Using Latent Semantic Analysis to Assess Social Competence

Fridolin Wild and Christina Stahl, *Vienna University of Economics and Business Administration, Austria*

***Abstract* — Assessing social competence is a time consuming and complex endeavor, which has prompted several institutions to implement automated assessment systems. Although other attempts have been made, the majority of automated assessments are still based on multiple-choice formats. Within this contribution, the authors therefore present a new method of how to assess social competence using LSA. The contribution first reports on the research design, then presents the findings which are subsequently discussed in brief.**

I. Introduction

STRUCTURAL changes in the qualification requirements for employees have led many countries to adopt competence-based education schemes. In particular, the promotion of social competencies is considered an important aspect of such education (e.g., [1], [2]). For the purpose of this research, we define 'social competence' as involving abilities that facilitate communicative and cooperative action and that aim at identifying, managing and mastering conflicts.

The increasing adoption of social competence education also poses the challenge of assessment. Assessing learners' social competence is a time-consuming endeavor, which has prompted several institutions to implement automated assessment systems. Although attempts to assess social competence by means of graph-based approaches [3], simulations [4], and natural language processing approaches [5] have been made, the majority of automated assessments is still based on multiple-choice formats.

We therefore present a new method of social competence assessment using Latent Semantic Analysis (LSA). The research design, results and a brief discussion are presented in the following sections.

II. Research Design

For our research we used a corpus of 337 textual contributions produced by students in an online discussion forum in the course of a university seminar. The contributions were split into sentences resulting in 1,012 individual messages. These messages were manually coded along ten dimensions of social competence (politeness, ability to motivate others, phatic communication, ability to express own opinion, cooperation competence, team competence, feedback competence, networking competence, ability to take initiative, and readiness to take on responsibility), assigning the code '1' when evidence of a dimension was found in a message, and'0' when no such evidence was identified. The dimension 'ability to motivate others' was found in only 37 messages and was therefore omitted from further analysis.

The goal of our analysis was to mimic the human coding with LSA. For this purpose the corpus coded along the remaining nine dimensions of social competence was split into a training corpus consisting of 490 messages and a test corpus of 522 messages. The training corpus was used to calculate the LSA space using the share-method proposed in [6] for space reduction. No document pre-processing (e.g., stemming) was performed.

From the 522 test messages, 16 had to be omitted from the LSA analysis as they did not include any terms from the LSA space. Each of the remaining 506 test messages was folded into the LSA space and Pearson's r was calculated to compare the test message with all training messages. For each test message the ten most highly correlated training messages were identified, which – for each dimension separately – were grouped according to their human coding ('1' or '0'). Each dimension of the test document was assigned the code that received the higher weighted sum within this comparison set (see [7] for a similar method).

III. Results

For seven of the nine remaining dimensions of social competence we received good to excellent results. For the dimension 'politeness', LSA was able to correctly predict human coding for 452 test messages (89.33%) out of 506. For the other dimensions, the respective results were 437 (86.36%) correctly coded messages for the dimension 'team competence', 430 (84.98%) correctly coded messages for the dimension 'cooperation competence', 432 (85.38%) correctly coded messages for the dimension 'networking competence', 393 (77.67%) correctly coded messages for the dimension 'ability to take initiative', 392 (77.47%) correctly coded

F. Wild and C. Stahl are with the Institute for Information Systems and New Media of the Vienna University of Economics and Business Administration, Augasse 2-6, A-1090 Vienna, Austria (fridolin.wild@wu-wien.ac.at and christina.stahl@wu-wien.ac.at).

messages for the dimension ‘readiness to take on responsibility’, and 320 (63.24%) correctly coded messages for the dimension ‘feedback competence’. The other two dimensions – ‘emphatic communication’ and ‘ability to express own opinion’ – produced inferior results of 229 (45.26%) and 164 (32.41%) correctly coded messages, respectively.

We also compared our results to the plain vector-space model, and surprisingly received similar results. In particular, the vector-space model performed better for three of the nine dimensions, although the results differed only by a maximum of 0.59%.

IV. Discussion

We have been able to successfully apply LSA to assess social competence in contributions to online discussion forums. LSA was able to mimic the human coding process for seven of the nine analyzed dimensions of social competence. Concerning the other two dimensions, we assume that the language used to express these dimensions is too diverse to be captured. Surprisingly we found similar results when comparing LSA to the Vector-Space-Model. Further research and discussion will have to investigate this issue, as well as the applicability of LSA to the assessment of social competence in particular, and the automation of human coding in general.

References

[1] EUROPEAN COMMISSION (2002): The Key Competencies in a Knowledge-Based Economy: A First Step Towards Selection, Definition and Description. A proposal by the Working Group on key competencies, set up by the European Commission in the framework of the ‘Objectives Report’.

[2] RYCHEN, D. S. & SALGANIK, L. H. (2003): Introduction. In: D. S. RYCHEN & L. H. SALGANIK (Eds.): Key Competencies for a Successful Life and a Well-Functioning Society. Hogrefe & Huber, Cambridge, pp. 1-12.

[3] FISHER, D., SMITH, M. & WELSER; H. T. (2006): You Are Who You Talk To: Detecting Roles in Usenet Newsgroups. In: R. H. SPRAGUE (Ed.): Proceedings of the 39th Annual Hawaii International Conference on System Sciences. IEEE Computer Society, Los Alamitos.

[4] JOHNSON, W. & BEAL, C. (2005): Iterative Evaluation of a Large-Scale, Intelligent Game for Language Learning. In: C.-K. LOOI, G. MCCALLA, B. BREDEWEG & J. BREUKER (Eds.): Artificial Intelligence in Education. IOS Press, Amsterdam, pp. 290-297.

[5] SHAW, E. (2005): Assessing and Scaffolding Collaborative Learning in Online Discussions. In: C.-K. LOOI, G. MCCALLA, B. BREDEWEG & J. BREUKER (Eds.): Artificial Intelligence in Education. IOS Press, Amsterdam, pp. 587-594.

[6] WILD, F., STAHL, C., STERMSEK, G. & NEUMANN, G. (2005): Parameters Driving Effectiveness of Automated Essay Scoring with LSA. In: Proceedings of the 9th International Computer Assisted Assessment Conference (CAA). Loughborough, pp. 485-494.

[7] LANDAUER, T. K., LAHAM, D. & FOLTZ, P. W. (2003): Automated Scoring and Annotation of Essays with the Intelligent Essay Assessor. In: M. D. SHERMIS & J. C. BURSTEIN (Eds.): Automated Essay Scoring. A Cross-Disciplinary Perspective. Lawrence Erlbaum Associates, Mahwah, pp. 87-112.