel tudják dolgozni, és arra akár egy általános témájú beszélgetés keretében értelmes mondatokkal tudnak válaszolni, vagy a mondatokban szereplő esetleges utasításokat fel tudják ismerni, és végre tudják hajtani. A dialógusrendszerben így nem csak a gép vezérlése válik érthetőbbé, hanem a visszajelzés is jobban



keretrendszerekhez illeszthető, sok esetben animált arcral kiegészített, természetes nyelvi kérdező-válaszoló rendszer, mely mind saját, mind külföldi kísérletek alapján megnöveli a tanulás hatékonyságát.

A természetesség érzetét növeli az, hogy ha a virtuális beszélgetőtárs a kommunikáció érzelmi szintjét is érzéklni tudja és ezt az információt válaszaiban kihasználja, például úgy, hogy hasonló vagy ellentétes érzelmi töltetű válaszokat próbál adni, a beállításoktól és helyzettől függően. Beszédfelismeréssel foglalkozó kollégáink számos információhoz jutnak az intonáció/prozódia vizsgálatával. Annak érdekében azonban, hogy egy szöveges információ érzelmi töltetét elemezzük, mellyel nem „jön” ilyen másodlagos információ, szükségünk van egy olyan adatbázisra, mely tartalmaz szavakhoz, kifejezésekhez rendelt érzelmi töltetet. Továbbá értelemszerűen egy kontextuselemző is elengedhetetlen annak érdekében, hogy érzéklni lehessen a szövegkörnyezet következtében fellépő érzelmi jelleg változását.

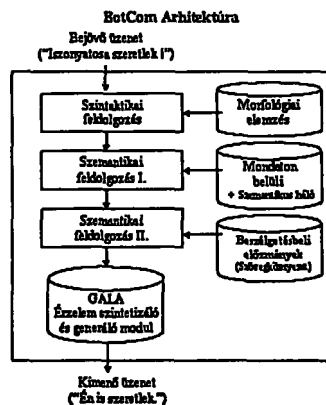
Cikkünkben bemutatjuk az erre a problémára született megoldásainkat, melyek természetesen még fejlesztés alatt vannak. Eddig eredményeink már számos nemzetközi publikációban megjelentek [2; 3; 4; 5], ezért fontosnak tartjuk azt, hogy a magyar szakmai közönség is jobban megismerje ezeket. A problémát egy szintetikus érzelmemodellező és kiértékelő rendszer, a GALA eszköz kifejlesztésével próbáljuk megoldani. Ez az eszköz tartalmaz egy Robert Plutchik által ismertett geometriai alapú érzelmi modellt [6; 7], erre építettünk egy további réteget, amely Michael A. Gilbert üzenet-aktus (*message act*) elméletére [8] támaszkodva teremt meg a kapcsolatot a nyelvi elemek és az érzelmkifejezések között. A üzenet-aktus elmélet lényege, hogy a kommunikáció nyelvi elemei érzelmi információt közvetítenek, amely az ismert beszédaktus (*speech act*) [1] elméletéhez hasonlóan több különböző komponenssel írható le (lokúciós, információs, illokúciós és perlokúciós aktusok). A két modell kiváló alap szintetikus érzelmek előállításához. A szintetikus jelleg miatt fontosnak hangsúlyozni, mert számos érzelmi modell létezik, továbbá emberenként is jelentős különbségek vannak, különösen a bonyolultabb érzelmek megjelenésében és működésében, ezért úgy gondoltuk, hogy egy egységes modell, mely ugyan számos egyszerűsítésen esett át, mindenképpen szükséges az érdemi felhasználáshoz. Ez azonban egyszerre jelenti azt is, hogy pontosan tisztában vagyunk azzal, hogy a GALA által felismert illetve generált érzelem nem feltétlenül feleltethető meg egy adott ember érzelmi állapotának, azonban tapasztalataink szerint statisztikailag egy jó általános, mesterséges érzelmi állapotváltozás modellezésére alkalmas. Elvégre valamennyien különbözőek vagyunk...

értelmezhető a felhasználó számára – természetesen abban az esetben, ha a gépi válasz adekvát.

(A diskurzusmenedzsment, a szemantikus információ kinyerése és követése stb. ismert és nehéz problémakör, melyet érintőlegesen tárgyaltunk csupán. Egy másik hasznos alkalmazási területre példa az e-learning

## 2 A BotCom beszélgetőrendszer felépítése

Beszélgetőrendszerünk többretegű felépítésű, a kezdeti szakaszban az üzenetet magát, illetve annak szemantikai tartalmát azonosítjuk és csak a legutolsó szakaszban kerül sor az érzelmi töltet meghatározására. Az első szinten szintaktikai feldolgozás történik: az egyes szavak azonosítjuk be szótövesítés [5] és egyéb morfológiai elemző eszközök alkalmazásával. Eztán kerül sor az üzenetnek a párbeszéd-szekvenciákból álló tudásbázisunkkal való összevetésére. Itt a bejövő üzenet részeinek mind a mondaton belüli grammatikai szerepét, mind az egyes szavak jelentés tartományába eső egyéb kifejezéseket figyelembe vesszük. Ez utóbbit az OSZK Tezauruszával [9], illetve egyéb szabadon felhasználható szinonima szótárak segítségével végezzük. Az üzenet szemantikai tartalmának pontosabb eldöntése a következő szinten valósul meg. Az azonosítást előre tárolt témakörök kulcsszavaival végezzük el. Annak érdekében, hogy pontosabban meg tudjuk határozni az adott dialógus szakasz fő témáját, figyelembe vesszük a dialógus korábbi szakaszában szereplő közlések már beazonosított szemantikai tartalmát is. Ez igen lényeges momentum, hiszen a dialógus során fenn kell tartani egyfajta perzisztenciát a témát illetően, annak érdekében, hogy a beszélgető robot se ugráljon túlságosan gyorsan más témára, akár félreértés, akár proaktív működés következtében. Így nagyobb biztonsággal tudunk a felhasználó üzenetéhez jól illeszkedő válaszokat adni. Az érzelmi feldolgozás csak ezután következik a később ismertett GALA modul segítségével.



### 2.1 A BotCom rendszer témakezelése

A BotCom válaszgenerálása igen nagymértékben függ a bejövő üzenetek helyes szemantikai beazonosításától. Ennek érdekében arra törekedtünk, hogy tudásbázisunk minél jobban lefedje a szokványos chat-en történő beszélgetés közben felmerülő témákat. A tudásbázist részben automatikusan, a weben szereplő chatlog fájlokból, részben 70 diák nyári-munka keretében történő, előre megadott témákról való beszélgetéseinek naplójából töltöttük fel. A feltöltésnél mind a témafelvetésekben, mind a válaszalternatívák variációjában próbáltunk megfelelő számú változatot felhalmozni. Előzetes kísérletek az egy kérdésre adott 4-5 sablon választ elégségesnek mutatták ahhoz, hogy egy fél órás beszélgetés kellően változatos legyen. A helyes szemantikai tartalom kinyerésének másik sarokköve a bejövő üzenet előzményeinek, a dialógus *history*-jának nyomon követése. A korábbi feldolgozási szakaszok nyújtotta alternatívák közötti választásnál fontos szempont a beszélgetés menetébe, mind a kommunikációs funkció, mind a szemantikai szempontból való

kációs funkció, mind a szemantikai szempontból való illeszkedés. A rendszer ily módon tulajdonképpen elvárásokat alakít ki a felhasználó üzenetét illetően és ezen elvárások mentén keresi a megfelelő választ.

Beszélgető rendszerünk így kevésbé „téved” el a fókuszukat nem tartalmazó közlések következtében. Például:

BotCom: Hogy vagy?

Felhasználó: Boldog vagyok.

B: Ezt jó hallani.

F: Nagyszerű érzés elmerengni a lét forgatagának végtelen lehetőségein.

B: Örülök, hogy jól érzed magad.

Ugyanez a dialógus a history figyelmen kívül hagyásával így zajlana le:

BotCom: Hogy vagy?

Felhasználó: Boldog vagyok.

B: Ezt jó hallani.

F: Nagyszerű érzés elmerengni a lét forgatagának végtelen lehetőségein.

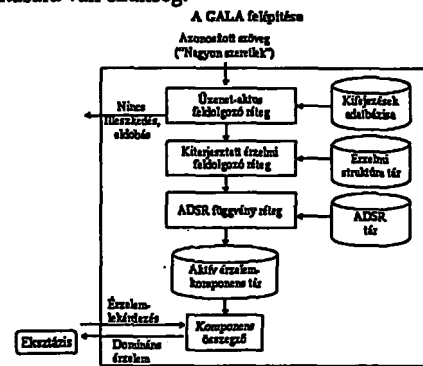
B: Milyen befektetési lehetőségeket ismersz?

A második példán látszik, hogy merengés, lé és a forgatás szavak együttállásának megfelelő szemantikai hálóval történő kiértékelése igen könnyen más síkra terelheti a beszélgetést. Míg ha a korábbi szekvenciák jelentésének tartalmát is figyelembe vesszük, a „nagyszerű érzés” sokkal jobban illeszkedik a korábbi szemantikus tartalomra, így érdemes megkockáztatni a mondat kevésbé illeszkedő részeinek akár a figyelmen kívül hagyását, és a „pozitív kedélyállapot” témában való válaszkeresés preferálását.

### 3 A GALA modul felépítése

A GALA érzelm szintetizáló és generáló rendszer a BotCom keretrendszerhez hasonlóan egy többlépcsős folyamat eredményeként nyeri ki a felhasználó üzeneteként azonosított mondatból a kommunikációs szándékot, illetve az érzelmi tartalmat. A modul önállóan is kezelhető és más keretrendszerekbe is könnyen integrálható, ahol írott szöveg érzelmi töltetének az azonosítására van szükség.

Az első rétegben az érzelmekkel és kommunikációs szereppel megfelelően felcímkézett adatbázisból kikeresünk az azonosított, így az adatbázisunkban már meglévő szöveghez tartozó üzenet-aktust. Ha ez nem szerepel az adatbázisban, akkor értelemszerűen a modul nem gyakorol hatást a válasz kiválasztására. Ha megtaláljuk az adatbázisban az üzenetet, vagy néhány szegmensét, akkor ezeket továbbítjuk az érzelm feldolgozó rétegnek. Ez vagy az egész szöveget, vagy az egyes szegmenseit megkeresi az Érzelmi



struktúra tárban, és kiértékeli a szöveg érzelmi töltetét. Ez a töltet, ahogy azt mindennapi beszélgetésekben is tapasztalhatjuk, nem feltétlenül csupán egyetlen alapérzelem, hanem több különböző erősségű alapérzelem együttes jelenlétével jellemezhető. Az érzelmek reprezentációját a Plutchik féle 24 alapérzelemmel operáló modell segítségével végeztük. A modellben több változtatást hajtottunk végre, illetve kiegészítettük egy minden érzelmet időbeli karakterisztikával jellemző függvénykezeléssel.

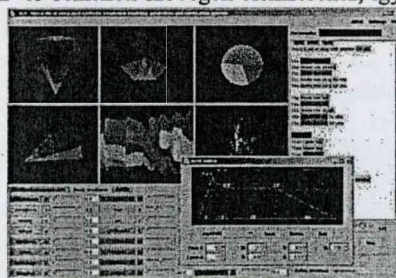
Az egyes érzelmekhez, illetve az egyes kifejezésekben hordozott összetett érzelmekhez (értsd több alapérzelem együttes jelenléte), különböző lefutású karakterisztikus függvényeket rendelünk. A függvények a könnyebb kezelhetőség kedvéért egyszerű ADSR (Attack Decay Sustain Release – Felfutás Csökkenés Kitartás Elengedés) egyenes szakaszokból állnak. Ezáltal konfigurálható, hirtelen hargú, de könnyen felejtő, vagy haragtartó, de nehezen feldühödő viselkedésű robot. A zenei hangok lecsengésének modellezésére is használt ADSR függvényekkel elérjük azt, hogy az egyes szöveg részletek által okozott hangulati hatások időben változnak, nem pillanatnyi a befolyásuk, és a korábbi üzenetek érzelmi töltete is hatást gyakorol a későbbi válaszok érzelmi töltetére. A komponens összegzés fázisához érve meghatározhatjuk egy adott pillanatban a korábbi, már lecsengőben lévő érzelmi hatások, és a legutolsó üzenet kiváltotta érzelmi hatások eredőjét. Ezt az eredő érzelmet leképezünk a már említett Plutchik-féle 24 alapérzelemmel operáló modellt, kinyerve a beszélgetés eddigi szakasza keltette domináns érzelmet. A robot a továbbiakban eszerint próbál viselkedni.

Mivel a BotCom rendszer már azonosította a bejövő üzenetet, és így kikereste az ehhez tartozó lehetséges válaszokat is, rendszerünknek csupán már az a dolga, hogy a válasz alternatívák közül a GALA által kiértékeltnek megfelelő kommunikációs tartalommal, illetve érzelmi töltettel rendelkezőt válassza ki az elérhetőek közül és közölje a felhasználóval. Egyezés hiánya esetén neutrális, vagy a legközelebbi érzelmi töltetű választ adja a rendszer.

#### 4 A GALA vizuális modellező felülete

A GALA modellező eszköz ennek megfelelően lehetőséget biztosít arra, hogy mondatokat, kifejezéseket, szavakat egy grafikus felületen keresztül érzelmi töltettel lássunk el. Továbbá tartalmazza az érzelmek dinamikájának kezeléséhez szükséges többretegű konfigurációs lehetőséget. A GALA használatával felcímkézett adatbázist a dialógusrendszer fel tudja használni, a ki- és bemeneti szövegek elemzéséhez, így azok érzelmi tartalmának megbecsléséhez.

A modellező eszköz képes grafikusán ábrázolni egy-egy beolvasott szöveg (pl. vers, párbeszéd) érzelmi töltetének időbeli alakulását, ezáltal segítséget nyújt a címkézés vagy párbeszéd során előálló hosszabb szövegek, dialógusok elemzéséhez és egyben a szintetikus érzelmi állapot változásának



realisztikussági fokának meghatározásához. Az első lépés kiértékelését 65 magyar nyelvű vers/dal és 70 félórás chat-dialógusból rögzített beszélgetés érzelmi címkézésével végeztük. Természetesen a címkézést végző személy értelmezésétől sosem lehet függetleníteni a rendszert, de előre kiadott irányelvek és egy személy által történő utólagos ellenőrzés elfogadható általános érzelmi jelölést eredményez.

A címkézés során bővülő érzelmi adatbázis egyre nagyobb mértékben teszi lehetővé ismeretlen szövegek automatikus felcímkézését, ugyanakkor ez még sok problémát hordoz magában. Hiába az alkalmazott több tízezer bejegyzést tartalmazó szótár, szinonimaszótár és szemantikus háló, az érzelmek címkézése elkerülhetetlen és időigényes, hiszen az egyes szinonimák számos esetben az általuk hordozott érzelmi töltetekben különböznek egymástól. E mellett sokszor rövid kifejezések címkézése szükséges ahhoz, hogy a valóságos érzelmi töltetet felfedjük, mely esetleg különbözik a kifejezés tagjai által önállóan hordozott érzelmi töltettől.

## 5 Összegzés

Ahogy ez várható volt, az egyszerűbb, egyértelműbb szavak, illetve világos, rövid mondatok esetén még viszonylag pontos találatokra képes a rendszer, ugyanakkor minél bonyolultabb az ismeretlen szöveg, annál gyakrabban ad kissé szűrreális érzelmi reakciókat. Jelenlegi kutatásaink ezért arra irányulnak, hogy a GALA kontextus-elemzőjének továbbfejlesztésével ezekben a komplikáltabb esetekben is jobb eredményt tudjon elérni.

## Irodalomjegyzék

1. Searle, J. R. What is a Speech Act. In *The Philosophy of Language*, J. R. Searle, Ed. Oxford University Press, London (1979)
2. Tatai G., Csordás A., Kiss Á., Laufer L., Szaló A.: Happy chatbot, happy user. *Proceedings of the 4th International Working Conference on Intelligent Virtual Agents (IVA'03)*, Irsee, Germany, Springer Verlag (2003).
3. Szaló A., Csordás A., Laufer L., Tatai G.: The GALA layered emotion model for advanced HCI interfaces. *Proceedings of the 3rd International Conference on Hybrid Intelligent Systems (HIS'03)*, Melbourne, Australia, IOS Press (2003).
4. Tatai G., Csordás A., Szaló A., Laufer L.: The chatbot feeling - Towards a usable emotional model for Internet ECAs. *Proceedings of EPIA'03 - 11th Portuguese Conference on AI*, Beja, Portugal, Springer Verlag (2003).
5. Magyar Ispell/Myspell Szótármodul: <http://www.szofi.hu/gnu/magyarispell/>
6. Plutchik R.: The nature of emotions. *American Scientist* 89(4):344-350 (2001).
7. Plutchik R.: A general psychoevolutionary theory of emotion. Plutchik, R., Kellerman, H. (szerk.): *Emotion theory, research, and experience, Theories of emotion*, pp 3-33, Academic Press (1980).
8. Gilbert M. A.: Language, words and expressive speech acts. van Eemeren, F., Grootendorst, R., Blair, J. A., Willard, C. A. (szerk.): *Proceedings of the Fourth International Conference of the International Society for the Study of Argumentation*, pp 231-234 (1999).
9. Országos Széchényi Könyvtár Tezaurusz: <http://www.oszk.hu/ujdonsag/tezauryj.html>