World Transactions on Engineering and Technology Education
Vol.3, No.1, 2004

A complexity analysis of collaborative turn-taking patterns that evolve during computer-mediated collaboration

Sofia J. Hadjileontiadou† & Leontios J. Hadjileontiadis‡

Demokritos University of Thrace, Xanthi, Greece†
Aristotle University of Thessaloniki, Thessaloniki, Greece‡

ABSTRACT: Peer-to-peer computer-mediated collaboration evolves by means of peers' successive submission of contributions, resulting in collaborative turn-taking sequences. Pattern identification in these sequences and their possible changes under a different set-up of collaboration may provide useful information regarding peers' collaborative activities. In this article, the authors describe the mapping of peers' collaborative contributions for turn-taking sequences, which, in turn, are transformed to symbol-sequences and analysed for pattern extraction. The use of a normalised complexity measure, when applied to collaborative data from the field of environmental engineering, reveals peers' tendencies to more complex turn-taking patterns when they receive appropriate feedback. The efficiency and simplicity of the proposed complexity analysis in pattern identification makes it useful for tracking non-stationarities in peers' collaborative activity.

INTRODUCTION

Virtual collaboration develops on the basis of a turn-taking sequence of collaborative computer-mediated interactions. Specific conventions that govern the turn-taking process are explicitly differentiated between synchronous and asynchronous modes of communication. In the first case, these conventions need to ensure that only one person speaks at a time until the completion of a turn, ie a continuous and complete unit of speech by a single speaker during the conversation [1].

However, in the second case, the collaboration is carried out mainly in textual form and the turn-taking process refers mostly to the sequential submission of elaborated units of information, namely collaborative *contributions*. In this case, the possible options of turn-taking between two collaborators, dubbed A and B, following a submission of A, includes a resubmission or submission of the next contribution by A or B respectively. Under this perspective, quite complex patterns of *collaborative turn-taking* may evolve, with the [A-B-A-B-…] being the most deterministic one.

An analysis of these patterns at a level of abstraction may reveal useful information for the realisation of the collaborative procedure. The calculation of an index of the complexity of the collaborative turn-taking patterns is used as an indicator of its systematicity [2]. According to this approach, when the index of complexity equals zero, it reflects the certainty that the next peer makes the next submission, whereas when it equals one, no information concerning the next turn-taking option is available, ie there is an almost non-systematicity of collaborative turn-taking [2].

In this article, the Lempel-Ziv complexity analysis for the identification of the evolving collaborative turn-taking patterns during asynchronous, computer-mediated collaboration is proposed [3]. Given this information, useful conclusions can be drawn concerning the collaborative attitudes of the peers, in conjunction with the structure of the collaboration and the provision, or lack thereof, of feedback, as suggested by the presented empirical findings.

METHODOLOGY

An increase in the number of different patterns seen in a time series results in a direct increase in its corresponding complexity. To this end, by adopting an appropriate complexity measure, the development of temporal activity patterns in the turn-taking process during peer-to-peer collaboration can be efficiently identified and characterised. So far, many complexity measures, mainly drawn from non-linear dynamics theory, have been reported in the literature. They are used to capture the *dynamic* complexity of a time series and the most known include approximate entropy, neural complexity, and *KL*-complexity [4-6]. However, most of these measures usually lack effective computational methods for their implementation. Fortunately, the Lempel-Ziv complexity measure, $C(n)$, is an alternative complexity measure that circumvents any implementation limitations, since it is simpler to understand and easier to be implemented [3]. The Lempel-Ziv complexity measure was adopted here as a tool for the turn-taking sequence analysis within a computer-mediated collaborative environment (CMCE).

Collaborative Turn-Taking Sequence Acquisition

The collaboration of two peers, ie A and B, within a CMCE allows the acquisition of a collaborative turn-taking sequence by means of logging the submission of the collaborative contribution of each peer. In particular, in this case, the Lin2k Web-based CMCE was used as a bed-set to collect the collaborative interactions of the two peers A and B [7][8]. As already thoroughly described in Ref.s [7][8], Lin2k includes successive parts of collaboration, namely steps (s), and supports the experiential learning of *proper* collaboration towards a balanced quality of peer collaborative activity in order to produce the best dynamic of the pair under consideration. In this regard, regulation is challenged through a follow-up of each peer's interactions, an evaluation of his/her collaborative activity and feedback provision. Peer communication is facilitated by individual workspaces that are semi-structured, ie possess predefined areas (eg buttons, frames) dedicated to editing and submission of certain types of contribution. The latter include *proposal* (*P*), *contra-proposal* (*CP*), *comment* (*CM*), *clarification* (*CL*), *agreement* (*AG*), *low (LQ)* and *high (HQ) level question* [7][8].

Since these collaborative interactions are automatically linked to each peer (A or B), their time sequence reflects the collaborative turn-taking sequence during the peers' collaboration activity. Consequently, by simply corresponding a vector of positive integer values to the collaborative interactions of A (eg [P→1; CP→2; CM→3; CL→4; AG→5; LQ→6; HQ→7]), and the symmetric negative one to the collaborative interactions of B, a collaborative turn-taking sequence can be constructed. An example of this procedure is shown in Figure 1, which clearly denotes the mapping of the alteration in the collaborative interactions between the two peers A and B to the collaborative turn-taking sequence.

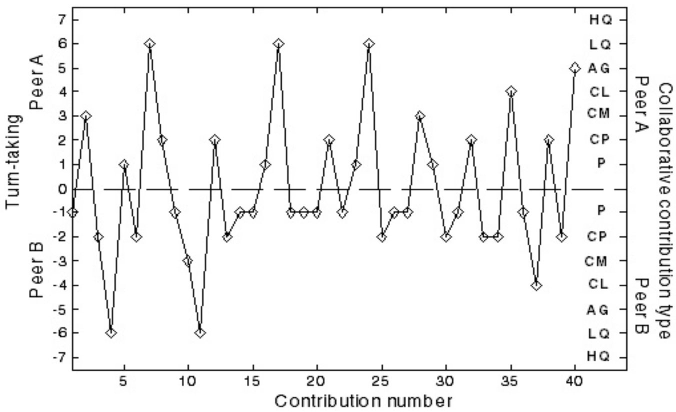

Figure 1: Construction of the collaborative turn-taking sequence from peers' collaborative contributions.

Lempel-Ziv Complexity Analysis

The adopted Lempel-Ziv complexity analysis is based on a coarse graining of the turn-taking sequence, ie its transformation into a sequence whose elements are only a few symbols [3]. By employing a complexity counter, $cc(n)$, the number of the distinct patterns contained in the constructed symbol-sequence, ie $P = s_1, s_2, ..., s_n$, (where s_1, s_2, etc, denote characters, for example A or B), is measured by scanning the P sequence from left to right and increasing by one the $cc(n)$ every time a new subsequence of consecutive characters is encountered in the scanning process; this process is described below.

Let K and L denote, respectively, the subsequence of the sequence $P = s_1, s_2, ..., s_n$, and let KL be the concatenation of K and L. Further, let $KL\neg$ denote the sequence derived from KL after its last character is deleted ($\neg$ corresponds to the operation to delete the last character in the sequence) and $v(KL\neg)$ denotes the vocabulary of all the different subsequences of $KL\neg$. Initially, $cc(n)$=1, $K = s_1$, $L = s_2$ and $KL\neg = s_1$. In a general case, $K = s_1, s_2, ...s_r$, and $L = s_{r+1}$; hence, $KL\neg = s_1, s_2, ..., s_r$; If $L \in v(KL\neg)$, then L is a subsequence of $KL\neg$, thus not a new sequence. The K subsequence has not changed, so the L subsequence is renewed to $L = s_{r+1}, s_{r+2}$; it is then examined if $L \in v(KL\neg)$ and the procedure is continued until $L \notin v(KL\neg)$; in that case, (index=i) $L = s_{r+1}, s_{r+2}, ..., s_{r+i}$ and it is not a subsequence of $KL\neg = s_1, s_2, ..., s_r, s_{r+1}, ..., s_{r+i-1}$. Consequently, the complexity counter is increased by one, ie $cc(n)=cc(n)+1$. Thereafter, K is combined with L and K is renewed to be $K = s_1, s_2, ..., s_r, s_{r+1}, ..., s_{r+i}$; accordingly $L = s_{r+i+1}$. This procedure is repeated until L is the last character of the sequence. At this time, the total number of different subsequences is stored in the complexity counter $cc(n)$ As it is clear from the described procedure, the Lempel-Ziv algorithm employs only two simple operations (comparison and accumulation). Consequently, the implementation of the computation of $cc(n)$ is highly facilitated. The derived $cc(n)$ complexity counter depends on the sequence length n. To this end, a normalised complexity measure, $C(n)$, was used instead that was independent of the sequence length. In Ref. [3], it is shown that the upper bound of $cc(n)$ is given by the following:

$$cc(n) < \frac{n}{(1-\varepsilon_n)\log_a(n)} , \tag{1}$$

where, ε_n is a small quantity and $\varepsilon_n \to 0$ for $n \to \infty$, and the base a of the logarithm denotes the number of different symbols in the sequence. Consequently, by applying limit normalisation for a=2 (A, B), and n sufficient high, the normalised complexity measure is given by the following:

$$C(n) = \frac{cc(n)}{\lim_{|n\to\infty} cc(n)} = \frac{cc(n)}{n/\log_2(n)} \tag{2}$$

The above complexity measure reflects the resultant rate of new patterns, along with the sequence, and, due to (2), it is usually less than one [3].

IMPLEMENTATION ISSUES

Practical Considerations

The symbol-sequence P, upon which the complexity analysis is applied, is constructed from the acquired turn-taking sequence (see Figure 1) after employing a threshold of zero-value (see dashed line in Figure 1) to separate between the alterations of the two symbols A and B. Consequently, the symbol-sequence P, which corresponds to the turn-taking sequence of the example in Figure 1, becomes: {P=BABBABAABBBABB BAABBBABAABBBAABBABBABBABA}, and the Lempel-Ziv complexity analysis identifies the following different subsequences (* denotes the end of each different subsequent) {B*A*BB*ABA*ABBB*ABBBAA*BBBABA*ABBBAAB* BBABBABA*} with $cc(n)$=9. As already stated, limit normalisation in (2) assumes sufficient high sequence length n.

However, it is quite often the case that, after a session of collaborations, there are not so many collaborative interactions. This is also the case, especially when the collaborative process is split into subsequent steps (as in the case of Lin2k), where, unless the collaboration is very productive, the expected number of contributions is quite small.

In order to circumvent possible bias in the estimation of the normalised complexity measure $C(n)$, an interpolation procedure is employed that extends the data length without destroying the morphology of the time series [9]. However, the appropriate extended length should be exploited in such a way that the estimated $C(n)$ attains stability. Figures 2(a) and (b) depict the curve of $C(n)$ against the length n obtained for different interpolation lengths, when a collaborative turn-taking sequence acquired from the whole collaboration session (contributions from s=1:6) and from one step (s=1) were used, respectively. As it is clear from Figures 2(a) and (b) that the adoption of the interpolation length of $n_{se} = 300$ and $n_{st} = 400$ samples, respectively, leads to stable estimations of $C(n)$, since over these lengths, the fluctuation in the estimated $C(n)$ values is negligible. This was also true for all available data; hence, these lengths were adopted for all of the calculations of $C(n)$ throughout this study.

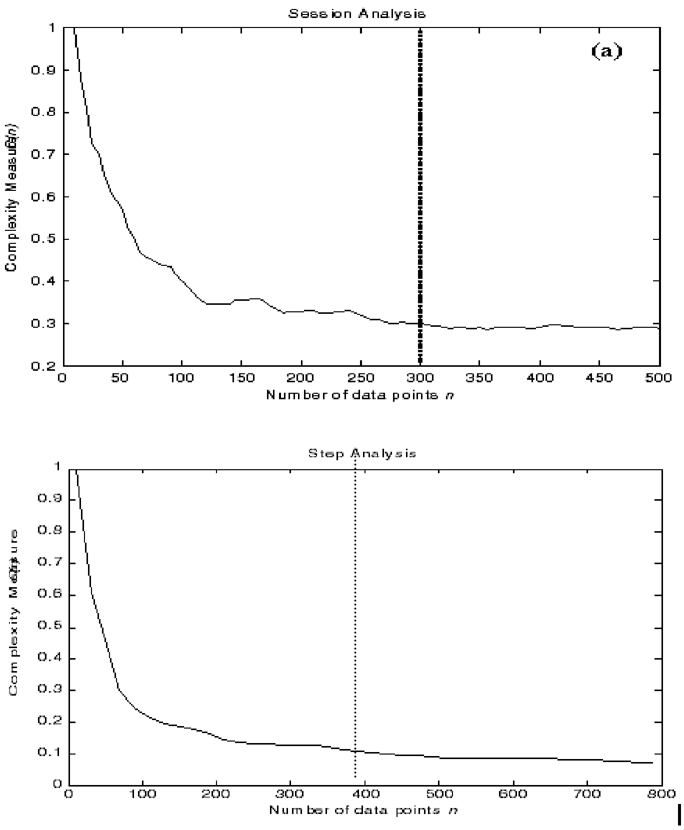

Figure 2: The estimated $C(n)$ against the interpolation length n for a turn-taking sequence acquired (a) at the end of a session of collaboration (all steps) (top), and (b) at the end of one step of collaboration (bottom).

Lin2k Feedback and Experimental Dataset

An individual feedback in terms of an encouraging or a warning message is foreseen in the CMCE Lin2k at the end of each step [7][8]. In particular, this feedback aggregates the values of the two variables employed to quantify the parameters of the collaborative activity. The first variable refers to the evaluation of individual performance compared to the total activity at the pair's level, whereas the second one refers to each peers' metacognitive awareness of the quality of his/her collaborative performance and his/her planning of improvement at the next step. The provision of feedback allowed for the experimental use of Lin2k is with a different set-up, ie with feedback (FD) or without feedback (NFD) provision [7].

The adopted experimental dataset, drawn from Ref. [7], refers to the NFD collaborative activity of ten pairs from the 6th semester and the FD collaborative activity of eight pairs of students of the 10th semester, all randomly selected from the Department of Civil Engineering at Aristotle University of Thessaloniki, Thessaloniki, Greece. All students collaborated using Lin2k [7][8]. This was achieved through a six-step (s=1:6) collaboration session on a case study from the environmental engineering field. The whole analysis was developed using *Matlab* 6.5 (Mathworks, Inc.).

RESULTS AND DISCUSSION

Figures 3(a) and (b) present the mean values per step of the estimated $C(n)$, $MC(n)$, for the two student-pair groups, ie, FD and NFD, respectively. The grey area denotes the estimated standard deviation (*std*) per step across the student pairs.

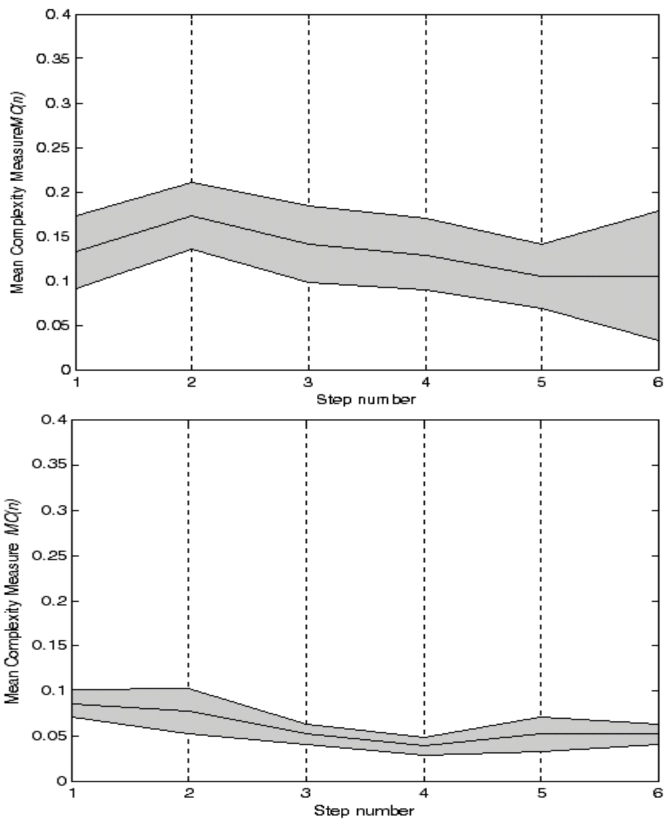

Figure 3: The mean value per step of the estimated $C(n)$, $MC(n)$, corresponding to turn-taking sequences from (a) FD (top), and (b) NFD student-pair groups (bottom).

In comparing these results, a shift towards a more simple turn-taking process is noted when no feedback is provided. Moreover, a statistical analysis (nonparametric Kruskal Wallis test, using SPSS 11.0) showed significant statistical difference (p<0.05) among all estimated $C(n)$ values per step derived from the FD and NFD cases (p values per step: 0.031; 0.004; 0.004; 0.003; 0.009; 0.010). The trend for increased complexity in the turn-taking patterns seen in the FD case could be explained as peers reacting to the feedback provision, since they are motivated to elaborate and get more deeply involved within the collaborative task. Consequently, they react in a way that produces less periodic zero-crossings in the turn-taking sequence, shifting towards less deterministic collaborative behaviour. The latter fires on more fruitful procedures, as far as the collaborative activity is concerned, fostering the non-stationary characteristics of the collaboration and triggering new initiatives in a less strict collaborative context. This assumption is further justified by the statistical analysis performed across the steps of collaboration.

In particular, the estimated $C(n)$ values per step derived from the FD and NFD cases were statistically compared separately, to identify possible significant difference across the steps of collaboration for each case (FD and NFD). Results from this analysis are presented in Table 1, alongside the task aim at each step of collaboration (*s*). As can be deduced from Table 1, from all of the examined data, the transition from the first to the second step for the case of FD only resulted in *C*(*n*) values with significant statistical difference ($p<0.05$). This explains the influence of the feedback to peers from the FD group, to shift from simple collaborative turn-taking patterns (like A-B-A-…) to more complex ones, in order to fulfil more demanding task aims in the subsequent steps. On the contrary, peers from the NFD group were *locked* within simple collaborative turn-taking patterns without any adjustment to the demands of the task aims across the steps of collaboration.

Table 1: Statistical analysis of the estimated *C*(*n*) values across the steps (*s*) of collaboration related to the task.

Steps (*s*)		1	2	3	4	5	6
Task aim		Under-standing of the situa-tion	Diag-nosis of the prob-lems	Produc-tion of alter-native solu-tions	Predic-tion of results of every solu-tion	Evalua-tion of alter-native solu-tions	Choice of best solu-tion upon reason-ing
Transi-tional *p**	FD	0.046		0.380		0.498	
			0.088		0.085		
	NFD	0.414		0.157		0.989	
			0.059		0.317		

*Using two-sample related nonparametric Wilcoxon test (SPSS 11.0). Statistically significant difference when $p<0.05$.

This is also clear from the differences seen in the variety of the identified patterns that emerged during the whole collaborative session of FD and NFD cases, as presented in Table 2. Moreover, Table 2 shows a significant difference between the FD and NFD cases, regarding the adoption of the turn-taking model, ie A*B or B*A; in the FD case the frequency of such simple model (seen mostly in the case of face-to-face communication) reaches 33.3% only, whereas in the NFD one is over 73%.

When analysing the whole session as one turn-taking sequence, the values of the estimated complexity measure *C*(*n*) are greater than those seen in the step-based analysis (both in FD and NFD cases), since the collaborative contributions are accumulated. In that case, the mean values (±*std*) of the estimated complexity measure *C*(*n*) for the FD and NFD cases are $MC(n)=0.5212\pm0.1684$ and $MC(n)=0.2853\pm0.0631$, respectively. Moreover, a statistical analysis (nonparametric Mann-Whitney U test, using *SPSS 11.0*) between the two groups across the student-pairs showed significant statistical difference ($p=0.006<0.05$) between the FD and NFD cases, justifying the active role of feedback in the complexity of the turn-taking patterns, not only at a micro-genetic level, but also at a macro-genetic one. Future research includes extension to the pattern categorisation according to the collaborative contribution type, ie by employing more threshold-levels in the symbol-sequence conversion, resulting in patterns like A_P*B_{CP}*A_PB_P*, etc.

CONCLUSIONS

A complexity-based analysis of the collaborative turn-taking sequences that emerge in a CMCE was presented in this article. Identification of the underlying patterns in peer turn-taking sequences by measuring the deviations in its complexity, and tracking any possible changes due to peers' feedback provision were explored. The results showed a clear transition to more complex turn-taking patterns (increased complexity) when appropriate feedback was provided to peers. As a result, a new way of monitoring peer role exchange during collaboration was introduced, which can be easily integrated within the context of intelligent mediator agent design.

Table 2: Type and frequency *F* (%) of the identified turn-taking patterns for the FD and NFD student-pair groups.

FD Student pair group		NFD Student pair group	
Pattern type	*F* (%)	Pattern type	*F* (%)
A*AAB*	2.1	A*AB*	3.3
A*AAB*ABABB*	2.1	A*AB*ABAA*	3.3
A*AB*	2.1	A*B*	40.0
A*AB*ABAA*BB*ABBAA*	2.1	A*B*BBA*	3.3
A*AB*ABB*ABBAA*	2.1	B*A*	33.3
A*AB*BA*ABA*	2.1	B*A*AB*	10.0
A*B*	20.8	B*BA*AB*	3.3
A*B*AA*BAB*	2.1	B*BBBA*	3.3
A*B*AA*BAB*ABB*	2.1		
A*B*ABABB*	4.2		
A*B*ABB*BAA*BABA*	2.1		
A*B*BA*	2.1		
A*B*BA*BAA*BBB*	2.1		
B*A*	12.5		
B*A*AB*	8.3		
B*A*AB*ABAA*	2.1		
B*A*AB*ABABABABABB*	2.1		
B*A*AB*BAB*	2.1		
B*A*BAA*	2.1		
B*A*BAA*BB*BABAB*	2.1		
B*A*BABB*	6.3		
B*A*BB*	2.1		
B*A*BB*ABBB*AA*	2.1		
B*BA*	2.1		
B*BA*AB*BAB*	2.1		
B*BA*BBAA*	2.1		
B*BA*BBABA*	2.1		
B*BBA*BBAA*AAAA*	2.1		

REFERENCES

1. Sacks, H., Schegloff, E. and Jefferson, G., A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, **4**, 696-735 (1974).
2. Dillenbourg, P., Jermann, P., Schneider, D., Traum, D. and Buiu, C., The design of MOO agents: Implications from an empirical CSCW study, http://tecfa.unige.ch/tecfa/publicat/dil-papers-/Dil.7.3.21.pdf
3. Lempel, A. and Ziv, J., On the complexity of finite sequences. *IEEE Trans. on Inform. Th.*, 22, 75-81 (1976).
4. Pincus, S.M., Approximate entropy (ApEn) as a complexity measure. *Chaos*, 5, 110-117 (1995).
5. Tononi, G. and Edelman, G.M., Consciousness and complexity. *Science*, 282, 1846-1851 (1998).
6. Klonowski, W., Jernajczyk, W., Niedzielska, K., Rydz, A. and Stepien, R., Quantitative measure of complexity of EEG signal dynamics. *Acta Neurobiologiae Experimentalis*, 59, 315-321 (1999).
7. Hadjileontiadou, S.J., Sakonidis, H.N. and Balafoutas, G.J., Lin2k: A novel Web-based collaborative tool-application to engineering education. *J. of Engng. Educ.*, 92, **4**, 313-324 (2003).
8. Hadjileontiadou, S.J., Nikolaidou, G.N., Hadjileontiadis, L.J. and Balafoutas, G.N., A fuzzy logic evaluating system to support web-based collaboration using collaborative and metacognitive data. *Proc. IEEE ICALT 2003*, Athens, Greece, 96-100 (2003).
9. *Matlab 6.5 Signal Processing Toolbox*. Natick: Mathworks (2002), www.mathworks.com