

***ReaderBench* - analiza bazată pe coeziune a implicării și a colaborării participanților în cadrul conversațiilor CSCL**

Mihai Dascălu, Ștefan Trăușan-Matu

Universitatea Politehnica din București

313 Splaiul Independenței, 060042

București, România

mihai.dascalu@cs.pub.ro,

stefan.trausan@cs.pub.ro

Philippe Dessus

LSE, Université Pierre-Mendès France

F-38040 Grenoble CEDEX 9, Franța

philippe.dessus@upmf-grenoble.fr

REZUMAT

Cu toate că tehnologiile care adresează învățarea colaborativă asistată de calculator (CSCL – Computer Supported Collaborative Learning) au dobândit un rol din ce în ce mai important în mediile educaționale, există doar câteva sisteme automate care vizează identificarea implicării participanților, a construirii de cunoștințe și evaluarea colaborării, dimensiuni centrale în procesul laborios, de lungă durată, de analiza a conversațiilor cursanților. Sistemul *ReaderBench* propune un model de analiză bazat pe coeziune ce integrează tehnici multiple de prelucrare a limbajului natural, un mecanism de notare a importanței intervențiilor și o metodă comprehensivă de evaluare a colaborării, derivate din procedeul de construcție socială a cunoștințelor, reflectate la nivelul intervențiilor participanților prin coeziune. În plus, prin combinarea perspectivei globale a întregii conversații cu o viziune mai rafinată axată pe fiecare participant, obținem o evaluare detaliată a conversațiilor de tip chat, cu accent pe modelarea interacțiunii și colaborării participanților. Ca rezultat al validării rezultatelor, un studiu preliminar reflectă faptul că analiza automată respectă evaluările tutorilor.

Cuvinte cheie

Învățarea colaborativă asistată de calculator (CSCL - Computer Supported Collaborative Learning), coeziunea textelor, dialogism, colaborare, analiza semantică latentă (LSA – Latent Semantic Analysis), alocarea latentă Dirichlet (LDA – Latent Dirichlet Allocation).

Clasificare ACM

I.2.7 Natural Language Processing.

INTRODUCERE

Studentii care colaborează pe forumuri și chat-uri folosesc limbajul ca element mediator pentru construirea cunoștințelor, în scop de negociere, dar și pentru definirea de noi artefacte culturale [24]. Activitatea de lectură reprezintă punctul forte al fiecărui scenariu educațional tipic CSCL. De obicei, studenții se implică în lectură pentru a dobândi noi informații, pentru a conștientiza modul în care colegii se familiarizează cu aceste informații sau pentru a înțelege ceea ce li se cere din partea profesorilor sau îndrumătorilor. Aceștia își pot reciti cu atenție intervențiile înainte și după publicarea lor în cadrul mediului colaborativ.

Coeziunea reliefată în relațiile de similaritate lexicală și semantică din cadrul textului reprezintă una dintre cele mai importante caracteristici ce conduce la înțelegerea textelor citite [25]. Lipsa de coeziune din cadrul materialelor poate spori dificultatea înțelegerii textului, în special pentru studenții ce dețin cunoștințe mai reduse [20]. Totodată, această lipsă de coeziune poate fi determinată și de o modificare de tematică, generând astfel diverse fire de discuție [10]. Cu toate acestea, evaluarea bazată pe coeziune a materialelor de studiu sau a intervențiilor studenților reprezintă o sarcină foarte solicitantă, iar sistemele care analizează în mod automat intervențiile în timpul sesiunilor de învățare prezintă un interes deosebit [13, 30]. Aceste sisteme utilizează adesea analiza semantică latentă (LSA – Latent Semantic Analysis) [17, 18] care poate fi utilizată pentru a identifica relațiile de coeziune din cadrul textelor [15].

Sistemul *ReaderBench* [11] adresează multiple scopuri și propune mecanisme bazate pe coeziune pentru a evalua contribuția sau nivelul de implicare al studenților la discuții, precum și colaborarea acestora pe parcursul conversației. Factorii anteriori sunt prezentați din două perspective: cea individuală, centrată pe fiecare participant, și cea holistică, globală care adresează discuția per ansamblu. Prima perspectivă permite studenților să înțeleagă în mod individual gradul în care contribuie la o anumită discuție, în timp ce cea de-a doua este utilă pentru profesori și îndrumători în vederea analizei eficacității în ceea ce privește construirea de cunoștințe și colaborarea participanților în cadrul discuției.

Mai mult, *ReaderBench* presupune furnizarea unei abordări comprehensive pentru evaluarea colaborării în cadrul unei conversații de tip chat cu mai mulți participanți, ca rafinare a evaluării inițiale a colaborării definită din prisma câștigului informațional [9, 28]. Modelul implementat în cadrul sistemului *ReaderBench* identifică în mod automat zonele de colaborare intensă și extrapolează conceptele de construire de cunoștințe personale și sociale [5, 23, 24] către o granularitate mai rafinată, în mod specific, la nivelul fiecărei intervenții în cadrul conversației. În acest context specific, coeziunea reprezintă legătura dintre diverse intervenții, permițând transferul de informații și construirea cumulativă de cunoștințe, atât din punct de vedere personal (intervenții ale aceluiași vorbitor), cât și social (intervenții aparținând unor vorbitori diferiți). În plus, modelul nostru de analiză înaintea în direcția reliefării modelului polifonic al

discuției [30, 31], luând, de asemenea, în considerare, o analogie cu conceptele cheie din dialogism [1, 3]. Pentru a furniza o echivalență rapidă ce va fi detaliată în continuare, topicurile conversației care aparțin diferiților interlocutori sunt considerate voci, în timp ce coeziunea induce efectul de ecou al fiecărei intervenții raportat la intervențiile ulterioare din cadrul discuției.

Următoarea secțiune este o prezentare a modelului nostru de analiză, concentrându-se pe construirea unui graf de coeziune și pe definirea unui mecanism de notare ierarhic. În continuare, secțiunea de evaluare individuală se axează pe identificarea topicurilor discuției și pe modelarea interacțiunii participanților în cadrul conversației. Ulterior, evaluarea și evoluția colaborării devin elementele centrale ale analizei. În cele din urmă, vom prezenta un experiment de validare preliminară și vom dezbate relevanța rezultatelor.

ANALIZA CONVERSAȚIEI BAZATE PE COEZIUNE

Coeziunea, văzută ca relație la nivel lexical și semantic dintre două segmente de text, este definită în cadrul modelului nostru în raport cu: 1/ inversul distanței dintre intervențiile participanților analizate în cadrul unui interval fix ce definește contextul local, 2/ proximitatea lexicală reflectată în identitatea lemelor și în distanțele semantice bazate pe ontologii [7] și 3/ proximitatea semantică identificată pe baza analizei semantice latente (LSA – Latent Semantic Analysis) și a alocării latente Dirichlet (LDA – Latent Dirichlet Allocation) [8].

Ulterior calculării funcției de coeziune dintre diferitele unități de analiză în cadrul discursului, următorul pas constă în construirea automată a grafului de coeziune, care joacă un rol central în cadrul analizei noastre [8, 11]. Graful multistratificat include următoarea ierarhie: document (conversația integrală de tip chat sau materialul

pentru lectură) > bloc (intervenția participantului la discuție sau paragraful din cadrul textului inițial) > propoziție. Graful anterior este construit utilizând următoarele categorii de relații: 1/ relații ierarhice stabilite prin funcțiile de incluziune (propoziții în cadrul unui bloc, blocuri în cadrul unui document) și 2/ relații obligatorii obținute din: 2.a/ legături explicite, definite de către participanți în cadrul interfeței cu utilizatorul – în cazul nostru, ConcertChat [16] sau 2.b/ blocuri adiacente de text. Mai târziu, coeziunea este determinată între toate propozițiile din fiecare bloc (coeziune intra-bloc) și între blocurile din cadrul unei ferestre glisante de analiză (măsurători ale coeziunii inter-bloc care în cadrul experimentelor noastre au luat în considerare un interval contiguu de 20 de intervenții). Legăturile rezultate au fost filtrate după o valoare de prag impusă, setată experimental în cadrul experimentelor noastre la suma dintre deviația standard și media valorilor coeziunii, pentru a păstra în final numai legăturile relevante din cadrul discursului per ansamblu.

Scopul nostru a fost acela de a furniza un model generalizat și personalizat de evaluare bazată pe coeziune a diferitelor tipuri de discurs: conversații chat cu mai mulți participanți, pe de o parte, și texte în general, pe de altă parte. În ceea ce privește conversațiile CSCL, am optat pentru perspectiva lui Dong [12] de separare a intervențiilor pe baza evenimentelor de schimbare a vorbitorilor în cadrul conversației. Cu toate că majoritatea intervențiilor participanților constau dintr-o singură propoziție, frecvent eliptică, am preferat să creăm un model extensibil, care să poată fi ajustat ulterior cu ușurință în vederea analizei de forumuri de discuții în cadrul cărora intervențiile sunt constituite din multiple propoziții.

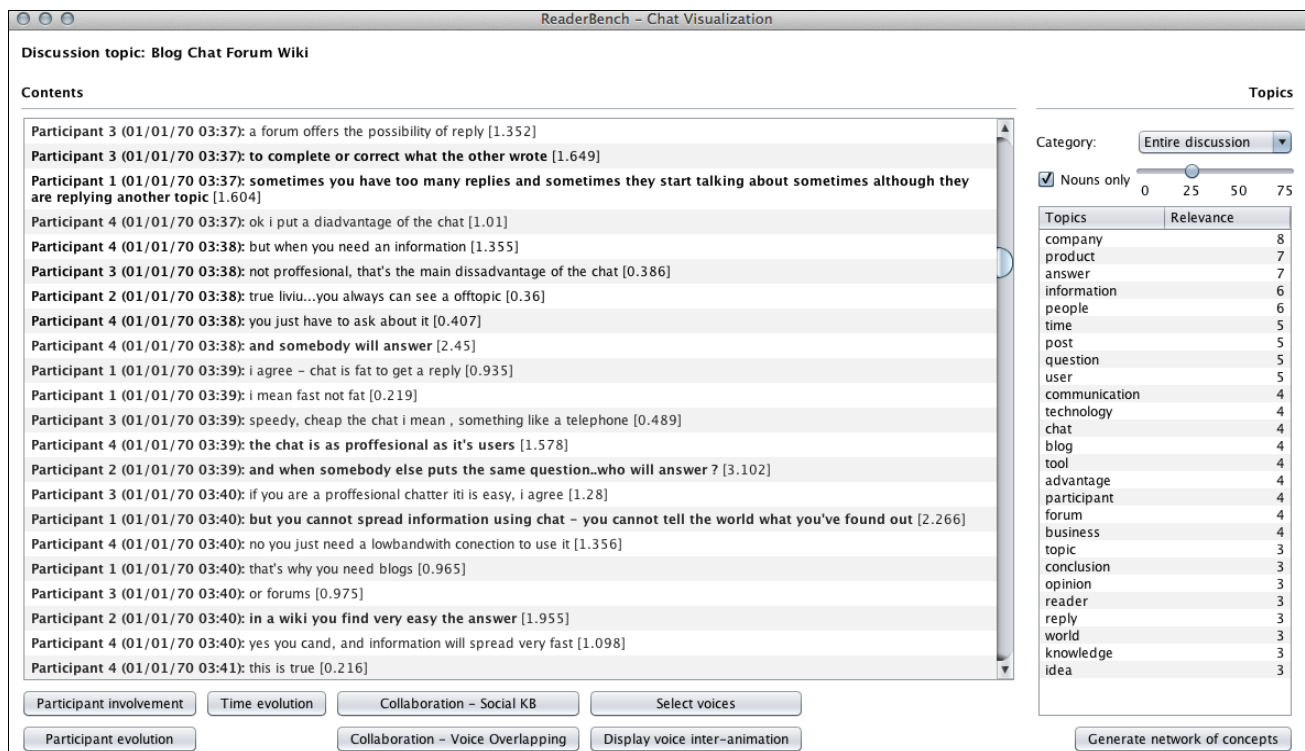


Figura 1. Interfața principală a sistemului ReaderBench pentru analiza conversațiilor de tip chat

O componentă importantă în procesul de evaluare a implicării fiecărui participant și a colaborării din cadrul conversației este reprezentată de mecanismul de notare bottom-up care asignează un scor fiecărei element de analiză (propoziție > intervenție > conversație). În primul rând, tehnici specifice de prelucrare a limbajului natural sunt aplicate fiecărei intervenții: selecția cuvintelor existente în dicționar, stemming, lematizare, recunoașterea entităților numite, adnotarea cu părți de vorbire, parsarea și rezoluția co-referențială [19]. Ulterior, fiecărei propoziții îi este atribuit un scor individual care reflectă acoperirea topicelor conversației și relevanța globală a propoziției raportat la discuția per ansamblu. Cu alte cuvinte, măsurăm prin coeziune gradul în care fiecare participant vizează fiecare concept, respectiv dacă este menținut focusul conversației, în special în cazul în care profesorul a definit un set de topicuri care trebuie să fie dezbătute.

În al doilea rând, la nivelul intervenției sau al blocului de text, scorurile individuale ale propoziției sunt ponderate prin măsuri de coeziune și însumate pentru a defini scorul intra-bloc. Raportat la modelul de descompunere a discursului (document > bloc > propoziție), valorile coeziunii inter-bloc sunt utilizate pentru a spori scorurile intra-bloc, luând în considerare toate legăturile din graful de coeziune definit anterior. În final, toate scorurile blocurilor sunt combinate pentru a stabili scorul total al conversației. Figura 1 identifică scorurile fiecărei intervenții între paranteze, imediat după prezentarea conținutului intervenției propriu-zise aferentă fiecărui participant. Conversațiile chat sunt încărcate dintr-un fișier XML exportat din ConcertChat [16], scorurile fiecărei intervenții fiind afișate între paranteze, iar topicurile identificate în mod automat pentru un anumit participant fiind prezentate în coloana din dreapta. Mecanismul propus de determinare a importanței elementelor de analiză este extensibil, întrucât ia în considerare importanța locală a fiecărei intervenții și combină aceste scoruri individuale prin intermediul coeziunii cu alte

elemente relaționate din punct de vedere semantic, din cadrul discursului.

EVALUAREA INDIVIDUALĂ A PARTICIPANȚILOR

Obiectivul de evaluare individuală a participanților presupune reliefaarea acoperirii topicurilor discuției și modelarea interacțiunii. Topicurile, percepute drept concepte-cheie ale conversației, joacă un rol marcant în obținerea unei viziuni de ansamblu asupra întregii conversații, dar și în observarea fiecărui punct de interes al participanților. Strâns legate de graful de coeziune și de mecanismul de măsurare anterior, topicurile sunt extrase per conversație și per participant. Relevanța fiecărui concept menționat în cadrul discuției este determinată prin combinarea unei multitudini de factori: 1/ frecvența normalizată a cuvântului atât în cadrul documentului, cât și în cadrul unui corpus de antrenare (Tf-Idf – term frequency-inverse document frequency) [19], 2/ similaritatea semantică bazată pe LSA și LDA cu discuția integrală, 3/ scorurile fiecărei propoziții, respectiv fiecărei intervenții în care este prezent conceptul în cauză. În cadrul acestei evaluări, procesarea inițială bazată pe tehnicile de prelucrare a limbajului natural, cu accent pe lematizare, a jucat un rol important în standardizarea intrărilor. Mai mult, scorurile fiecărui concept sunt normalizate raportat la lungimea lanțului lexical din care conceptul face parte. Lanțurile lexice anterioare sunt identificate din graful de dezambiguizare [14] modelat prin distanțele semantice [7] din WordNet [21] și WOLF (WordNet Libre du Francais) [22]. În plus, selectarea exclusiv a substantivelor din cadrul listei de topicuri a furnizat rezultate mai relevante în majoritatea cazurilor, întrucât substantivele tind să acopere mai bine conceptualizarea discuției.

În acest context, conceptele-cheie, alături de lanțurile lexice corespunzătoare pot fi considerate voci care se răsfrâng asupra conversației, în timp ce coeziunea simulează ecouri ale vocilor către alte intervenții inter-relaționate. Mai mult, prin schimbarea atenției asupra unui anumit participant, putem observa forța unei voci, direct proporțională cu relevanța topicurilor identificate anterior.

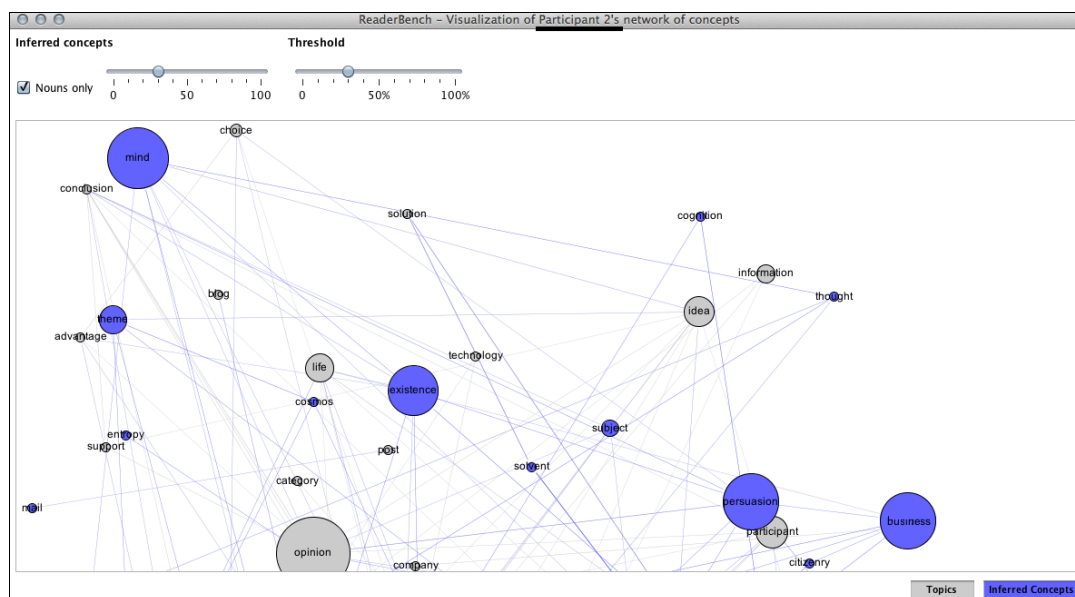


Figura 2. Vizualizarea conceptelor din spațiul semantic aferent unui utilizator

În urma tranziției de la o viziune globală a discursului către o perspectivă axată pe utilizator, componenta de vizualizare a spațiului conceptual pentru fiecare participant la chat, prezentată sub forma unui graf conex generat pe baza similarității semantice LSA și LDA dintre cuvinte, surprinde conceptele cheie utilizate de fiecare membru al discuției. Totodată, am îmbogățit topicurile identificate anterior cu concepte deduse, care nu sunt menționate în cadrul conversației, pe baza a trei distanțe semantice: 1/ similaritatea cu vectorul întregii discuții, 2/ similaritatea cu restul topicurilor, ponderate cu relevanța lor în cadrul conversației și 3/ similaritatea cu cel mai apropiat concept din cadrul discursului per ansamblu, întrucât corelarea strânsă cu un concept deja prezent sporește proximitatea cu discuția integrală. Aceste trei dimensiuni furnizează o argumentare empirică a metodei în care experții au selectat conceptele deduse din conversațiile evaluate manual: similaritatea cu un concept sau cu o idee din text, dar și cu întregul text (întregul text este reprezentat ca vector al documentului și ca extrapolare a topicurilor deja identificate).

Figura 2 prezintă spațiul de conceptualizare aferent unui utilizator, incluzând topicurile inițiale și conceptele deduse într-o manieră automată. Din punct de vedere computațional, stabilim toate similaritățile între cuvintele anterioare, considerate noduri în cadrul grafului, și selectăm drept arce numai pe legăturile semantice care se află deasupra pragului selectat de către utilizator în cadrul interfeței grafice (30% în acest caz particular). Dimensiunea reală a fiecărui nod este direct proporțională cu parametrii specifici analizei rețelelor sociale (SNA – Social Network Analysis) derivați din graful anterior (ex. betweenness centrality [6]). Toate măsurătorile și vizualizările multidimensionale ale grafului anterior sunt efectuate cu ajutorul platformei Gephi [4].

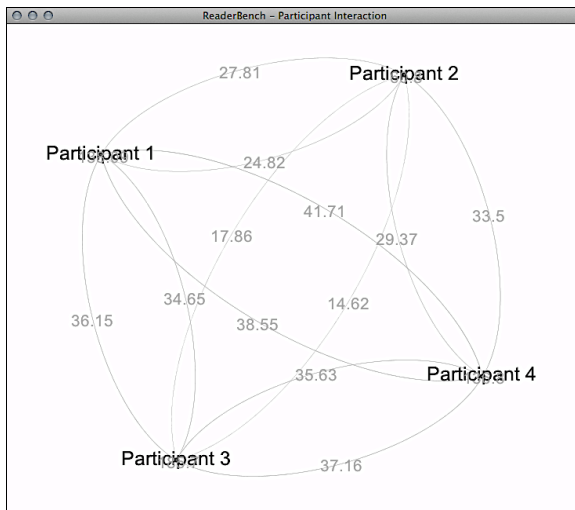


Figura 3. Vizualizarea grafului de interacțiune

Pe lângă identificarea topicurilor fiecărui participant la discuție, *ReaderBench* sprijină, de asemenea, modelarea interacțiunii participanților, acoperind două dimensiuni principale: o dimensiune preliminară cantitativă, generată de numărul real de interacțiuni per participant, și o dimensiune mai profundă calitativă, obținută după însumarea scorurilor intervențiilor aferente. Intern, un graf de interacțiune este construit având participanții drept

noduri, iar tăria legăturilor fiind egală cu suma scorurilor intervențiilor multiplicată prin funcția de coeziune cu elementul de analiză vizat, valoare desprinsă din graful de coeziune. Prin urmare, prin aplicarea de metrice specifice SNA asupra grafului de interacțiune a participanților, scara analizei este mutată către o perspectivă individuală (a se vedea Figura 3).

EVALUAREA COLABORĂRII

Pentru a obține o colaborare veritabilă între participanți, care poate fi percepută drept o polifonie reală conform teoriei lui Bakhtin [2] cu privire la dialogism, conversația trebuie să conțină o întretesere densă de voci emise de către diverși participanți [29]. Mai mult, transferul real de informații prin coeziunea dintre unitățile de analiză obține două valențe prin estimarea cantitativă a procesului de construcție a cunoștințelor sociale și personale [5, 23, 24] la nivelul intervenției și prin folosirea coeziunii ca legătură între intervențiile văzute drept suprapunere a vocilor. În primul rând, apare o dimensiune personală, prin luarea în considerare a intervențiilor aceluiași vorbitor, prin urmare, continuarea propriului discurs explicit. În al doilea rând, intervențiile interschimbate cu vorbitori distincți definesc o perspectivă socială care modelează colaborarea sub forma unui efect cumulativ.

Astfel, fiecare intervenție deține scorul definit anterior de importanță și un efect de construire a cunoștințelor, atât din punct de vedere personal, cât și social. Efectul personal este inițiat ca scor al intervenției, în timp ce efectul social este nul. Mai târziu, luând în considerare toate legăturile din graful de coeziune, fiecare dimensiune este amplificată corespunzător: dacă legătura este între intervențiile aceluiași vorbitor, cunoștințele construite anterior (atât din punct de vedere personal, cât și social) din intervenția respectivă sunt transferate prin intermediul funcției de coeziune către dimensiunea personală a intervenției curente; în caz contrar, dacă perechea de intervenții este între participanți diferiți, dimensiunea socială a construcției cunoștințelor aparținând intervenției analizate este amplificată cu aceeași cantitate de informație (scorul și efectul de construcție a cunoștințelor aferente intervenției anterioare, multiplicat cu măsura de coeziune). Cu alte cuvinte, continuarea ideilor sau a intervențiilor anterioare ale aceluiași vorbitor construiesc un dialog interior sau cunoștințe personale, în timp ce perspectiva socială măsoară interacțiunea cu alți participanți, încurajează schimbul de idei, dezvoltând creativitatea de grup [27] și influențând punctele de vedere ale altor participanți pe parcursul discuției, definind astfel bazele unei discuții cu adevărat colaborative. În acest context, colaborarea poate fi cuantificată prin însumarea efectelor de construcție a cunoștințelor sociale, pe baza scorului fiecărei intervenții coroborat cu funcția de coeziune.

Cu toate acestea, vom lua în considerare limitările modelului implementat în ceea ce privește construcția cunoștințelor personale. Colaborarea reiese în mod clar din transferul social de cunoștințe prin coeziune, ca influență a intervenției asupra discursului altui participant. Spre deosebire de aceasta, aproximarea construcției cunoștințelor personale reprezintă, mai degrabă, un prag

superior al transferului de informație dintre intervențiile personale. Asemănător abordării bazate pe câștig [9, 28], utilizăm o aproximare măsurabilă a dialogului interior, deși limitată în ceea ce privește procesele cognitive ale participantului. Construcția cunoștințelor personale este considerată o reflectare a propriilor gânduri exprimate în mod explicit în cadrul conversației derulate, sub forma unor legături coezive între intervențiile aceluiași participant la discuție. Însă acest fenomen nu induce neapărat construcția de cunoștințe personale, ci numai un discurs coeziv. Prin urmare, putem considera că valoarea numerică a construcției cunoștințelor personale reprezintă valoarea maximă a efectului explicit al construcției cunoștințelor personale, modelată pe parcursul discursului prin legături coezive.

Alături de estimarea efectelor construcției cunoștințelor personale și sociale pentru fiecare intervenție, *ReaderBench* identifică în mod automat zonele de colaborare intensă (a se vedea Figura 3) ce reprezintă intervale de intervenții în care participanții sunt implicați activ, colaborează și generează idei noi legate de contextul neîntrerupt al discuției. Primul pas în cadrul algoritmului

nostru de detecție a zonelor de colaborare intensă constă în identificarea vârfurilor de construcție a cunoștințelor sociale, ca valori maxime locale. Ulterior, fiecare vârf este extins lateral într-un areal predefinit, stabilit în mod experimental, la 2,5% din numărul de intervenții din întreaga conversație. Definirea acestui areal a fost importantă datorită focusului pe analiza colaborării la nivel macroscopic și din cauza posibilei existențe a unui amestec de multiple fire de discuție. În final, numai zonele cu minim 5 intervenții sunt selectate drept zone de colaborare intensă. Cu alte cuvinte, după identificarea intervențiilor cu cel mai important efect de construcție a cunoștințelor sociale, algoritmul extind fiecare zonă în stânga și în dreapta, în mod ne-suprapus raportat la zonele identificate anterior, luând în considerare numai intervențiile aflate deasupra valorii medii a construirii cunoștințelor sociale și care se află în arealul de proximitate definit anterior. Dintr-un punct de vedere distinct și strâns legat de construirea cunoștințelor sociale, coeziunea unește intervențiile în cadrul unei zone de colaborare intensă prin înrudirea topicelor abordate.

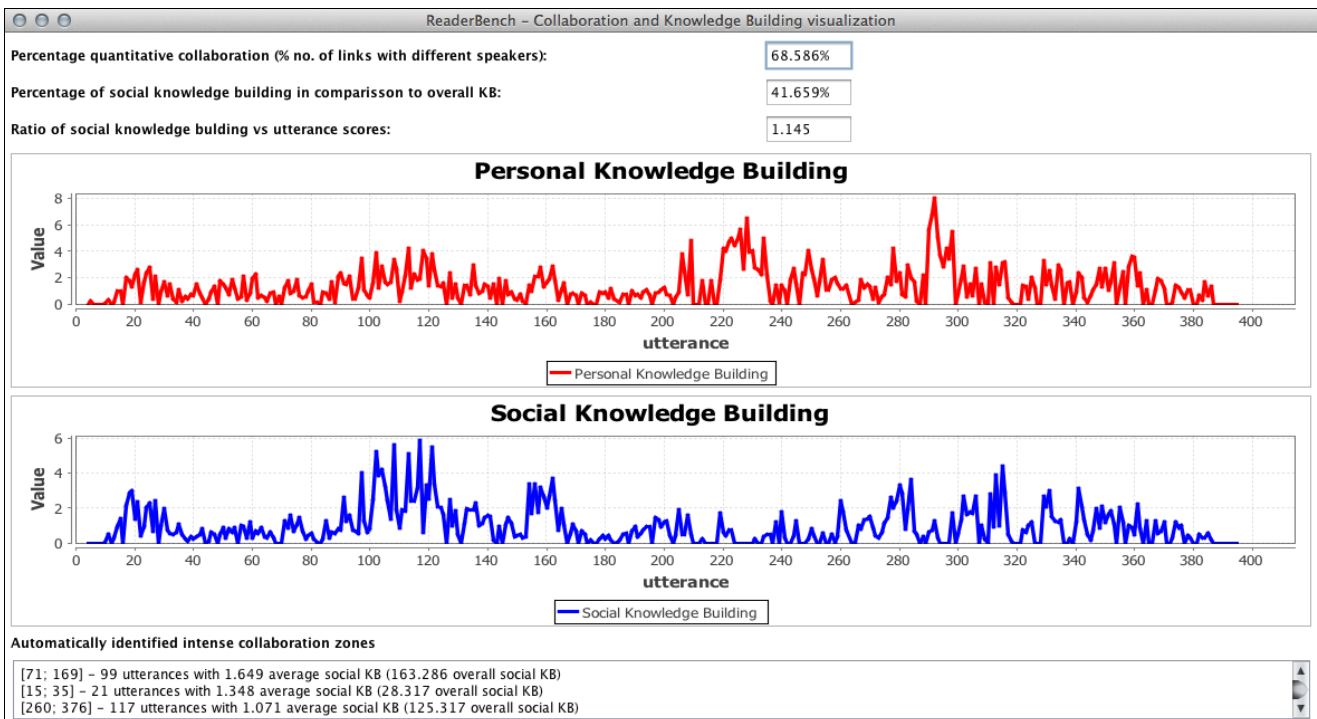


Figura 4. Evaluarea colaborării și evoluția acesteia în timp

La nivel global, au fost implementați trei factori pentru a caracteriza cel mai bine colaborarea generală în cadrul discuției. În primul rând, colaborarea cantitativă este stabilită ca procent al legăturilor dintre grafurile de coeziune cu vorbitori diferiți în comparație cu numărul legăturilor identificate în mod automat. Cu toate că este o estimare brută, această metrică asigură o viziune clară cu privire la schimbul efectiv de informații dintre participanți. În al doilea rând, scorul total al construcției de cunoștințe sociale este comparat cu efectul total al construcției de cunoștințe (social și personal). În al treilea rând, raportul dintre scorul total al construcției de cunoștințe sociale și scorurile totale ale importanței intervenției este determinat

pentru a reliefa cantitatea de informații transferate prin colaborare, în comparație cu ceea ce a fost prezentat inițial în cadrul fiecărei intervenții.

PRIMUL STUDIU DE VALIDARE

Experimentele preliminare au fost administrate pentru a valida modelul dialogic utilizat pentru evaluarea conversațiilor de tip chat, cu accent pe implicarea participanților și evaluarea colaborării. Trei conversații de tip chat derulate cu studenți din anul 4 de la cursul de Interacțiune Om-Calculator, centrate pe dezbaterea tehnologiilor colaborative, au fost evaluate manual de 4 tutori. În mod mai precis, fiecare student a trebuit să se

axeze pe o tehnologie (chat, wiki, blog sau forum), să prezinte și să dezbată beneficiile acestora pe baza unor scenarii specifice. Aceste trei conversații (Echipa 4, Echipa 34 și Echipa 36) au fost selectate pentru o analiză detaliată în urma unei priviri de ansamblu asupra a aproximativ 50 de discuții ce au implicat peste 200 de studenți. Cu toate că au fost observate discrepanțe mari în termeni de calitate raportată la conținut, implicare și colaborare a participanților, aceste conversații au fost considerate reprezentative pentru întreg eșantionul, iar evaluările preliminare au fost elaborate numai pe baza acestor conversații, din cauza duratei îndelungate implicate de evaluarea manuală a unei singure conversații pe chat care se poate extinde până la de două ori durata conversației purtate de către participanți [26].

În plus, interfața de modelare a evoluției în timp descrisă în Figura 5 a fost creată pentru a facilita evaluarea manuală a conversațiilor prin prisma zonelor de colaborare intensă. Astfel, graficul generat diferențiază intervențiile pe

diverși interlocutori în cadrul conversației și prezintă cronologic implicarea fiecărui participant.

În acest context, prezentarea conversației urmează derularea discursului și modelează întreteserea intervențiilor între diverși participanți, pe baza grafului de coeziune. Această componentă este utilă pentru identificarea manuală a: 1/ pauzelor din timpul conversației, a zonelor cu colaborare limitată sau nulă, din cauza faptului că în cadrul unui interval specific există un monolog al unui participant, fără nicio intervenție din partea altor utilizatori, și 2/ zonelor cu colaborare intensă, datorită inter-animării dense a intervențiilor dintre diferiți participanți. În cazul particular prezentate în Figura 5, toate intervențiile cu identificatori între 30 și 50 aparțin unui singur utilizator, într-un interval de timp limitat, reducând astfel efectul construirii cunoștințelor sociale la zero. Ulterior, pe măsură ce participanți suplimentari se implică în discuție, colaborarea devine mai puternică.

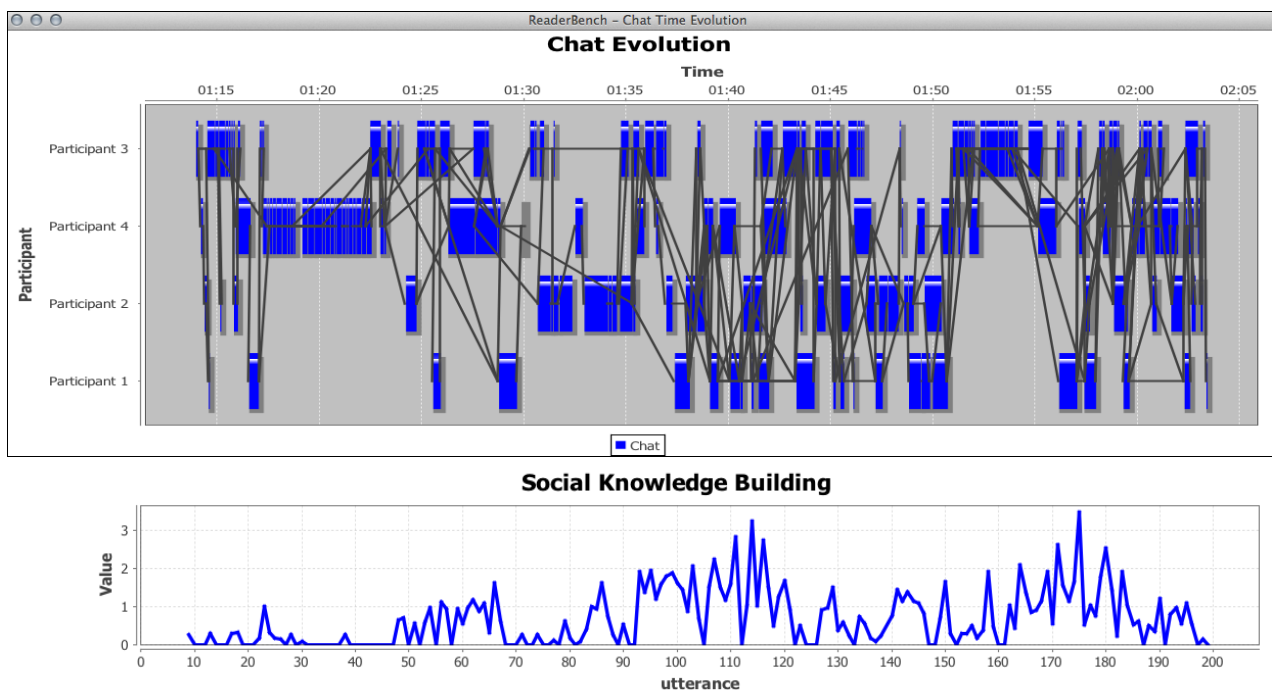


Figura 5. Intervalul de timp al unei conversații ce subliniază legăturile de coeziune și monologul din cadrul conversației.

Tabelul 1 prezintă corelația dintre diferiți factori de evaluare extrași din *ReaderBench* și notele finale acordate de către experți. Cu toate că identificatorii participanților coincid, la fiecare conversație au participat studenți distincți. Mai mult, pentru a asigura caracterul echitabil al analizei, corelarea dintre factorii stabiliți în mod automat de sistem și valorile medii ale notelor atribuite manual de către experți au fost determinată prin considerarea scorurilor participanților din toate conversațiile.

Ca interpretare a rezultatelor prezentate în Tabelul 1, putem observa o discrepanță în conversația echipei 4: deși implicarea participanților a fost satisfăcătoare, colaborarea actuală pe durata conversației este foarte dezechilibrată. Conversația echipei 34 prezintă cele mai scăzute scoruri raportat la toți factorii, în timp ce conversația echipei 36, considerată cea mai bună de către tutori, atât din punct de vedere calitativ, cât din prisma colaborării, a înregistrat cele mai bune scoruri atribuite de către sistem.

Din analiza corelării fiecărui factor este clar că tutorii au pus accent pe calitatea intervențiilor și nu pe numărul propriu-zis al acestor. Având în vedere că acordul dintre experți măsurat prin Intraclass Correlation Coefficient (ICC) a fost de .61 la măsurătorile individuale și de .86 în medie, scorurile de importanță cumulată de intervenții, respectiv de construcție a cunoștințelor sociale, se corelează extrem de bine cu notele medii ale experților. În plus, dimensiunea construcției de cunoștințe personale se corelează mai bine cu notele experților; acest lucru susține observația empirică conform căreia efectul personal reflectat în implicarea fiecărui participant are un impact mai mare în evaluarea experților decât dimensiunea socială, prin care se obține colaborarea. Cu alte cuvinte, este mai ușor de înțeles participarea activă decât interacțiunea socială din perspectiva construcției de cunoștințe prin colaborare.

Tabelul 1: Corelarea dintre evaluările manuale și automate ale participanților [11]

Numele Participantului	Numărul de intervenții	Scorul cumulat al intervențiilor	Construcția de cunoștințe personală	Construcția de cunoștințe socială	Notă expert				
					1	2	3	4	Medie
Echipa 4									
Participantul 1	90	215.34	262.19	27.63	9.0	10.0	8.0	9.5	9.13
Participantul 2	61	99.78	114.91	15.39	7.0	7.0	8.0	6.5	7.13
Participantul 3	120	220.92	248.99	12.76	8.0	9.0	8.5	10.0	8.88
Participantul 4	118	205.79	239.96	29.19	8.0	6.5	9.0	8.0	7.88
Echipa 34									
Participantul 1	23	60.13	77.91	11.42	6.0	6.0	7.0	6.0	6.25
Participantul 2	34	64.17	72.58	6.66	5.0	7.0	7.0	6.0	6.25
Participantul 3	73	120.03	139.51	13.22	8.0	7.0	8.0	7.0	7.50
Participantul 4	60	124.89	141.12	4.55	7.0	4.0	7.0	5.5	5.88
Echipa 36									
Participantul 1	54	160.89	200.18	37.95	9.0	-	9.5	8.0	8.83
Participantul 2	67	201.57	251.98	49.46	10.0	-	8.0	10.0	9.33
Participantul 3	119	327.45	412.78	58.83	9.0	-	8.0	8.5	8.50
Participantul 4	57	161.39	200.60	40.33	9.0	-	9.5	8.0	8.83
Corelația medie	.64	.83	.84	.77					

În ceea ce privește zonele intense de colaborare, adnotările manuale și zonele identificate automat sunt prezentate în Tabelul 2. Prin analiza tuturor rezultatelor anterioare, în afară de factorii de evaluare globală corelați cu evaluările umane, au fost observați următorii indicatori de slabă colaborare (majoritatea în conversația echipei 34): numărul crescut de zone identificate în mod automat conținând 1-3 intervenții care nu au fost considerate în final zone de colaborare intensă, valoarea medie scăzută a efectelor de construcție a cunoștințelor care nu au generat în mod automat o zonă de colaborare cu răspândire largă (peste 50 de intervenții, deși a fost cea mai scurtă

conversație din cele trei). În schimb, conversațiile cu colaborare activă (și anume, echipa 36 care a înregistrat cea mai bună colaborare) au mai multe zone de colaborare echilibrate, valori medii mai mari a dimensiunii de construcție a cunoștințelor și o distribuție și acoperire mai bună a întregii conversații.

Sunt necesare investigații ulterioare pe un eșantion extins de conversații, însă luând în considerare subiectivitatea sarcinilor trasate, rezultatele preliminare sunt promițătoare și demonstrează că modelul dialogic implementat de evaluare atât a implicării participanților cât și a colaborării asigură măsuri relevante.

Tabelul 2: Suprapunerea dintre identificarea manuală și automată a zonelor intense de colaborare [11]

Conversația	Număr de intervenții	Zonele de colaborare identificate manual	Zonele de colaborare identificate automat
Echipa 4	389	[90; 160], [320; 360]	[19; 43], [72; 146], [157; 166], [240; 280], [290; 317], [329; 379]
Echipa 34	190	[90; 120], [170; 178]	[51; 67], [74; 78], [94; 121], [127; 131], [158; 166], [181; 195]
Echipa 36	297	Distribuție uniformă	[38; 58], [69; 122], [136; 157], [202; 221], [286; 290]

CONCLUZII

Sistemul *ReaderBench* a fost proiectat pentru a sprijini profesorii în diferitele scenarii educaționale care implică utilizarea chat-urilor sau a forumurilor pentru rezolvarea diferitor sarcini, cu accent pe rezolvarea colaborativă a problemelor sau pe dezbateri. Pornind de la rezultatele primului studiu de validare, considerăm că *ReaderBench* poate fi folosit pentru sprijinirea proceselor consumatoare de timp vizavi de evaluarea manuală a conversațiilor chat.

Global, plecând de la un model dialogic al discursului, sistemul de analiză integrată poate fi utilizat pentru identificarea lipsei de coeziune dintre intervenții, pentru analiza implicării participanților și pentru evaluarea colaborării individuale și globale, prin procesul de

construire socială a cunoștințelor din mediile CSCL. Extrapolând, prin combinarea celor două percepții diferite asupra coeziunii, participanții la conversații CSCL pot utiliza sistemul *ReaderBench* pentru a evalua gradul în care coeziunea dintre intervenții reflectă coeziunea grupului drept rezultat al colaborării reflectate la nivelul grafului de interacțiune.

MULȚUMIRI

Cercetările prezentate au fost susținute de către proiectul FP7 264207 ERIC-Empowering Romanian Research on Intelligent Information Technologies și de către proiectul POSDRU/107/1.5/S/76909 Valorificarea capitalului uman din cercetare prin burse doctorale (ValueDoc).

REFERINȚE

1. Bakhtin, M.M., 1981. The dialogic imagination: Four essays. The University of Texas Press, Austin and London.
2. Bakhtin, M.M., 1984. Problems of Dostoevsky's poetics. University of Minnesota Press, Minneapolis.
3. Bakhtin, M.M., 1986. Speech genres and other late essays. University of Texas, Austin.
4. Bastian, M., Heymann, S., and Jacomy, M., 2009. Gephi: an open source software for exploring and manipulating networks. In International AAAI Conference on Weblogs and Social Media AAAI Press, San Jose, CA, 361–362.
5. Bereiter, C., 2002. Education and Mind in the Knowledge Age. Lawrence Erlbaum Associates, Mahwah, NJ.
6. Brandes, U., 2001. A Faster Algorithm for Betweenness Centrality. *Journal of Mathematical Sociology* 25, 2, 163–177.
7. Budanitsky, A. and Hirst, G., 2006. Evaluating WordNet-based Measures of Lexical Semantic Relatedness. *Computational Linguistics* 32, 1, 13–47.
8. Dascalu, M., Dessus, P., Trausan-Matu, S., Bianco, M., and Nardy, A., in press. ReaderBench, an Environment for Analyzing Text Complexity and Reading Strategies. In 16th Int. Conf. on Artificial Intelligence in Education (AIED 2013) Springer, Memphis, USA.
9. Dascalu, M., Rebedea, T., and Trausan-Matu, S., 2010. A deep insight in chat analysis: Collaboration, evolution and evaluation, summarization and search. In 14th Int. Conf. on Artificial Intelligence: Methodology, Systems, Applications (AIMSA 2010), D. Dochev and D. Dicheva Eds. Springer, Varna, Bulgaria, 191–200.
10. Dascalu, M., Trausan-Matu, S., and Dessus, P., 2010. Utterances assessment in chat conversations. *Research in Computing Science* 46, 323–334.
11. Dascalu, M., Trausan-Matu, S., and Dessus, P., in press. Cohesion-based Analysis of CSCL Conversations: Holistic and Individual Perspectives. In 10th Int. Conf. on Computer-Supported Collaborative Learning (CSCL 2013) ISLS, Madison, USA.
12. Dong, A., 2005. The latent semantic approach to studying design team communication. *Design Studies* 26, 5, 445–461.
13. Fujita, N. and Teplov, C., 2009. Automating the analysis of collaborative discourse: Identifying idea clusters. In 9th Int. Conf. on Computer supported collaborative learning (CSCL'09), C. O'Malley, D. Suthers, P. Reimann and A. Dimitracopoulou Eds. ISLS, Rhodes, 162–164.
14. Galley, M. and McKeown, K., 2003. Improving Word Sense Disambiguation in Lexical Chaining. In IJCAI'03, G. Gottlob and T. Walsh Eds. Morgan Kaufmann Publishers, Inc., Acapulco, Mexico, 1486–1488.
15. Graesser, A.C., McNamara, D.S., Louwerse, M.M., and Cai, Z., 2004. Coh-Metrix: Analysis of text on cohesion and language. *Behavioral Research Methods, Instruments, and Computers* 36, 2, 193–202.
16. Holmer, T., Kienle, A., and Wessner, M., 2006. Explicit Referencing in Learning Chats: Needs and Acceptance. In Innovative Approaches for Learning and Knowledge Sharing, EC-TEL 2006, W. Nejdl and K. Tochtermann Eds. Springer, Crete, Greece, 170–184.
17. Landauer, T.K. and Dumais, S.T., 1997. A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review* 104, 2, 211–240.
18. Landauer, T.K., Foltz, P.W., and Laham, D., 1998. An introduction to Latent Semantic Analysis. *Discourse Processes* 25, 2/3, 259–284.
19. Manning, C.D. and Schütze, H., 1999. Foundations of statistical Natural Language Processing. MIT Press, Cambridge, MA.
20. McNamara, D.S., Kintsch, E., Butler Songer, N., and Kintsch, W., 1996. Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction* 14, 1, 1–43.
21. Miller, G.A., 1995. WordNet: A Lexical Database for English. *Communications of the ACM* 38, 11, 39–41.
22. Sagot, B., 2008. WordNet Libre du Francais (WOLF) INRIA, Paris.
23. Scardamalia, M., 2002. Collective cognitive responsibility for the advancement of knowledge. In Liberal Education in a Knowledge Society, B. Smith and C. Bereiter Eds. Open Court Publishing, Chicago, 67–98.
24. Stahl, G., 2006. Group cognition. Computer support for building collaborative knowledge. MIT Press, Cambridge, MA.
25. Tapiero, I., 2007. Situation models and levels of coherence. Erlbaum, Mahwah, NJ.
26. Trausan-Matu, S., 2010. Automatic Support for the Analysis of Online Collaborative Learning Chat Conversations. In 3rd Int. Conf. on Hybrid Learning, P.M. Tsang, S.K.S. Cheung, V.S.K. Lee and R. Huang Eds. Springer, Beijing, China, 383–394.
27. Trausan-Matu, S., 2010. Computer Support for Creativity in Small Groups using chats. *Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information* 3, 2, 81–90.
28. Trausan-Matu, S., Dascalu, M., and Rebedea, T., 2012. A system for the automatic analysis of Computer-Supported Collaborative Learning chats. In 12th IEEE Int. Conf. on Advanced Learning Technologies (ICALT 2012), C. Giovannella, D.G. Sampson and I. Aedo Eds. IEEE, Rome, Italy, 95–99.
29. Trausan-Matu, S. and Rebedea, T., 2009. Polyphonic Inter-Animation of Voices in VMT. In Studying Virtual Math Teams, G. Stahl Ed. Springer, Boston, MA, 451–473.
30. Trausan-Matu, S., Rebedea, T., Dragan, A., and Alexandru, C., 2007. Visualisation of learners' contributions in chat conversations. In Blended learning, J. Fong and F.L. Wang Eds. Pearson/Prentice Hall, Singapur, 217–226.
31. Trausan-Matu, S., Stahl, G., and Zemel, A., 2005. Polyphonic Inter-animation in Collaborative Problem Solving Chats. Drexel University