

PolyCAFE: Sistem avansat de evaluare a conversațiilor de tip chat bazat pe modelul polifonic

Traian Rebedea, Mihai Dascălu, Ștefan Trăușan-Matu, Ana Teodorescu, Alexandra Ene

Universitatea Politehnica din București

313 Splaiul Independenței, 060042 București

traian.rebedea@cs.pub.ro, mihai.dascalu@cs.pub.ro, stefan.trausan@cs.pub.ro,
ana.teodorescu@cti.pub.ro, alexandra.ene@cti.pub.ro

REZUMAT

Conversațiile de tip chat au devenit foarte populare printre membrii diferitelor comunități online. Odată cu creșterea în popularitate a învățării colaborative asistate de calculator (*Computer Supported Collaborative Learning sau CSCL*), conversațiile de tip chat au început să fie utilizate frecvent în mediile educaționale. Cu toate acestea, problema principală în încercarea de a evalua studenții prin intermediul chaturilor la care participă este că acest proces este extrem de complicat și mare consumator de timp. Astfel, nevoia de a realiza analize și evaluări automate ale studenților pe baza chaturilor a crescut foarte mult.

În consecință, în ultimii ani au fost depuse eforturi pentru dezvoltarea unui instrument care să poată analiza automat conversații chat colaborative utilizate în diverse contexte educaționale. În această lucrare propunem o soluție bazată pe prelucrarea limbajului natural, pe analiza rețelelor sociale, pe analiza semantică latentă, precum și pe metode specifice de data-mining și de regasire de informații. Toate aceste perspective sunt integrate cu scopul de a oferi următoarele facilități: evaluarea participanților, a replicilor aferente și a gradului de colaborare, vizualizarea topicelor discuției și căutare semantică.

Cuvinte cheie

Colaborare, evaluarea participanților în cadrul mediilor colaborative, rețele sociale, analiză semantică latentă, căutare semantică

Clasificare ACM

H5.2. Information interfaces and presentation: Training, help, and documentation and Natural language

INTRODUCERE

Drept rezultat al evoluției continue a web-ului, noi instrumente de colaborare au apărut și, concomitent, s-a dezvoltat și dorința de a putea evalua automat cantitățile imense de informație generate. În momentul de față, chatul este probabil cea mai simplă și cea mai utilă tehnologie utilizată în comunicare.

Computer Supported Collaborative Learning (CSCL) consideră chatul drept elementul central, din ce în ce mai uzitat în mediile academice, pentru susținerea procesului de învățare colaborativă [1]. Astfel, chatul a fost introdus în educația formală și este totodată folosit de către studenți în rezolvarea diverselor probleme, în dezbateri, cu scopul de a-și spori cunoașterea într-un anumit domeniu și de a

învăța de la interlocutorii lor. Mai mult, chatul poate fi folosit de către profesori în evaluarea studenților lor.

Această situație a generat însă anumite probleme: evaluarea manuală a chaturilor consumă foarte mult timp, de aceea nevoia de a utiliza aplicații care să ofere suport în procesul de evaluare a crescut foarte mult.

Raportându-ne la contextul creat, se urmărește dezvoltarea unui sistem care să evalueze fiecare participant al discuției utilizând metode specifice de prelucrare a limbajului natural, reducând astfel timpul necesar în procesul de evaluare manuală.

La ora actuală există foarte puține sisteme de analiză automată și de generare de feedback pe baza istoricului conversației. Explicația este dată de faptul că pentru a evalua o discuție în profunzime este nevoie de prelucrarea limbajului natural, iar tehnologiile existente în analiza lingvistică nu sunt încă suficient de mature, în special în domeniul analizei conversațiilor de tip chat, domeniu care prezintă dificultăți suplimentare față de analiza textelor convenționale.

Până acum au fost dezvoltate câteva sisteme CSCL de analiză a interacțiunilor la nivelul conversațiilor de tip chat sau forum. Printre acestea se numără CORDTRA [22], COALA [23], și alte aplicații utilizate în sistemul Argonaut [24], metoda de vizualizare, KSV [25] și Polyphony [11].

PolyCAFE este unul dintre acestea și se bazează pe o integrare de diverse metode de prelucrare a limbajului natural, analiza rețelelor sociale, analiză semantică latentă, precum și pe metode specifice de data-mining și de regasire de informații [6]. Versiunea curentă a fost dezvoltată și validată în cadrul proiectului FP7 LTFLL (<http://ltfll-project.org>).

Pentru testarea aplicației, studenți din facultatea noastră (din doi ani diferiți) au fost rugați să discute beneficiile instrumentelor de colaborare online (în mod particular wiki-uri, bloguri, chat-uri și forumuri), tratând atât aspectele pozitive cât și cele negative ale fiecărei tehnologii. Mediul folosit a fost ConcertChat (<http://concertchat.sourceforge.net/>), fiecare student fiind reprezentantul unei tehnologii din cele abordate. Corpusul obținut a cuprins peste 40 de chat-uri care au fost apoi evaluate folosind PolyCAFE.

În această lucrare vom prezenta pe larg abordarea folosită de PolyCAFE în analiza chaturilor, care se bazează în principal pe LSA (Latent Semantic Analysis), dar care implică și tehnici specifice de NLP (Natural Language

Processing) [2, 3, 4] și SNA (Social Network Analysis) [4, 5]. De asemenea vom prezenta și o metodă nouă de evaluare a nivelului de colaborare în cadrul unei conversații de tip chat, metode de căutare semantică, aspecte specifice interacțiunii om-calculator și rezultate cheie obținute în urma validării sistemului.

CONCEPTELE DE BAZĂ ALE DIALOGISMULUI BAKHTIN

Analiza din spatele sistemului PolyCAFE este derivată din polifonie și este centrată pe trei concepte inter-dependente: replici - definite pe scurt ca unități de analiză, voci – puncte distincte de vedere care reies din discuție, și ecouri – replicarea unei anumite voci în timp cu implicații ulterioare în discurs [8].

Replicile pot fi definite ca porțiuni de text ale căror limite sunt reprezentate de schimbarea subiectului de discuție [7] și care înglobează elementul principal de analiză din discuție. Replicile exprimă atât acte de comunicare [9], cât și evoluția conversației din perspectiva dezvoltării ei continue. Analiza noastră aderă la perspectiva de separare a replicilor propusă de Dong [10]: introducerea unui punct nou de vedere sau a unei intervenții din partea unui alt participant segmentează discursul, schimbând perspectiva proprie a vorbitorului.

Graful de replici utilizat în cadrul evaluării [11] este construit pe baza a două tipuri de legături între replici: implicite și explicite. Participanții pot adăuga legături explicite pe parcursul sesiunii de chat folosind o facilitare oferită de mediul de conversație – în cazul nostru ConcertChat. Legăturile implicite sunt identificate automat prin intermediul co-referințelor, repetițiilor, lanțurilor lexicale, tiparelor de inter-animare și similarității semantice [6]. În graful aciclic orientat rezultat, fiecare replică este un nod, iar ponderile arcelor sunt date de similarități între replici multiplicat cu încrederea asociată fiecărei legături (valoare predefinită atribuită în funcție de metoda de detecție a legăturii). Orientarea fiecărui arc este dată de evoluția în timp a discuției.

O **voce** exprimă o poziție distinctă, un punct de vedere cu impact asupra conversației [5]. O voce poate fi exprimată și drept o perspectivă sau un topic [9] a unui singur participant sau a unui grup ai căror membri partajează idei similare. Un anumit individ poate personaliza, exprima și adera la diferite voci prin interacțiunea sa cu alți indivizi. Totodată, vocile generale integrează pe lângă vocile personale și vocile exterioare ale altor indivizi, voci care exprimă influența celorlalți asupra opiniei proprii. O replică poate deveni o voce și poate încorpora ecourile unor voci precedente [7].

Ecoul reprezintă replicarea unei voci în timp, cu o putere suficient de mare pentru a influența alte voci în unul sau mai multe contexte. Pot fi identificate două tipuri de ecouri: ecouri individuale, când participantul personalizează o voce, și ecouri colective, când mai mulți participanți reacționează la o voce, îmbogățind contextul. Ecourile vocilor curente pot influența replicile ulterioare și pot, de asemenea, modela vocile personale ale fiecărui participant.

După analizarea tuturor conceptelor de bază, au fost identificate și considerate două efecte majore. În primul rând identificăm un *efect retrospectiv*, sinergic, bazat pe vocile care se suprapun și se întretaș într-o manieră polifonică ale replicilor precedente. Astfel, ecourile lor modelează și influențează replica curentă într-un context dat. În al doilea rând, un *efect prospectiv* exprimă implicații viitoare în firul de discuție cu privire la propriul ecou și perspectivă în contextul considerat, subliniind natura impredictibilă și maleabilă a unei discuții.

Insumând cele spuse mai sus, putem concluziona că o colaborare este bazată pe întretașerea vocilor într-o manieră polifonică și pe inter-animare. Raportat la abordarea structuralistă, perspectiva dialogică conferă sistemului nostru capacitatea de a evalua profund interacțiunea dintre participanți, într-un mediu colaborativ.

DESCRIEREA SISTEMULUI

PolyCAFE este o platformă online ce cuprinde widget-uri Web construite în conformitate cu standardul W3C Widgets 1.0. Widget-urile pot fi integrate în orice platformă web care integrează un container W3C pentru widget.

Din punct de vedere tehnic, PolyCAFE este constituit dintr-o serie de procesări, pe diverse dimensiuni, după cum urmează:

Componenta Cantitativă presupune realizarea de procesari specifice prelucrării limbajului natural: verificarea ortografiei, aducerea la radacina (stemming), tokenizare, adnotare cu parti de vorbire (POS Tagging). Deasemenea, aceasta presupune si determinarea numărului de caractere pentru fiecare rădăcină și a numărului de apariții pentru fiecărui concept.

Componenta Calitativă presupune determinarea similarității semantice (similaritate bazată pe spațiul de vectori din LSA (Latent Semantic Analysis)), utilizarea de topice predefinite folosite pentru măsurarea completitudinii replicilor, urmărirea evoluției firului de discuție cu privire la impactul viitor în cadrul discuției și cu privire la coerența sa internă, evaluarea gradului de relevanță a replicii măsurată prin impactul său asupra discursului.

În sfârșit, *Componenta Socială* presupune realizarea unei analize a rețelelor sociale aplicată asupra grafului replicilor.

Sarcinile principale ale sistemului presupun detectarea legăturilor implicite folosind tipare, repetiții și distanțe semantice bazate pe WordNet și LSA [6], evaluarea replicilor și analiza colaborării bazată pe graful replicilor și pe scorurile asociate acestora. Rezultatul acestor calcule generează feedback pe câteva nivele distincte: pentru fiecare replică din conversație, pentru fiecare participant și pentru conversația luată ca întreg.

Prezentarea feedbackului către utilizatori este realizată în widgeturi web simple, care pot fi utilizate independent sau împreună și care pot fi integrate în majoritatea mediilor de învățare [15]. Există widgeturi pentru fiecare nivel de feedback discutat mai sus, precum și alte widgeturi ajutătoare.

Widgetul de feedback pentru conversație (Figura 1) prezintă statistici despre întregul chat: cele mai frecvente seturi sinonimice, cele mai relevante concepte care sunt prezente în chat și în corpusul specific domeniului folosit în antrenarea LSA-ului, o sugestie de concepte derivate din spațiul de vectori LSA care sunt similare din punct de vedere semantic cu cele discutate în cadrul chat-ului și statistici privind densitatea grafului de replici, procentajul de opinii personale și argumentări, etc.

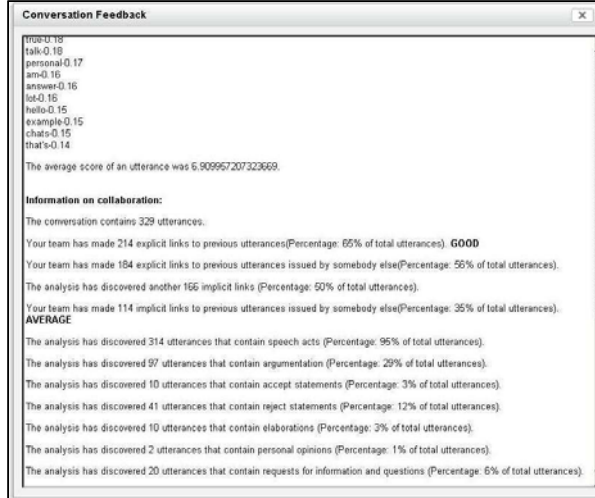


Figura 1. Widgetul de feedback pentru conversație

Widgetul de feedback pentru participanți (Figura 2) oferă o evaluare pentru fiecare participant pe multiple nivele: relevanța replicilor cu referire la domeniul din corpus, prezența socială și importanța, coerența, etc.

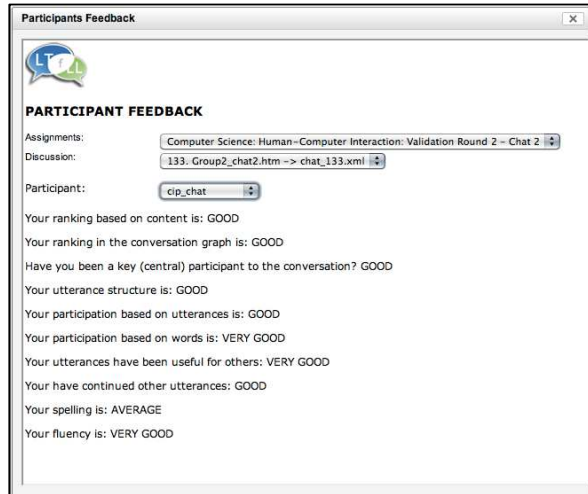


Figura 2. Widgetul de feedback pentru participanți

Widgetul de feedback pentru replici (Figura 3) oferă indicatori ai valorii specifice fiecărei postări: acte de vorbire și pattern-uri de argumentare care sunt prezente în replică, plus rolul social și semantic.

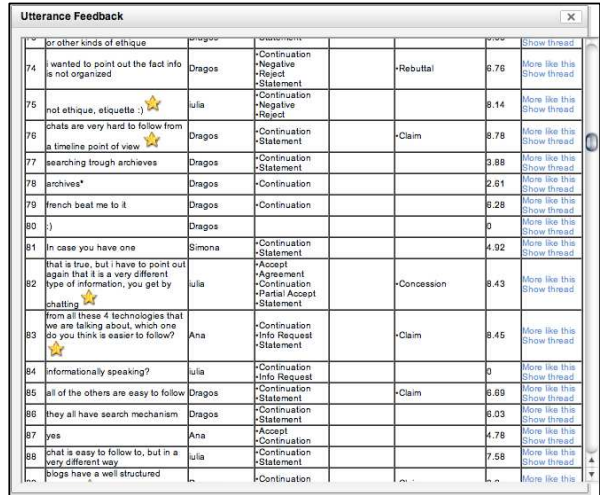


Figura 3. Widgetul de feedback pentru replici

Vizualizarea conversației este un widget ajutător, care conține o afișare intuitivă a grafului replicilor și a grafului colaborării, așa cum arată Figura 4:



Figura 4. Widgetul de vizualizare a unei conversații

Widgetul de căutare în cadrul discuțiilor oferă un mecanism de catalogare și ranking a participanților ținând cont nu numai de itemi lexicali, ci și de relațiile semantice, respectiv de scorul fiecărei replici după cum au fost detectate în cadrul procesului de evaluare a replicilor:

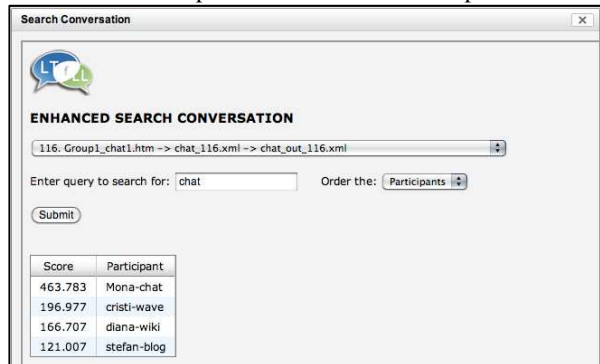


Figura 5. Widgetul de căutare semantică

PROCESUL DE EVALUARE

Informația lexicală, informația sintactică și cea semantică reprezintă cele trei niveluri folosite pentru a extrage trăsăturile replicilor din cadrul conversațiilor studenților. Acestea sunt luate în considerare în analiza implicării și a gradului de cunoaștere / familiaritate al participanților la discuției.

La nivel lexical sunt calculate metrici de suprafață pentru toate replicile unui participant în vederea determinării unor factori precum fluența în vorbire, ortografie, dicție sau structura replicii [12, 13]. Toți acești factori sunt combinați liniar și un scor parțial este obținut pentru fiecare participant, fără a lua analiza din perspectivă sintactică sau semantică ce se discută efectiv.

Aceste metrici sunt derivate din studiile lui Page care a explorat posibilitatea de a folosi calculatorul pentru a evalua și nota în mod automat eseurile studenților la fel de eficient ca un tutore uman [13]. Scopul de a îmbunătăți capacitățile studenților prin practică, folosind capacitatea calculatoarelor de a realiza statistici și de analiza discursul a fost preluat și în cadrul PolyCAFE. Pentru a realiza o analiză statistică, Page a corelat două concepte: prox-uri (aproximările intereselor realizate de calculator) cu trins-uri (variabilele intrinseci – măsuri date de evaluatori umani, folosite pentru evaluare).

Pe lângă factorii obținuți din studiile lui Page, în cadrul acestui nivel este aproximată și lizibilitatea drept ușurința cu care o persoană poate parcurge și înțelege discuția purtată [14]. Aceasta are un rol important întrucât cercetările în acest domeniu arată că un text ușor de citit de către audiență are un mare impact asupra înțelegerii, reținerii, vitezei de citire și a persistenței informațiilor citite. Lizibilitatea este de multe ori folosită ca modalitate de evaluare în mod inconștient, pe baza unei perspective asupra celorlalți participanți la chat, dar pentru evaluarea ei este folosit un set de formule care realizează o potrivire între discuția globală și nivelul așteptat al capacităților de citire ale audienței. Aceste formule estimează complexitatea globală a textului, oferind astfel mijloacele folosite în identificarea unei audiențe țintă [14].

Următorul nivel adresează **analiza morfologică** bazată pe verificarea ortografiei, aducerea la rădăcina comună, tokenizare și **etichetarea părților de vorbire** [20].

În cele din urmă este realizată o evaluare semantică folosind **analiza semantică latentă** (Latent Semantic Analysis - LSA) care presupune construirea unui spațiu de vectori [17, 18, 19]. Prin aplicarea funcției cosinus se evaluează importanța unei replici pe baza similarității semantic dintre replică și întreaga discuție, aceasta similaritate evaluându-se ținându-se cont de un set predefinit de topice care se caută să fie acoperite în cadrul discuției.

Utilizarea LSA în cadrul PolyCAFE este bazată pe o matrice termen-document construită dintr-un corpus de chaturi, adnotat automat cu părțile de vorbire corespunzătoare. Odată ce această matrice este populată, se aplică *Tf-Idf* (frecvența termenilor – inversul frecvenței documentelor), urmat de descompunerea în valori singulare (SVD) și proiecția spațiului după k dimensiuni [21]. Valoarea empirică utilizată pentru k este 300,

aproximativ egală cu rădăcina pătratică a numărului de concepte existente în spațiul antrenat.

Segmentarea presupune împărțirea discuțiilor în unități cu înțeles, luând în considerare coeziunea și unitatea replicilor. În implementarea curentă folosim replicile acelorași participanți datorită consistenței lor, divizate după o fereastră fixă.

O aplicație imediată a LSA presupune **identificarea topicelor** care se bazează pe similaritatea de tip cosinus a fiecărui cuvânt cu întregul document. Valoarea maximă reflectă cea mai bună corelație între concept și toți termenii utilizați în cadrul discuției.

De asemenea, în cadrul PolyCAFE sunt utilizați o serie de factori specifici **analizei rețelelor sociale**. În consecință este generat un graf în care nodurile sunt reprezentate prin participanții din cadrul mediului colaborativ, iar legăturile dintre aceștia sunt generate pe baza grafului de replici [15].

Din punctul de vedere al rețelelor sociale, se determină cel mai competitiv și implicat participant al discuției folosind: gradul, centralitatea (fie de apropiere, fie în graf) și rangul utilizatorului similar cu bine-cunoscutul algoritm *Google Page Rank* [16].

Aceste metrici sunt aplicate pe:

- numărul efectiv de replici inter-schimbate între participanți, care oferă o abordare cantitativă a implicării;
- suma scorului replicilor care presupune o evaluare calitativă a conținutului fiecărei replici.

Toate notele parțiale obținute prin aplicarea metricilor din analiza rețelelor sociale sunt relative în sensul că ele oferă note comparative între participanții la aceeași conversație și nu între chaturi diferite.

Procesul de notare a replicilor

Scorul efectiv al fiecărei replici are la bază 3 componente distincte: una cantitativă, una calitativă, și una socială [21].

Perspectiva cantitativă evaluează replica la nivel de suprafață. Scorul asignat în acest fel ia în considerare lungimea în caractere a cuvintelor rămase după eliminarea cuvintelor frecvente și irelevante, verificarea ortografiei și aducerea la rădăcina comună. Pentru a reduce impactul repetițiilor inutile, funcția logaritm este aplicată asupra numărului de apariții ale fiecărui cuvânt.

O dimensiune mai interesantă este cea **calitativă** care implică calculul de similarități între replica curentă și:

- vectorul asociat întregului chat pentru a determina corelarea cu discuția purtată;
- vectorul unui set predefinit de topice desemnate de către tutore drept cuvinte cheie ale discuției (dacă sunt alese topice libere, acest factor devine 1)
- fiecare replică ulterioară legată cu cea curentă pentru a evalua impactul și ponderea legăturii dintre ele.

Astfel, această dimensiune implică utilizarea LSA în determinarea a *4 componente diferite*: coerența firelor de discuție, impactul viitor, relevanța unei replici și

completitudinea acesteia. Nota calitativă finală este obținută prin înmulțirea acestor factori.

Plecând de la graficul replicilor, **coerența** unei replici reprezintă procentul de replici interconectate cu aceasta care au o similaritate mai mare decât un prag impus. Pentru a asigura coeziune internă și continuitate, similaritățile dintre orice replici adiacente ale aceluiași fir de discuție trebuie să depășească un anumit prag (în cazul nostru 0.1)

Impactul viitor măsoară transferul de informație și reprezintă ecoul unei anumite voci, fiind estimat prin măsurarea similarității între două replici interconectate printr-o legătură fie implicită, fie explicită.

Relevanța exprimă importanța fiecărei replici cu privire la întreaga discuție. Aceasta poate fi ușor măsurată calculând similaritatea între replica curentă și vectorul asociat întregului chat, determinând astfel corelația cu întreaga discuție.

Completitudinea este obținută prin evaluarea similarității între replică și vectorul setului de cuvinte cheie specificat de către tutore.

Perspectiva **socială** presupune aplicarea metricilor specifice analizei rețelelor sociale asupra graficului de replici. În acest context există anumite restricții: spre exemplu gradul de centralitate nu este foarte relevant deoarece toate legăturile urmăresc cursul conversației, fiind, deci, orientate în aceeași direcție.

Scorul final al fiecărei replici este obținut prin multiplicarea sub-scorurilor obținute pe fiecare dimensiune în parte.

Notarea participanților

Fiecare factor al analizei, indiferent de nivelul la care este identificat, este convertit în procente, iar nota finală per participant presupune realizarea unei combinații lineare între toate scorurile calculate anterior.

EVALUAREA COLABORĂRII

Colaborarea într-un mediu de tip chat poate fi evaluată pe baza următoarelor aspecte: coeziune și colaborare socială, colaborare cantitativă, colaborare bazată pe câștig informațional, toate descrise în continuare.

Coeziunea și colaborarea socială

Plecând de la analiza rețelelor sociale, atât la nivelul de suprafață cât și la nivel semantic, această metrică este folosită pentru evaluarea echilibrului și coeziunii între toți participanții la discuție.

Un fapt universal acceptat despre mediul colaborativ este acela că, având cât mai mulți participanți implicați similar în discuție (atât ca dimensiune socială cât și la nivelul cunoștințelor deținute), ei vor colabora cu atât mai bine, toți având interese asemănătoare și cunoștințe comune sau interschimbabile.

Pentru a avea o măsură normalizată a valorilor fiecărui factor luat în considerare în cadrul analizei rețelelor sociale, coeficientul de variație este calculat pentru fiecare metrică în parte. Rezultatul final reprezintă 100% minus valoarea medie a tuturor rezultatelor parțiale, întrucât

colaborarea este cu atât mai bună cu cât mai mulți participanți se vor implica similar în discuție.

Colaborarea cantitativă

Cea mai simplă abordare a evaluării nivelului de colaborare a unui chat se obține utilizând numărul de legături implicite și explicite, cu un coeficient de încredere asociat fiecăreia. Colaborarea este ilustrată perfect când anumite idei sau puncte de vedere din replica curentă sunt preluate sau transferate într-o replică viitoare a unui alt participant la discuție. Cu alte cuvinte, colaborarea se poate măsura prin numărarea legăturilor implicite sau explicite dintre două replici date de doi participanți diferiți.

Colaborarea bazată pe câștig informațional

Cunoașterea poate fi construită în două moduri diferite: *construire internă a cunoașterii*, atunci când cunoașterea este rezultat al studiului și experienței individuale, și *construire colaborativă a cunoașterii* prin interacțiunea cu alte persoane [21].

Plecând de la cele două tipuri de construire a cunoașterii, se pot defini două tipuri de câștig: **câștig informațional personal** când replicile interconectate aparțin aceleiași persoane și **câștig colaborativ**, când informațiile din firul de discuții sunt date de către participanți diferiți.

Din perspectivă computațională, fiecare replică are un scor determinat pe baza procesului de evaluare descris anterior și pe baza câștigului informațional descompus în câștig personal și câștig colaborativ.

Mai mult câștigul informațional se obține adunând, pentru toate replicile anterioare, nota și câștigul aferent înmulțit cu similaritatea semantică dintre replica anterioară și cea curentă.

Combinând nota cu câștigul informațional obținem o estimare bună a importanței actuale a replicii într-un context dat. În principal similaritatea reflectă puterea, impactul și transferul de cunoștințe între două replici conectate printr-o legătură implicită sau explicită.

Ideea de câștig este preluată din teoria informației și este folosită pentru a sublinia importanța și impactul viitor al replicii curente, luând în considerare toate replicile anterioare interconectate [21].

Următoarele două metrici sunt folosite pentru evaluarea gradului de colaborare dintr-o conversație chat:

$$\begin{aligned} & \text{colaborare bazată pe note} \\ & \frac{\sum_{\text{replicile}} \text{toate } \text{câștig colaborativ}(u)}{u} \\ & = \frac{\sum_{\text{replicile}} \text{toate } \text{notă}(u)}{u} \end{aligned}$$

$$\begin{aligned} & \text{colaborare bazată pe câștig} \\ & \frac{\sum_{\text{replicile}} \text{toate } \text{câștig colaborativ}(u)}{u} \\ & = \frac{\sum_{\text{replicile}} \text{toate } \text{câștig}(u)}{u} \end{aligned}$$

Prima metrică se referă la estimarea procentului total al importanței unei replici relativ la informația construită sau transferată într-o manieră colaborativă. Cea de-a doua este

folosită pentru a evalua colaborarea relativ la câștigul informațional total. Aceste două metrici determină procentajul de colaborare efectivă, eliminând partea de monolog intern la nivelul fiecărui individ.

Nota **colaborării totale** se obține înmulțind cele patru metrici anterioare. În concluzie, câștigul informațional măsoară puterea ecoului, scorul exprimă importanța individuală a fiecărei unități analizate și, prin combinarea acestora, metoda propusă evaluează colaborarea, mai ales din punctul de vedere al inter-animării și al întrețeserii vocilor.

Căutare semantică

Posibilitatea de a căuta în cadrul unei conversații este o altă componentă importantă a sistemului PolyCAFE.

Pornind de la o interogare dată, două tipuri de rezultate sunt evaluate relativ la o singură conversație de tip chat: o clasificare a participanților după cea mai bună performanță totală cu privire la cererea formulată și o listă de replici, în ordinea descrescătoare a relevanței. În cazul de față, aceste facilități de căutare sunt centrate pe mecanismul de notare al fiecărei replici.

Pentru a evalua relevanța răspunsului față de o interogare dată sunt parcurși trei pași. Mai întâi, interogarea este îmbogățită sintactic și semantic folosind synset-uri din WordNet și selectând vecinii cei mai apropiați din spațiul de vectori LSA. Pentru a reduce numărul de elemente noi adăugate în interogare, o valoare de prag și un maxim pentru numărul de vecini sunt utilizate. Obiectivul final este de a obține o interogare extinsă și un vector asociat în spațiul LSA.

Pasul următor se adresează nivelului sintactic, evaluând numărul aparițiilor fiecărui concept în fiecare replică.

Pasul al treilea se bazează pe aplicarea similarității cosinus între vectorul interogării și vectorul fiecărei replici. Scorul final se obține înmulțind scorul semantic cu cel social și cu suma dintre scorul și a câștigului informațional determinate anterior.

CONCEPTE SPECIFICE INTERACȚIUNII OM-CALCULATOR

Oamenii își crează modele mentale despre cum merg lucrurile din jurul lor sau despre cum ar trebui să meargă într-o anumită situație dată. În cazul de față, fiecare utilizator al sistemului ConcertChat, provenind dintr-un mediu academic axat pe domeniul IT, are o imagine clară despre aspectul și utilitatea acestuia, sistemele de conversații chat fiind momentan foarte răspândite și folosite în cadrul universității.

Pornind de la aceste premise, în cadrul experimentelor de validare a fost utilizat un mediu de chat ușor de folosit și intuitiv, dar care conține și o serie de îmbunătățiri precum: posibilitatea de a adăuga legături explicite între replici și o componentă de vizualizare de tip whiteboard (cu care utilizatorul poate interacționa) care se poate dovedi extrem de utilă în cazul colaborării în cadrul unui proiect.

Prima facilitate – specificarea de legături explicite între replici – este foarte utilă întrucât, în cazul unei conversații de tip chat cu mulți participanți, utilizatorii rareori au răbdarea să îi lase pe ceilalți să își termine rândul la

replică. Ca urmare, conversațiile devin greu de urmărit, mai ales dacă se încearcă ajungerea la un consens, baza oricărei comunicării eficiente.

Totodată, componenta de vizualizare de tip whiteboard se poate folosi la realizarea unor schițe despre cum arăta în viziunea unui anumit participant o interfață grafică dintr-un proiect mai amplu sau pentru a da indicații referitoare la locații, aspecte cu adevărat utile în conturarea ideilor în cadrul unor sesiuni de brainstorming specifice mediului academic.

De asemenea, modelele pentru widget-urile de căutare în cadrul unei conversații, vizualizarea întregii conversații și a gradului de colaborare, feedback-ul oferit pentru o replică și individual, la nivel participant, au rezultate ușor interpretabile, bazându-se pe reprezentări clare ale cunoștințelor dobândite în urma analizei discursului.

Modelele de interfețe om-calculator au la bază ideea de învățare: utilizatorul descoperă lucruri noi parcurgând un anumit itinerariu, în ordinea gândită de dezvoltatori, până la rezultatul final.

Un alt concept al interfețelor om-calculator este învățarea. Învățarea se bazează pe achiziția de noi cunoștințe, deprinderi și informații. În cazul de față, învățarea reprezintă gradul de familiarizare ale utilizatorilor cu interfețele proiectate să-i ajute în rezolvarea diferitelor probleme și să obțină mult mai rapid feedback și informații relevante cu privire la discuțiile purtate.

Potrivit teoriei, de la o anumită vârstă învățarea începe să devină un proces din ce în ce mai dificil. Din această perspectivă este crucială ergonomia unei interfețe la nivel de plasare a anumitor butoane, pictograme, opțiuni acolo unde utilizatorul se așteaptă să le regăsească.

Învățarea în cazul PolyCAFE se poate face utilizând diverse widget-uri:

- utilizatorul poate folosi marcajul aferent replicilor cu scorul cel mai mare la nivelul conversației pentru a reține sau pentru a recapitula lucrurile cele mai importante care s-au discutat;
- fiecare participant poate analiza implicarea sa în discuție și gradul său colaborare efectivă folosind widgetul de vizualizare a conversațiilor; pornind de la feedback-ul asupra participanților, acesta își poate da seama de lipsurile sale, fie că țin de scriere ortografică, implicare, fluentă sau modul în care i-a sprijinit pe alții de-a lungul discuției. Evident, fiecare participant își poate îmbunătăți astfel performanțele în discuțiile viitoare;
- tutorele evaluează mult mai repede participanții folosind feedback-ul conversației, sau, dacă este interesat de performanțele individuale, obține prompt informații din widgetul de feedback pentru participanți;
- participantul sau tutorele poate de asemenea folosi widget-ul de căutare pentru a identifica fix informația care îl interesează din discuția respectivă.

Astfel, PolyCAFE sprijină achiziția de noi cunoștințe pentru fiecare categorie de utilizator în parte, bineînțeles cu constrângerea ca participantul să acorde un minim interes necesar.

În altă ordine de idei, interfețele om-calculator au la bază și principiul rezolvării de probleme. Rezolvarea de probleme se referă la stabilirea unui obiectiv și a unui drum de urmat pentru a obține rezultatul dorit, drum ulterior parcurs iterativ. În acest caz sunt relevante:

- extragerea celor mai importante replici ale conversației pe baza scorului și al câștigului informațional – obiectivul presupune sumarizarea conversației date, iar drumul de urmat este reprezentat de alegerea unui procentaj de detaliu la care se dorește rezumatul discuției respective;
- căutarea în cadrul unei conversații: este dat drept obiectiv de căutat un set de cuvinte în cadrul unei anumite conversații, iar drumul este similar cu cel descris anterior;
- feedback-ul asupra participanților: se dorește evaluarea și notarea utilizatorilor pe baza unei conversații de tip chat (obiectiv). Drumul urmat presupune vizualizarea feedbackului pentru fiecare participant în parte și compararea acestora, pentru a-l descoperi pe cel cu rezultatele cele mai bune.

O altă caracteristică luată în considerare presupune modelarea spațiului vizual prin folosirea de metafore ce redau cât mai bine ideea prezentată. De exemplu, replicile importante sunt marcate cu o steluță – steaua strălucește pe cer, similar cum replica se remarcă prin importanța ei în discuție. În plus, la vizualizarea conversațiilor, firele de discuție sunt redade ca un drum într-un graf neorientat, concepte și percepție familiare utilizatorilor sistemului.

EXPERIMENTUL DE VALIDARE ȘI REZULTATELE SALE

Un prim experiment de validare a fost efectuat la cursul de Interacțiune Om-Calculator al Catedrei de Calculatoare și a implicat 9 studenți din anul IV, 5 tutori și un profesor care au folosit sistemul PolyCAFE.

Experimentul a fost structurat după cum urmează: studenții s-au documentat folosind diverse materiale pe un subiect dat și apoi, în grupe de 4-5, au luat parte la o dezbatere folosind ConcertChat. După finalizarea discuției, studenții au folosit feedback-ul automat dat de sistem pentru a afla informații suplimentare despre discuția anterioară, de exemplu: rolul fiecăruia în conversație și ce pot face fiecare pentru a-și îmbunătăți prestațiile.

Această activitate a fost monitorizată de doi tutori care au oferit ajutor studenților și au luat notițe în legătură cu modul de folosire al componentelor, ce întrebări au fost puse și alte comentarii din partea studenților. A fost urmat de un chestionar cu 32 de afirmații de validare pentru studenți, cu răspunsul pe o scală de la 1 la 5 (1-nu sunt deloc de acord, 2-nu sunt de acord, 3-neutru, 4-sunt de acord, 5-sunt total de acord).

După terminarea chestionarului, echipa de validare a mai realizat o întâlnire de grup pentru a identifica cele mai importante avantaje și dezavantaje ale sistemului, precum și sugestii de îmbunătățire.

Experimentul de validare al tutorilor a avut loc într-o manieră puțin diferită. Tutorii au fost puși să ofere feedback pentru o conversație folosind PolyCAFE și

pentru o altă discuție fără a folosi sistemul. După acest pas, le-a fost dat un chestionar cu 35 de afirmații de validare, folosind aceeași scală. Ulterior, toți tutorii au luat parte la o întâlnire de grup în care au partajat punctele lor de vedere cu privire la utilitatea fiecărei componente, seriozitatea feedback-ului și îmbunătățiri ulterioare care pot fi aduse sistemului.

Amândouă experimentele de validare au avut succes și s-au finalizat prin rezultate încurajatoare. Cele 35 de întrebări adresate tutorilor au fost validate în totalitate, obținând punctaje între 3,50 și 5 (din maxim 5). Cel mai important rezultat a fost acela că tutorii consideră sistemul relevant și folositor în activitatea lor, timpul total de oferire a unui feedback studenților scăzând simțitor în condițiile unei calități mai bune a feedback-ului.

Astfel, caracteristicile principale ale sistemului evidențiate sunt: calitatea superioară a feedback-ului, timpul de acordare a unui feedback de către tutore pentru studenți scade cu minim 30%, sistemul este relevant și folositor pentru munca tutorilor.

Pentru o mai bună analiză a datelor, afirmațiile au fost împărțite în cinci categorii: eficacitate pedagogică, bună funcționare, utilizabilitate, satisfacție și încărcătură cognitivă, regăsite în tabelul 1.

Tabel 1: Rezultatele detaliate pe categorii

Categorie	Medie
Eficacitate pedagogică	4.11
Bună funcționare	5.00
Încărcătură cognitivă	4.60
Utilizabilitate	4.36
Satisfacție	4.57
Total	4.53

De cealaltă parte, studenții au validat 28 din cele 32 de întrebări, cu punctaje între 3.66 și 5. În cadrul grupurilor țintă ei au adus la cunoștință câteva informații ce îi induceau în eroare, minore raportat la întreaga aplicație. De asemenea, se observă că tutorii au dat note per ansamblu mai mari decât studenții. Studenții au considerat că partea cea mai utilă a sistemului o reprezintă vizualizarea conversației.

A doua etapă de validare a sistemului PolyCAFE a implicat 35 de studenți dintre care 25 în grupul experimental și 10 în grupul de control, împreună cu 6 tutori în grupul experimental. S-au realizat câte 2 conversații chat pentru fiecare echipă de 5 studenți, deci în total 2 X 7 conversații chat. Majoritatea studenților nu mai folosiseră PolyCAFE până atunci, dar toți tutorii erau familiari cu sistemul.

Rezultatele obținute au fost și mai bune, obținându-se o corelare foarte mare între clasamentele realizate automat de sistem și cele date de participanții înșiși.

CONCLUZII

În această lucrare am prezentat PolyCAFE, un sistem care oferă feedback și suport pentru cei care folosesc discuțiile de tip chat sau forum în activitățile lor de învățare. Pornind de la teoria dialogismului și polifoniei, sistemul utilizează o multitudine de procedee pentru a evalua

replici, participanți și gradul de colaborare. Sistemul a fost testat într-un context educațional formal în vederea validării utilității și verificării corectitudinii feedback-ului oferit utilizatorilor.

Rezultatele obținute în urma rundelor de validare sunt încurajatoare și ne permit să concluzionăm că evaluarea globală a contribuției unui participant într-o conversație purtată într-un mediu colaborativ poate fi realizată, chiar dacă factorul subiectiv nu poate fi complet eliminat din cadrul analizei.

MULȚUMIRI

Ne dorim să mulțumim tuturor studenților din cadrul Facultății de Automatică și Calculatoare, Universitatea "Politehnica" din București, care au participat la experimente. Activitățile de cercetare prezentate în această lucrare au fost parțial efectuate în cadrul proiectului PC7 LTfLL (Language Technologies for Lifelong Learning).

REFERINȚE

1. Stahl, G.: *Group Cognition: Computer Support for Building Collaborative Knowledge*. MIT Press, 2006
2. Rose, C. P., Wang, Y.C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., Fischer, F.: *Analyzing Collaborative Learning Processes Automatically: Exploiting the Advances of Computational Linguistics in Computer-Supported Collaborative Learning*, ijCSCL, 2007
3. Dong, A.: *Concept formation as knowledge accumulation: A computational linguistics study*. *Artif. Intell. Eng. Des. Anal. Manuf.* 20, 1, 35-53, 2006
4. Dascalu, M., Chioasca, E.V., Trausan-Matu S.: *ASAP - D. Dochev, M. Pistore, and P. Traverso (Eds.): AIMS 2008, LNAI 5253, Springer, 58-68, 2008*
5. Trausan-Matu, S., Rebedea, T.: *Polyphonic Inter-Animation of Voices in VMT*. In: Stahl, G. (Ed.), *Studying Virtual Math Teams*, pp. 451-473, Boston, MA, Springer US, 2009
6. Trausan-Matu, S., & Rebedea, T.: *A Polyphonic Model and System for Inter-animation Analysis in Chat Conversations with Multiple Participants*. In A. Gelbukh (Ed.), *Cicling 2010, LNCS. 6008*, pp. 354-363, Springer Berlin / Heidelberg, 2010
7. Bakhtin, M.M.: *Speech genres and other late essays*, University of Texas, Austin, 1986
8. Bakhtin, M.M., *Problems of Dostoevsky's Poetics*, University of Minnesota Press, Minneapolis, 1993
9. Linell, P.: *Rethinking language, mind, and world dialogically*, Information Age Publications, Greenwich, 2009
10. Dong, A.: *The latent semantic approach to studying design team communication*, *Design Studies*, 26(5): 445-461, 2005
11. Trausan-Matu, S., Rebedea, T., Dragan, A., & Alexandru, C.: *Visualisation of Learners' Contributions in Chat Conversations*. In J. Fong & F. L. Wang (Eds.), *Blended Learning: Addison-Wesley, 2007*
12. Anderson, J. R.: *Cognitive psychology and its implications*, New York, Freeman, 1985
13. Page, E. B. Paulus, D. H.: *Analysis of essays by computer. Predicting Overall Quality*, U.S. Department of Health, Education and Welfare, 1968
14. http://www.streetdirectory.com/travel_guide/15672/writing/all_about_readability_formulas_and_why_writers_need_to_use_them.html
15. Rebedea, T., Dascalu, M., & Trausan-Matu, S., *Overview and preliminary results of using PolyCAFE for collaboration analysis and feedback generation*. In *Proceedings of ECTEL 2010*, 2010
16. Dascălu, M., Chioașcă, E.-V. Trăușan-Matu, S.: *ASAP – An Advanced System for assessing chat participants*. In: D. Dochev, M. Pistore, and P. Traverso (Eds.): *AIMSA 2008, LNAI 5253, Springer*, pp. 58–68, 2008
17. Landauer, K. Th., Foltz, W. P., Laham, D.: *An Introduction to Latent Semantic Analysis*. *Discourse Processes*, 25, 259-284, 1998
18. Miller, T.: *Latent semantic analysis and the construction of coherent extracts*. In: Nicolov, N. Botcheva, K., Angelova, G. and Mitkov, R., (eds.), *Recent Advances in Natural Language Processing III*. John Benjamins, pp. 277–286, 2004
19. Wiemer-Hastings, P., Zipitria, I.: *Rules for syntax, vectors for semantics*. In: *proceeding of the 23rd Annual Conference of the Cognitive Science Society*, 2001
20. Manning, C., Schütze, H.: *Foundations of statistical Natural Language Processing*. MIT Press, Cambridge (Mass.), 1999
21. Dascalu, M., Rebedea, T. & Trausan-Matu, S.: *A Deep Insight in Chat Analysis: Collaboration, Evolution and Evaluation, Summarization and Search*, *AIMSA 2010, LNAI 6304, 191-200, Springer, 2010*
22. Hmelo-Silver, C. E., Chernobilsky, E., & Mastro, O.: *Representations for Analyzing Tool-mediated Collaborative Learning*. *Proceedings of ICLS 2006*, New Brunswick, 2006
23. Dowell, J., & Gladisch, T.: *Design of argument diagramming for case-based group learning*. *ACM International Conference Proceeding Series; Vol. 250*, 99–105, 2007
24. Harrer, A., Hever, R., & Ziebarth, S.: *Empowering researchers to detect interaction patterns in e-collaboration*. *Frontiers in Artificial Intelligence and Applications; Vol. 158*, 503–510, 2007
25. Teplovs, C.: *The Knowledge Space Visualizer: A Tool for Visualizing Online Discourse*. *Proceedings of the International Conference of the Learning Sciences*, 2008