

Sistem de evaluare a participării unui student la un forum de discuții

Iulia Pașov, Ștefan Trăușan-Matu, Traian Rebedea

Universitatea Politehnica București

Splaiul Independenței nr. 313, București

iulia.pasov@cs.pub.ro stefan.trausan@cs.pub.ro traian.rebedea@cs.pub.ro

REZUMAT

Scopul acestui studiu a fost acela de a oferi o unealtă utilizată în comunicarea mediată de computer (CMC) precum și în sprijinirea experiențelor educaționale pe calculator. Acest studiu se bazează pe un model numit comunitate de interes (“*community of inquiry*”) ce constă în trei elemente esențiale ale unui proces educațional: prezența cognitivă, socială și educativă. Indicatori (cuvinte, expresii) pentru fiecare dintre cele trei elemente se obțin în urma analizei textelor. Acești indicatori reprezintă șabloane pe care cercetătorii le folosesc pentru a analiza niște comunicări scrise, dar și pentru a ghida personalul didactic spre a utiliza conferințele computerizate cât mai eficient, ca pe un mediu ce facilitează dezvoltarea educației. Prin acest studiu se demonstrează că orice conferință computerizată are potențialul de a crea o comunitate cu scopuri educaționale.

Cuvinte cheie

Prelucrarea Limbajului Natural, Inteligență Artificială, Forum, Comunicarea Mediată de Calculator, CSCL

Clasificare ACM

I.2.7 Natural Language Processing, K.3.1 Collaborative Learning

INTRODUCERE

Unul dintre principalele motive pentru care universitățile optează pentru învățământul la distanță este democratizarea accesului la educație al oamenilor ce provin din medii dezavantajate. Când s-a introdus *e-learning*, sau e-învățarea (e-educația), același motiv a justificat procesul. Un atribut important al tehnologiilor folosite este capacitatea lor de interacțiune în timp real.

Mult timp s-a crezut că procesul de e-educație constă în transpunerea în format digital a unui curs predat în clasă, ce poate fi apoi trimis unui număr oricât de mare de elevi, fără prea multă implicare din partea profesorilor. Această perspectivă nu se apropie deloc de adevăr, deoarece nu corespunde sensului educației. Simpla publicare a cursurilor nu înlocuiește procesul de educație, ci doar face parte din acesta, fiind completată de interesul și dedicația profesorului, precum și de colaborarea între acesta și elevi sau doar între elevi.

Comunicarea mediată de calculator (CMC) se remarcă între metodele de educație la distanță din cauza capacității de a suporta nivele înalte de interacțiune între facultate și studenți, în același timp oferind libertate în timp și spațiu

pentru implicare [10]. Indiferent de semnalele folosite, informația trebuie transmisă interactiv între sursă și destinație, între student și profesor, studenți și studenți sau către studenți [14].

Comunicarea mediată de calculator este definită ca orice formă de comunicare interpersonală ce se produce cu ajutorul unei rețele formate din două sau mai multe calculatoare. Cercetările în CMC se axează în special pe efectele sociale pe care le au diferite tehnologii de comunicare cu ajutorul calculatoarelor. Multe studii recente iau în considerare faptul că rețelele sociale bazate pe Internet sunt în strânsă legătură cu softurile sociale.

Multe instituții de învățământ superior folosesc sau intenționează să folosească CMC, în special conferințele în mod text, considerându-le un mediu capabil pentru distribuția programelor educaționale, oriunde, oricând. În timp ce dezvoltatorii acestui nou mediu sunt convinși de potențialul său, nici efectele asupra calității procesului de învățare, nici rezultatele sale, nu au fost încă bine studiate.

Cum CMC au început să fie folosite în scopuri cât mai diferite și inovative în domeniul educației, și nu numai, a apărut și nevoia prelucrării acestora, dar în timp ce numărul instrumentelor de colaborare pe Web, precum și utilizarea acestora, crește tot mai mult în fiecare an, ne confruntăm cu problema analizei acestora și extragerii informațiilor din ele. În acest scop un rol important îl are prelucrarea limbajului natural.

SCOPUL LUCRĂRII

Conferințele computerizate au devenit tot mai utilizate în ultimii ani, mai ales datorită posibilității de a simula conversațiile “față în față” într-un spațiu virtual. Conform studiilor, printre beneficiile conferințelor computerizate ar fi: confortul, independența spațială și temporală, posibilitatea utilizatorilor de a face parte dintr-o comunitate virtuală [2].

Dacă prin învățarea colaborativă studentul este cel mai avantajat participant, pentru profesor predarea online necesită cel puțin același efort ca și predarea într-o sală de clasă. Acesta trebuie să pregătească materialele, să le facă disponibile pe calculator, să-i motiveze și să-i ghideze pe studenți printr-o interacțiune continuă, la distanță. Deși atât studenților, cât și profesorilor le este permis să participe la cursuri și să lucreze din orice locație, și în intervale mari de timp, efortul depus de profesor pentru fiecare elev în parte crește semnificativ, mai ales când apare problema evaluării. Profesorul ajunge în situația de a evalua participarea fiecărui elev într-o comunicare online,

ce poate ajunge la dimensiuni considerabile. Simpla parcurgere a contribuției fiecărui participant necesită foarte mult timp și concentrare din partea profesorului.

Această lucrare urmărește dezvoltarea unei aplicații care să ajute profesorul să evalueze participarea unui student la un forum de discuții. Printre avantajele acestei aplicații de precizat ar fi în special reducerea timpului și creșterea eficienței.

Aplicația folosește modelul comunității de interes (*Community of Inquiry*), pentru a identifica în fiecare mesaj dintr-un forum prezențele sociale, educative sau cognitive ale autorului. Prin ea se dorește să se încurajeze utilizarea CMC în scopul realizării obiectivelor educaționale putând fi folosită pentru a analiza diverse strategii educaționale și efectul acestora asupra elevilor și studenților.

PROIECTE SIMILARE

Datorită necesității de a extrage informații din diverse comunicări verbale sau textuale, au apărut mai multe proiecte cu scopuri similare, majoritatea dorind să detecteze prezența emoțiilor. Mai multe companii ca *Right Now Technologies* sau *NICE Systems* au produs aplicații precum *SmartSense™* sau *NICE Perform™* care recunosc emoțiile folosind tehnici de detecție a cuvintelor cheie, apoi marchează, prioritizează și dirijează investigațiile, bazându-se pe conținutul emoțional.

Tot în domeniul detecției emoțiilor se remarcă și agentul care preia știri de la diverse resurse, clasificând titlurile în opt categorii de emoții, folosind interpretoare semantice și *SenseNet* [12].

Un alt studiu [8] folosește un program de învățare automată, numit *BOOSTEXTER* să clasifice automat emoții. *BOOSTEXTER* este un program destinat clasificării de texte. Acesta folosește un algoritm de stimulare, ce formează o clasificare ipotetică combinând rezultatele mai multor iterații [11].

Un grup de cercetare de la Universitatea din Illinois a dezvoltat un soft numit *Semantic Role Labelling* [15] ce identifică verbul și structura argumentului într-o propoziție. Acesta a fost utilizat, alături de un motor de căutare (Google), în continuare cu scopul detecției emoțiilor, încercând în același timp o înțelegere a textului [16]. Din păcate Google nu furnizează mereu informațiile așteptate (exemplu: *war + happy* are mai multe rezultate decât *war + sad*) și de multe ori rezultatele nu sunt cele corecte.

E-rater este un program pentru notarea automată a eseurilor, dezvoltat de ETS (*Educational Testing Service*) și inaugurat în 1999, ajungând în prezent la versiunea v.2. Inițial a fost folosit pentru notarea testelor GMAT (*Graduate Management Admission Test*), necesare pentru admiterea la programele de masterat. Programul analizează exemple de scris, sintaxă, topologie, discurs și vocabular. Eseurile ce trebuie notate sunt comparate cu un set de învățare ce conține sute de eseuri scrise corect, fiecare set corespunzând unei întrebări. Cum programul se bazează pe un set de test, el nu poate să înțeleagă sau să decidă dacă textul este coerent, sau dacă ceea ce este scris are sens. Deși versiunea v.2 este cu mult superioară celei

din 1999 și rezultatele sunt foarte bune, programul nu face conexiuni logice și nu aduce argumente.

MODELUL "COMUNITY OF INQUIRY"

Interacțiunea este considerată esențială în învățarea experiențelor din perspectiva constructivistă socioculturală [13], ceea ce presupune că membrii unei comunități ce participă într-o comunicare își dezvoltă mai mult cunoștințele. În plus, capacitatea de dezvoltare intelectuală a elevului (a celui care învață) se consideră a fi sporită de prezența unei ghidări, reprezentate de un tutore/coleg, prin interacțiune. În contextul educației online, sau e-educației, schimbarea mediului de la educația tradițională, în sala de clasă, la una plasată într-un spațiu virtual, prezintă numeroase beneficii și provocări pentru pedagogi, întrucât interacțiunile nu mai sunt limitate de decizii tehnologice.

În 1989 Moore a introdus trei tipuri de interacțiune, acum cunoscute la scară largă și acceptate în domeniul învățământului la distanță: cursant – conținut, cursant – instructor și cursant – cursant.

Datorită optimizărilor tehnologiilor web/CMC, interacțiunea cursant – cursant reprezintă o nouă dimensiune pentru învățământul la distanță, ce ocolea de obicei grupurile/abordarea învățării colaborative datorită separării geografice a participanților.

Dezvoltând aceste trei clase de interacțiune, Anderson și Garrison [17] au sugerat că orice comunicare online ar trebui să se desfășoare între cele trei macro-componente: cursantul, instructorul și conținutul, creând astfel tipurile: interacțiune cursant – conținut, cursant – instructor și instructor – conținut. În plus, comunicarea din interiorul fiecărei macro-componente poate determina un subset de interacțiuni: cursant – cursant, instructor – instructor și conținut – conținut. Pentru a obține un context pentru aceste interacțiuni, Garrison, Anderson și Archer [17] au creat modelul *Community of Inquiry*, care a fost elaborat și explicat în publicațiile ulterioare: Anderson, 2004; Garrison, 2003; Garrison și Anderson, 2003; Kanuka și Garrison, 2004 etc.

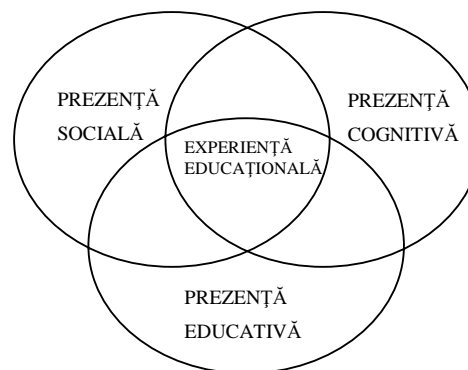


Figura 1. Mediul de comunicare într-o comunitate de interes [4]

Precum se observă în Figura 1, o experiență utilă de învățare este cuprinsă într-o comunitate de interes compusă din profesori și studenți, participanții esențiali într-un proces educațional. Modelul presupune că învățarea se realizează în interiorul comunității prin interacțiunea celor trei elemente. În aceeași figură sunt

prezentate cele trei elemente esențiale: prezența socială, cognitivă și educativă.

În analiza conferințelor folosite pentru scopuri educative, se caută mesaje sau segmente din acestea, care arată că aceste trei elemente esențiale sunt prezente, adică se caută indicatori ale celor trei elemente. Acești indicatori sunt reprezentați de anumite fraze, cuvinte cheie, sau sinonime ale acestora. Pentru o structurare mai ușoară acești indicatori au fost grupați în subcategorii care indică mai clar aspectul fiecărui element ce este reprezentat de fiecare grup de indicatori [5].

Prezența Cognitivă este măsura prin care elevii sunt capabili de a construi și confirma sensul, susținând reflecția și discursul [4].

Tabelul 1. Indicatori prezență cognitivă

Definirea Categoriilor	Cod	Indicatori
Declanșare Eveniment (<i>Triggering Event</i>) conceptualiza rea inițială a unei probleme	cte	Recunoașterea unei probleme, probabil din experiență, exprimarea nedumeririi sau disconfortului, întrebări, cererea explicațiilor e.g. <i>"I am not quite sure what [...] is about – should we start by clarifying what we mean by[...]."</i>
Explorare (<i>Exploration</i>) încercarea de a găsi sensul unei probleme	ce	Schimb de informații, clarificări de situații sau termeni, discutarea ambiguităților, căutarea explicațiilor; caracterizat de schimbarea informațiilor; e.g. <i>"I think this is an interesting topic, but others have a different perception of it, to mine and we have to take this into account"</i>
Integrare (<i>Integration</i>) conectarea ideilor și începerea de a lega concepte, ce duce la găsirea de explicații	ci	Integrarea cunoștințelor și ideilor în explicații coerente; testarea posibilelor intuiții asupra problemelor eg <i>"I observed all types of behaviors on the hospital wards, but you have to reflect on the reasons why people behave as they do in order to [...]and the easiest course of action is not always the best"</i> .
Rezoluție (<i>Resolution</i>) estimarea soluțiilor problemelor (Prioritate maximă)	cr	Reflectarea asupra eficienței soluțiilor sau dilemelor, explorarea consensurilor, a înțelegerilor sau a diferențelor eg <i>"Perhaps the best way of approaching this problem is not by [...], but by asking questions in such a manner [...], which would improve inappropriate practices by [...] by making them think about their actions"</i> .

În cadrul prezențelor cognitive, deși fiecare subcategorie are o prioritate (cea mai mica pentru declanșarea

evenimentelor si cea mai mare pentru rezoluție), uneori mai multe apariții ale unei categorii cu prioritate mai mică pot să depășească în importanță o categorie cu prioritate mai mare, dacă acest lucru este mai relevant (Dacă cr = 6, ci = 4, ce = 3, cte = 2 și o propoziție poate primi adnotarea cr sau 3*ce, ar fi de preferat să alegem ce).

Prezențele sociale și educative nu includ o ierarhie, așa că putem avea oricâte instanțe ale fiecare subcategorii.

Prezența Socială este capacitatea participanților de a se uni și identifica cu comunitatea, de a-și comunica intențiile într-un mediu de încredere și de a dezvolta relații interpersonale printr-o metodă prin care își expun propriile personalități unice [5].

Tabelul 2. Indicatori prezență socială

Definirea Categoriilor	Cod	Indicatori
Expresie emoțională (<i>Emotional Expression</i>) indică sentimentul de siguranță într-un mediu online	see	Împărtășirea expresiilor și a trăirilor, atât exprimarea convențională, cât și neconvențională a emoțiilor, a umorului, ironiei și deschiderea spre auto-dezvăluire ce indică vulnerabilitate. e.g. <i>"I was so angry[...] I could not understand him[...]"</i> .
Comunicare Deschisă (<i>Open Communication</i>)	soc	Recunoașterea altor participanți și a contribuției acestora, încurajarea participanților în legătură cu mesajele lor eg <i>"In your last message you referred to[...]."</i> I really liked your interpretation of that situation"
Colaborare în Grup (<i>Group Collaboration</i>)	sgc	Încurajarea interacțiunii în grup, acceptând diferențe de opinie, indicate de adresarea cu "we"(noi), referirea la participanți cu ajutorul numelor,, folosind "our" eg <i>"I think that John summarised our discussions very well.."</i>

Prezența Educativă este proiectarea, facilitarea și direcția proceselor cognitive și sociale având ca scop realizarea rezultatelor învățării într-un mod semnificativ din punct de vedere personal și al meritelor educaționale [1].

O problemă a aplicației este greutatea identificării prezențelor educative, ale profesorului. O soluție simplă ar fi fost alegerea persoanelor cu rol de tutore, sau identificarea lor cu prima persoană care postează, dar ar fi putut apărea neconcordanțe in cazurile cu mai mulți îndrumători.

Tabelul 3. Indicatori prezență educativă

Definirea Categoriilor	Cod	Indicatori
Managementul Instrucțiunilor (<i>Instructional Management</i>) crearea de structuri de bază	tim	Facilitează dezvoltarea organizării în cadrul grupului, inițiind discuții, setând reguli sau etichete eg “ <i>In our inițial face to face meeting we decided to deal with....</i> ” “ <i>We must finish this discussion by Friday...</i> ”
Instruire Directă (<i>Direct Instruction</i>) ține pasul cu discuția, confirmă că grupul a înțeles, răspunde la probleme tehnice	tdi	Înțelege când grupul a ajuns la un punct critic și îl ajută să îl depășească, folosind cunoștințe din afară și alte referințe pentru a menține discuția aprinsă, răspunzând la probleme tehnice eg. “ <i>If you want to upload an attachment just click on....</i> ”
Construirea Discursurilor și a înțelegerilor în grup (<i>Group Cohesion</i>)	tbu	Facilitează colaborarea în grup, identificând acordul sau dezacordul, asigurând un mediu pentru discuție, făcând rezumate, folosind întrebări-cheie pentru a dezvolta discuția, confirmând comentariile interesante de la alți participanți, încurajând pe toți să participe eg “ <i>Any thoughts on this issue?</i> ” “ <i>Anyone care to comment?</i> ” “ <i>I think we’re getting a little off track.</i> ”

Potrivit lucrării lui Garrison și Anderson [4], termenul “Community of inquiry” a fost folosit pentru prima dată de Lipman [7] referindu-se la o comunitate facilitată de profesori, unde studenții își ascultă colegii cu respect, rezultat din exprimarea opiniilor personale, se provoacă unii pe alții pentru a obține motive pentru opinii contrare, care în alt context nu ar fi fost acceptate, se ajută reciproc în obținerea concluziilor la ceea ce a fost spus, și caută să-și identifice și clarifice reciproc ipotezele [7]. Bazându-se pe acest concept larg de comunitate educațională, care cuprinde studenți și profesori, și pe ipoteza constructivistă că formarea de cunoștințe este un proces social contextualizat, care se desfășoară într-o asemenea comunitate, modelul *Community of inquiry* este conceput astfel încât să cuprindă trei elemente ce interacționează între ele și se consolidează: prezența social (*social presence*), prezența cognitivă (*cognitive presence*) și prezența educativă (*teaching presence*). Acestea sunt dezvoltate în mediile de instruire online prin tehnologii CMC (mediate de calculator). Formarea unei astfel de comunități în contextul învățării online reprezintă un mediu pentru “discurs critic și reflecție” [5], unde experiența educației bazate pe dialog, de a împărtăși și

negocia atât ceea ce este cât și ceea ce nu este înțeles, poate conduce la “niveluri mai ridicate de învățare” [10].

APLICAȚIA

Aplicația realizată are scopul de a analiza discuții online, și de a evalua participarea studenților la acestea. Inițial programul a fost aplicat pe forumurile de discuții ale studenților de la Facultatea de Medicină din Manchester. În urma studiului asupra acestor forumuri s-a ajuns la concluzia că modelele descrise anterior pot fi identificate în funcție de anumite șabloane.

Pentru a implementa aplicația au trebuit considerate mai multe elemente: interfața grafică, care ar trebui să fie cât mai simplă, intuitivă și ușor de utilizat, și modulele apelate de aceasta cu rolul detecției celor trei tipuri: prezență socială, cognitivă și educativă, din modelul comunității de interes (*Community of Inquiry*), conectate prin structura *Forum*.

Programul se bazează pe un sistem de reguli prin care se identifică într-o serie de mesaje din forumuri prezențele sociale, educative și cognitive, căutând șabloane fixe în text, pentru a putea evalua participarea studenților la discuție. Studenții care prezintă toate tipurile de prezențe: socială, cognitivă și educativă sunt cei care se implică mai mult în conversație. Cele trei tipuri diferite de prezențe fac posibilă o analiză mai complexă a mesajelor, pentru a nu considera cele mai implicate acele persoane care discută mult pe teme irelevante subiectului de tratat. Acestor persoane sunt favorizate acele care se implică în conversație din toate aspectele (social, cognitiv și educativ).

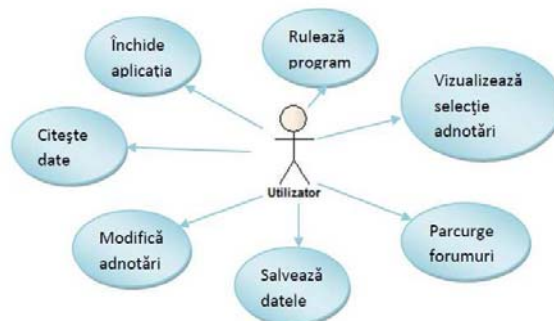


Figura 2. Use-case

Pentru a rula aplicația, este nevoie doar de o mașină virtuală Java. Comunicarea cu interfața este cât se poate de intuitivă:

File->Open: citește datele dintr-un XML ce conține diverse mesaje în forumuri și afișează informațiile într-un tabel în primul tab, numit sugestiv, *Forums*.

File->Save: opțiunea este dezactivată în lipsa modificărilor asupra tabelului, iar când este activă salvează noul tabel.

File->Exit: iese din aplicație

Edit->Run: adaugă în ultima coloană adnotările calculate

Dintre coloanele tabelului din primul tab, se poate modifica doar cea cu adnotările citite din XML, putându-se seta oricare dintre adnotări. Al doilea tab permite afișarea pentru un anumit mesaj. Selectând un anume

mesaj din tabel, prin click pe ID, în al doilea tab se poate vizualiza textul și adnotările calculate. Prin selectarea unei opțiuni din lista de RadioButtons (unde putem alege o subcategorie din modelul Community of Inquiry) și apăsarea butonului “Ok”, propozițiile în care s-a identificat această categorie sunt evidențiate cu font de culoare roșie.

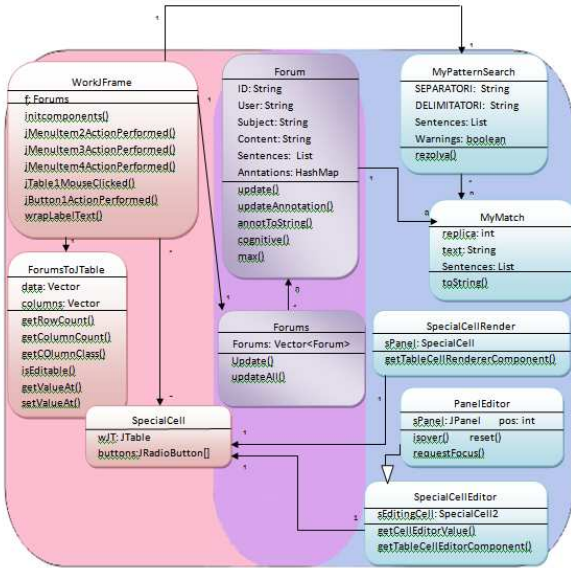


Figura 3. Diagrama claselor

Într-o primă instanță, utilizatorul nu poate decât să închidă aplicația sau să citească forumuri. Open declanșează citirea din XML, prin constructorul Forums() care populează lista de forumuri.

Forum este clasa principală a aplicației, ce reprezintă un mesaj pe forum. Forums este o colecție de obiecte de tip Forum. Fiecare astfel de Forum are un număr de identificare unic, un autor, subiect, text, adnotări preluate din XML și adnotări calculate. Textul este împărțit într-o lista de propoziții: sentences. Rezultatele obținute în urma adnotărilor automate se păstrează într-un HashMap, annotations, ce asociază fiecărei categorii un Vector de MyMatch-uri.

Căutarea propriu-zisă este realizată cu ajutorul clasei PatternSearch, apelând metoda rezolva() care întoarce un vector de potriviri (Match). O potrivire conține indice, expresia căutată și textul în care a fost găsită expresia. Clasele sunt destinate lucrului cu chat-uri.

Aceste clase au fost modificate pentru a corespunde cerințelor, și se aplică pe propoziții, întorcând o singură subcategorie de un anumit tip pentru o propoziție dată, întrucât am considerat că nu putem avea mai multe instanțe. Noua clasă MyMatch va corespunde potrivirii făcute între o expresie și o propoziție. Vectorii de MyMatch-uri întorși de metoda rezolva() sunt parcurși pentru a elimina instanțele duble și a fi memorati în HashMap-ul fiecărui forum în parte.

Interfața cu utilizatorul este simplă, intuitivă, pentru a putea fi folosită de oricine. Aceasta conține un meniu, cu câteva sub-meniuri și un TabbedPane cu două taburi.

Tabelul (jTable1) are coloane de tip nr întreg (ID), text (Utilizator, Subiect, Conținut) sau panel. A cincea coloană, cea cu adnotările citite din XML conține un tabel cu 3 linii și 3 coloane, conținând categoriile din prezențele sociale și educative, și numărul aparițiilor, și un panel cu RadioButtons – categoriile de prezență cognitivă.

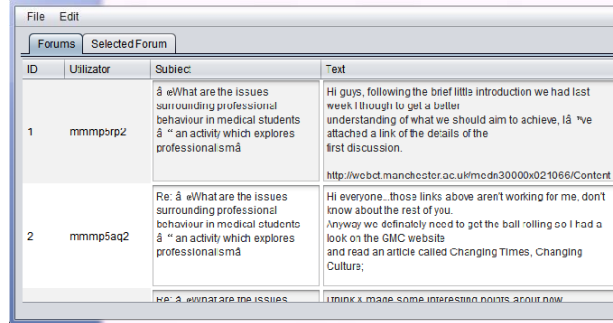


Figura 4. Citirea datelor

Utilizatorul poate modifica adnotările citite din fișier, selectand pe rand mesaje si apoi categoriile dorite.

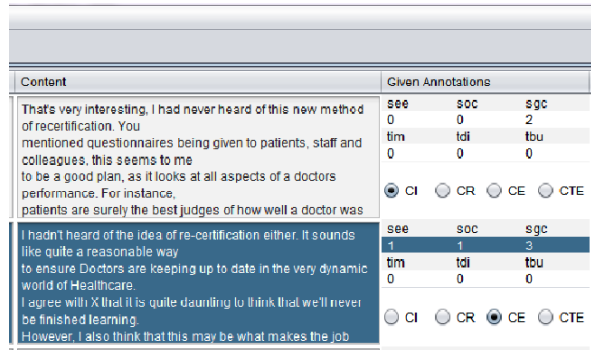


Figura 5. Selecția

Al doilea tab conține un label, unde va fi afișat textul forumului, o serie de RadioButtons și un TextArea, ce va conține adnotările. Selectând o categorie de prezență socială, educativă sau cognitivă, la apăsarea butonului Ok, în label propoziția în care aceasta a fost identificată se va colora cu font roșu. Acest lucru este posibil prin setarea textului corespunzător label-ului în format HTML. Tot aici a fost definită o metodă de extindere (wrapLabelText) a textului pe label, pentru a putea afișa mai multe rânduri de lățime fixă.

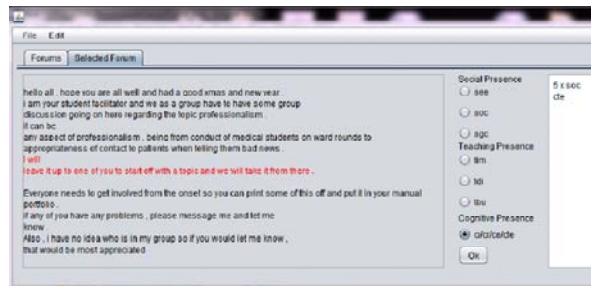


Figura 6. Evidențiere categorii selectate



Figura 7. Evidențierea prezenței cognitive

Pentru a determina categoriile din modelul lui Garisson, corespunzătoare fiecărui mesaj al forumului am ales să adnotez fiecare propoziție din textul postat pe forum. Astfel, o propoziție poate să fie marcată cu o singură subcategorie a prezențelor sociale, educative și cognitive. În final, dintre prezențele cognitive identificate pe un mesaj se alege una singură, în funcție de prioritate și cantitate.

Pentru a lucra mai bine cu propoziții și cuvinte, mai ales datorită faptului că aplicația se bazează pe recunoaștere de șabloane, am apelat la *WordNet*.

WordNet [9] este o bază de date lexicală interactivă, o ontologie lexicală, electronică, considerată în momentul de față cea mai importantă resursă pusă la dispoziția cercetătorilor în lingvistica computațională, analiza de texte și alte domenii similare. Aceasta a fost dezvoltată de un grup de cercetători de la Universitatea Princeton, conduși de George Miller [3]. *WordNet* poate fi privit ca un dicționar semantic deoarece cuvintele sunt localizate pe baza afinităților conceptuale cu alte cuvinte, spre deosebire de cazul dicționarelor, unde găsim cuvintele în ordine alfabetică. În această bază de date se pot găsi majoritatea cuvintelor limbii engleze (substantive, adjective, verbe și adverbe), organizate în mulțimi de sinonime numite *synset-uri*. Fiecare dintre aceste *synseturi* reprezintă un concept, ceea ce ne duce la concluzia că *WordNet* organizează informația lexicală în termeni de sensuri ale cuvintelor. Cuvintele polisemantice aparțin mai multor *synseturi*. *WordNet* face legătura între formele cuvintelor și sensurile lor, folosindu-se de categoria lexico-gramaticală (clasa morfologică) pentru a grupa termeni cu același interes în același *synset*.

În această aplicație *WordNet* este folosit în detalierea și identificarea șabloanelor din setul de reguli, pentru a face posibilă o căutare evoluată.

REZULTATE

Aplicația a avut rezultate foarte bune în special pentru prezențele sociale și cognitive, probleme apărând mai ales la prezențele educative, datorită șabloanelor alese, care sunt destul de ambigue. Rezultatele se pot vedea în tabelul de mai jos. Acesta a fost realizat în urma analizei a 400 de mesaje postate pe forumuri.

Pentru centralizarea rezultatelor am considerat că un forum a fost adnotat corect dacă prezintă aceleași subtipuri de categorii (see, soc, sgc, tim, tdi, tbu, si, ce, cr, cte) cu datele citite din XML. Diferențe apar de obicei datorită faptului că unele persoane au adnotat paragrafe și nu propoziții sau au considerat un singur item de o anumită

categorie pentru două, sau mai multe, propoziții consecutive de același tip.

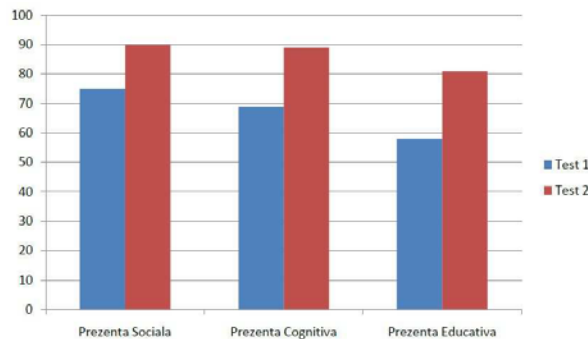


Figura 8. Rezultate eficiență

Al doilea studiu a fost efectuat pe 40 de mesaje adnotate după teoria de la 3.2, cu specificația că fiecare propoziție poate fi descrisă de aceste categorii. Rezultatele au fost considerate corecte pentru erori mai mici de 10%.

De remarcat este faptul că analiza este independentă de conținut: indiferent de subiectul conversației, performanțele rămân constante.

În urma rulării programului pentru aproximativ 400 de mesaje din forumuri și cu o listă considerabilă de șabloane de identificare, obținute în urma analizei textelor, s-a ajuns la concluzia că acesta necesită destul de mult timp pentru a adnota textele, în ciuda eforturilor depuse pentru a ameliora timpii de execuție. Avantajul este că, spre deosebire de oameni, programul nu obosește și performanțele sale sunt constante în timp.

A fost realizat un studiu pe cinci persoane (studenți la Facultatea de Medicină) pentru a compara performanțele programului. Fiecare a trebuit să adnoteze câte 40 de mesaje din forumuri, având la dispoziție descrierile de la capitolul 3.2. S-a ajuns la concluzia că timpul necesar unei persoane de a face această analiză este cu mult mai mare, deoarece pentru numai 40 de mesaje s-a obținut un timp de aproximativ 10 de ori mai mare decât cel necesar programului pentru un număr de zece ori mai mare de mesaje. În plus, după primele 10-15 mesaje toți participanții la studiu au cerut o pauză motivând probleme de vedere și oboseală. La cei care au adnotat aproximativ 30 de forumuri fără pauză au apărut multe erori în prelucrarea datelor, categoriile alese fiind de multe ori inexistente în textele respective.

În plus, este suficient ca programul să fie rulat o singură dată pentru a determina categoriile pentru fiecare forum, după care evaluarea se face foarte ușor apelând la rezultatele păstrate în obiectele de tip Forum.

OBSERVAȚII

De remarcat și notat este faptul că două persoane diferite nu pot adnota în același mod același text, deoarece oamenii sunt foarte subiecțivi. De multe ori adnotările citite din XML nu s-au potrivit analizei efectuate de cele cinci persoane testate pentru subcapitolul anterior. Această subiecțivitate este rezultată din faptul că un om citește textul și alege categoriile în funcție de înțeles. Din moment ce încă nu s-a dezvoltat un algoritm de inteligență artificială care să poată înțelege un text, acest lucru nu este

posibil de făcut în cadrul unei aplicații. Metoda aleasă (sistem bazat pe reguli) întoarce rezultate foarte bune pentru scopul analizei.

Se poate observa că persoanele în ale căror intervenții pe forum se pot identifica toate cele trei categorii: prezența socială, prezența cognitivă și prezența educativă, indiferent de numărul acestora, sunt acelea cu o activitate cât mai productivă. În cadrul unei evaluări a participării fiecărui participant la forum aceste persoane vor avea punctaje cât mai mari. În zona de dezvoltare comună a celor trei elemente independente se creează o experiență de învățare profundă. Deși elementul cheie ce stă la baza procesului de învățare este cel cognitiv, prezența celorlalte denotă faptul că studentul respectiv se implică în discuție, și are șanse mari să se dezvolte mai mult, deoarece forumurile sunt folosite în scopul învățării colaborative.

CONCLUZII

CSCL dorește să ajungă la o *construire colaborativă a cunoașterii* [15] prin susținerea și ajutarea tuturor elevilor sau studenților să învețe eficient împreună. Se încurajează ideea comunicării ideilor și informațiilor între studenți, a accesării colaborative a informațiilor și documentelor, și a feedbackului constant de la colegi și tutore. CSCL sprijină activitățile în grup și dinamica grupului, în moduri ce nu pot fi aplicate în comunicarea față în față.

Cercetătorii au demonstrat că învățarea colaborativă are un rol foarte important în educație și în dezvoltarea personală a participanților. Dacă învățarea colaborativă reprezintă ideea de a forma grupuri care să lucreze și să învețe împreună, într-un mod colaborativ, atunci uneltele folosite de învățarea colaborativă sprijinită de calculator își îndeplinesc sarcina fie sincron, fie asincron. Astfel chaturile, forumurile, blogurile nu sunt folosite doar pentru relaxare și distracție, ci și în scopuri educative. Spre deosebire de sala de clasă, mediul virtual este de multe ori privit cu mai multă încredere și simpatie. Totuși, CSCL înseamnă *Computer-Supported Collaborative Learning*, iar procesul de educație (“learning”) nu presupune doar avantaje pentru elevi sau studenți, întrucât la fel de implicat în acest proces este și profesorul. CSCL ar trebui să aducă avantaje pentru toți participanții la procesul de educație. În concluzie, și profesorul ar trebui să aibă parte de o ușurare sau evoluare a muncii, datorată folosirii calculatoarelor (“Computer-Supported”) și în ajutorul lui vin aplicațiile de inteligență artificială.

Rolurile tutorelui în CSCL, precum și responsabilitățile sale cresc foarte mult, deoarece el nu se mai poate exprima la fel de ușor, și nu poate să apeleze la gesturi, mimică sau tonalitate. El va trebui să aibă grijă să inspire încredere pentru a insufla cursanților dorința de a învăța pe cont propriu, să animeze grupul de care este responsabil, încitând toți membrii să realizeze raporturi virtuale, să motiveze elevii în a găsi cele mai potrivite strategii și metode de învățare, să identifice și să rezolve neînțelegerile sau problemele lor, să le răspundă la nelămuriri și să livreze feedback legat de progresele realizate etc.

În cadrul unui sistem bazat pe comunicare este bine de știut cât de mult se implică fiecare participant în colaborare. Am prezentat în capitolele precedente ce înseamnă învățare colaborativă, ce avantaje aduce aceasta și ce rezultate au fost înregistrate. Adevărul este că, într-un sistem colaborativ în care colaborarea este la un nivel minim, niciunul dintre aceste rezultate nu poate fi atins. De aceea este foarte important să analizăm gradul de participare al fiecărui membru la discuție.

Lucrarea de față își propune nu doar să găsească alternative teoretice de rezolvare, ci și să aducă o implementare a acestora, care să vină în sprijinul profesorilor sau cercetătorilor. Aplicația și-a atins scopul, putând să identifice prezențe sociale, cognitive sau educative în diverse mesaje de pe forumuri, făcând posibilă analiza contribuției fiecărui participant.

Aplicația poate fi îmbunătățită din mai multe puncte de vedere. Interfața ar putea să cuprindă mai multe modalități de expunere a rezultatelor (de exemplu în forma diagramei de la capitolul 3.2). Modelul șabloanelor poate fi optimizat dacă fiecare regulă ar primi o prioritate. Acest lucru ar avantaja algoritmul deoarece uneori expresiile nu sunt suficiente pentru a se potrivi cu modelul corect. Trebuie luate în calcul optimizările care conduc spre un mai bun echilibru între precizie și timp. În acest scop se dorește și evaluarea altor clasificatori, și includerea mai multor facilități, în special semantice, precum și extinderea utilizării metodelor de învățare automată până la un echilibru precizie-timp.

Tot în scopul de a optimiza modelul de analiză cu ajutorul regulilor, se poate recurge la combinarea cu un algoritmul de învățare automată pentru a filtra expresiile incorecte, dar și pe cele corecte. O abordare bazată pe reguli este utilă pentru a identifica expresii fixe, dar dependentă de existența unui expert care să definească setul de reguli. Acestea se obțin ca rezultat a sute de ore de muncă (analiză de text). Analiza gramaticală actuală este suficient de avansată pentru a identifica mai multe tipuri diferite de șabloane care să se muleze perfect pe o conversație obișnuită, dar limbajul natural nu este predictibil. Comunicarea în limbaj natural se remarcă în special prin variații și incertitudini, apariții ale expresiilor noi, fiecare domeniu având propriile expresii utilizate pentru a descrie o anumită activitate, obiect, etc. Includerea tuturor expresiilor în modelul bazat pe reguli nu numai că ar fi imposibil de realizat, dar și dacă ar fi îndeplinită, ar pune problema unei complexități temporale foarte ridicate, în special datorată expresiilor inutile (Nu are sens să căutam șabloane specifice unui domeniu în analiza forumurilor dintr-un domeniu independent de acesta). Într-adevăr aceste câteva expresii, per domeniu, ar fi într-un număr mult mai mic decât al celor generale (pentru orice grup), dar ținând cont că pot exista extrem de multe grupuri, fiecare cu un limbaj propriu, se poate ajunge chiar la ambiguitate. Deși completarea algoritmului cu o analiză de învățare automată ar presupune încetinirea calculelor, aceasta ar fi de preferat completării listei de reguli cu toate posibilitățile, și ar conduce la rezultate mai bune.

REFERINŢE

1. Anderson, T., Rourke, L., Garrison, D. R., & Archer, W., 2001. Assessing teaching presence in a computer conferencing environment, *Journal of Asynchronous Learning Networks* ,5(2).
2. Berge, Z. și Collins M. 1993. Computer Conferencing and Online Education. *The Arachnet Electronic Journal on Virtual Culture*, 1(3).
3. Gangemi, A. , Navigli, R., și Velardi, P., 2003. *The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet*, p. 820-838. disponibil pe: <http://www.loa-cnr.it/Papers/ODBASE-WORDNET.pdf> la data 7.07.2010.
4. Garrison, D. R., Anderson, T., și Archer, W., 2000. Critical inquiry in a text-based environment: Computer conferencing in higher education. *Internet and Higher Education*, 2(2-3), p.87 - 105.
5. Garrison, D. R. și Anderson, T., 2003. *E-Learning in the 21st Century: A Framework for Research and Practice*. 1. London: RoutledgeFalmer.
6. Koschmann, T., Stahl, G., and Zemel, A., 2004. The video analyst's manifesto (or the implications of Garfinkel's policies for the development of a program of video analytic research within the learning sciences), *International Conference on Learning Sciences*, Editia a 6-a, p. 278-285, disponibil pe: <http://www.cis.drexel.edu/faculty/gerry/publications/journals/manifesto.pdf>
7. Lipman, M., 1991. *Thinking in education*. 2. New York: Cambridge University Press.
8. Liscombe, J., Riccardi, G., și Hakkani-Tur, D. 2005. *Using Context to Improve Emotion Detection in Spoken Dialog Systems*, disponibil pe: http://www1.cs.columbia.edu/nlp/papers/2005/liscombe_al_05a.pdf la data 7.07.2010.
9. Miller, G. A., Beckwith, R., Felbaum, C., Gross, D., și Miller, K., 1993. *Introduction to WordNet: An Online Lexical Database*. Disponibil pe: <http://www.mit.edu/~6.863/spring2009/readings/5papers.pdf> la data 7.07.2010.
10. Rourke, L., Anderson, T. Garrison, D. R., Archer, W. 1999. Assessing social presence in asynchronous, text-based computer conferencing. *Journal of Distance Education*, 14(3), p.51-70.
11. Schapire, R. E., și Singer, Y., 2000. BoosTexter: A boostingbased system for text categorization, *Machine Learning*, 39(2/3), p. 135–168.
12. Shaikh, M. , Prendinger, H. și Ishizuka, M. 2007. Emotion sensitive news agent: An approach towards user centric emotion sensing from the news. *The 2007 IEEE/WIC/ACM International Conference on Web Intelligence (WI-07)*, p 614-620
13. Vygotsky, L., 1962. *Thought and language*. 1. Cambridge: MIT Press.
14. Wagner, E.D. 1994. In support of a functional definition of interaction. *American Journal of Distance Education*, 8(2), p.6 – 26.
15. *Cognitive Computation Group, University of Illinois*, disponibil pe: <http://l2r.cs.uiuc.edu/~cogcomp/srl-demo.php> la data 7.07.2010.
16. *Emotion Detection in Textual Information by Semantic Role Labeling and Web Mining Techniques*, disponibil pe : <http://hal.archives-ouvertes.fr/docs/00/10/56/49/PDF/4TFIT2006-CYL-Final.pdf> la data 7.07.2010.
17. *Community of inquiry*, disponibil pe : http://coursedesign.ucalgaryblogs.ca/files/2008/10/community_of_inquiry_model.jpg la data 7.07.2010