NON-INVASIVE DATA TRACKING IN EDUCATIONAL GAMES: COMBINATION OF LOGFILES AND NATURAL LANGUAGE PROCESSING

Stephanie B. Linek

University of Graz
Graz / Austria
stephanie.linek@uni-graz.at

Georg Öttl

Graz University of Technology
Graz / Austria
georg.oettl@tugraz.at

Dietrich Albert

Graz University of Technology
Graz / Austria
dietrich.albert@tugraz.at

Abstract

Within educational games one essential problem is how to gather user information without destroying the game play. Even though, there exist several possibilities of tracking the user behaviour, most of the objective behavioural data are ambiguous and lack subjective meaning. However, the assessment of subjective data by, e.g., questionnaires either destroys the game play or could be only made when the game has finished. The proposed methodological framework illustrates how the combination of data tracking by logfiles and Natural Language Processing (NLP) could lead to a more precise interpretation of user behaviour which in turn enables an optimization of adaptive user support in the context of serious games.

Keywords - Game-based learning, non-invasive data tracking, logfiles, Natural Language Processing, methodology.

1 INTRODUCTION: CHALLENGES OF DATA TRACKING IN EDUCATIONAL GAMES

Educational games are strongly dependent on an adaptive design which in turn requires an ongoing monitoring of player, not only his/her learning progress but also his/her actual socio-affective and physiologiocal state since the joy in gaming is essential for game-based learning. In contrast to conventional e-learning environments the possibilities of data-tracking are limited in the sense that the measurement of user behaviour has to be accomplished without interrupting the ongoing learning/gaming process to avoid destroying the overall game-play and flow-experience.

Thus, non-invasive methods like logfile-protocols or Natural Language Processing should be used. However, such non-invasive methods normally measure only pure behavioural data without subjective meaning. Accordingly, the assessed behavioural data are often ambiguous and thereby useless for appropriate interventions and adaptations. However, a more sophisticated combined approach could reduce ambiguousness and thus enables the possibility to an adequate intervention/adaptation during the gaming process in order to enhance learning as well as enjoyment.

Besides the requirement for an ongoing monitoring of the players behaviour in order to enhance the game play, there is also a requirement for an online data-tracking to evaluate the overall quality of the educational game. Evaluation of e-learning (and serious games) is often accomplished by questionnaires. However, questionnaires should only be presented before (to assess user-data) or after an educational game (to assess the gaming experience) to avoid the disturbance of the game play. Thus, questionnaires on the game-experience deliver only retrospective data. Additionally, questionnaires can be considered as reactive methods with several other drawbacks. Also in this case the combination with other non-invasive methods can provide a more holistic view of the gaming process and might deliver a surplus meaning which in turn enables a more sophisticated evaluation and accordingly recommendations for further development. The fruitful combination of questionnaires and logfiles was described in [1]. The following paper goes a step further by combining two non-invasive methods and thus, become more independent from invasive, reactive methods like questionnaires that are of only limited use in the context of games.

In the following, we will discuss essential aspects for the game play and the enjoyment of the game and how these aspects are (partly) depend on a non-invasive ongoing monitoring of the player. Afterwards, two methods of non-invasive data tracking will be described that could be used within educational games. Both methods have their advantages and drawbacks. A common problem of these (and many others) assessment methods is, that they deliver partly ambiguous information and thus, the players' actual state remains unclear. In this article we will describe, how the combination of both data-sources could lead to a more sophisticated non-invasive data-tracking that reduce ambivalence of information and allows distinct judgement of the actual state of the player. An accordingly open framework will be outlined and suggestions for further combinations will be given.

2 CORE FACTORS FOR ENJOYMENT IN GAMES AND EDUCATIONAL GAMES

The core idea of educational games is using the motivational potential and the joy of conventional games for learning. When talking about the joy of gaming, the general terms enjoyment, immersion and flow are popular staring points. When taking about the concrete reasons for playing games it becomes obvious, that there are multiple aspects that attract players, for example active participation, parasocial interaction, fantasy, escapism, collaborative interaction, pleasure of control, and/or breaking social rules and conventions in a safety setting [2][3][4][5][6][6][7].

One the one hand, many of these aspects relate to the parasocial interaction during the game and other player-specific preferences and needs like fantasy or active participation. Parasocial interaction is accomplished by different dialog systems and also the overall game play often strongly depends on a clever interaction and communication-design with the other game characters. Furthermore, many of these aspects are rather general and could be accomplished in multiple divers instances within different games, e.g. active participation is part of role playing games (by killing monsters) as well as part of racing games (by driving a car). The concrete player-specific preferences are normally reflected in the selection of a specific game genre. However, most gamers play more than one game genre, and the concrete selection of a specific game and the game-episode per se strongly depends of the actual mood, i.e. games (like other media) are used for mood management. Besides the subjective preferences, the joy of gaming, especially the pleasure of control, strongly depends on an adaptive game design that neither bores nor overwhelms the player. In conventional games this is normally accomplished by the different game levels.

Since game-based learning tries to exploit the motivational potential of conventional video games, the question arises how to address these aspects in educational games: the player's need for a reasonable parasocial interaction, the player's individual preferences and the player's actual mood as well as the adaptivity on the player's capabilities.

On first view this seems to be rather easy and in some cases it is in fact simply to accomplish, e.g. to be adaptive on the players capabilities. However, on a second view one could easily recognize the challenges a game designer is confronted with in creating an educational game: The gaming context has not only to fit with the learning context but also have to serve the needs of different subgroups of players. One solution to this problem could be to adapt to different player preferences. This in turn requires an ongoing individual monitoring of the user/player without destroying game play by interrupting the gaming process. In other word, there is a need for a more sophisticated non-invasive

data-tracking that allows reacting immediately to the player's actual state (cognitive, socio-affective and physiological).

2.1 Adressing the player's need for reasonable parasocial interaction

When addressing the players need for social or parasocial interaction, one has to keep in mind that different player-subgroups have different communication preferences. In conventional games this fact is often reflected in the selection of a specific game genre. However, educational/serious games are normally designed and selected with respect to their learning contents regardless of the concrete preferences of the player for specific style of game genre or specific style of interaction. This fact might at least partly explain why educational games often fail and are subjectively judged as rather boring. Thus, if one wants to design an educational game that attracts different subgroups of players one has to be adaptive to the concrete communication preferences of the player. Thereby, the history of the parasocial interaction of the player could be helpful. For example, if the former dialogue with a specific NPC was rather negative and the player's behaviour provides evidence, that the NPC was perceived as boring or too old fashioned in its communication style, then the game should adapt by changing either the communication of the actual NPC or by introducing another NPC.

2.2 Adressing the player's individual preferences and actual mood

Games as well as other media are often used for mood-management. Mood-management relates to the so-called mood-management theory by Zillmann [8][9] which states that people use the media experiences for optimizing their emotional state and their arousal level, respectively. For example, if a person is under-stimulated (i.e. in a low level of arousal), he/she will choose an exciting action film in order to higher his/her arousal near to the optimal level. Thereby, media (experiences) are used for regulation of one's own mood (by means of optimizing the stimulus environment).

If a game designer wants to adapt to the actual mood of the gamer, this in turn requires monitoring of the actual mood. Normally, the actual mood strongly influences gaming and communication behavior and thus, also in this case, the recent activities and communication of the player provides evidence about his/her mood and the need for a change. Analogously, individual preferences could be identified by the actual behavior and communication of the player.

2.3 Adressing the players capabilities in educational games

Regarding the adaptivity on the players capabilities, the so-called flow theory [10][2] is of special interest for educational games since it regards not only to a motivational experience but also to an optimization of learning. The flow theory is a unified framework on subjective experience in the course of mastering everyday challenges. According to the flow theory, the quality of subjective experience is determined by the balance between the perceived challenges (intrinsic demands) on the one hand and the perceived skills (self-perceived capacity to meet demands) on the other hand. If challenges and skills are both at a low level, the person experiences apathy. If the perceived challenges are lower than the perceived skills, the person is bored. If the perceived challenges overrun the perceived skills, the person feels anxiety. If the perceived challenges as well as the perceived skills are simultaneously high, the person experiences flow or flow in consciousness, and this represents the highest quality of subjective experience. Thus, according to the flow theory the balance between challenges and skills is crucial for intrinsic motivation and enjoyment of the activity. In general, research supports this assumption [11][12][13].

The experience of flow or being immersed the activity, is intrinsically rewarding and can be seen as a hedonic state. This implies two dynamics (assumed that the person tries to reach this marvelous state): If the person is bored (skills > challenges) then he/she should seek new challenges. On the other hand, if the person is overloaded and anxious (skills < challenges) he/she should try to enhance his/her skills and knowledge, or in other words he/she will engage in learning activities.

Overall, the so-called flow-experience can be characterized by intense, focused concentration, clear goals, direct, immediate feedback (success and failures are apparent), balance between the perceived challenges and the perceived skills, feeling of control, distortion of the sense of time, loss of self-

consciousness (because of being immersed in the activity), and hedonic quality of the experience (intrinsically rewarding). Thereby the flow theory states that there are several conditions that facilitate the experience of flow: Structured activities with manageable rules and clear goals, the possibility to adjust the required actions in respect to one's own abilities, the provision of clear feedback, as well as the elimination of distraction and the provision of the possibility for concentrated work.

With respect to (educational) games it can be assumed that the challenges in the game should be adaptive with respect to the gamer's skills – not only in order to enhance learning but also to foster enjoyment of the gamer. The given characterization of the optimal conditions for experiencing flow underlies that a game has to react to the players progress during the game, i.e. to be adaptive on the players behaviour and provide him/her immediate feedback on his/her actual progress. This in turn requires an ongoing monitoring of the actual state of the player, i.e. not only the cognitive capabilities but also the learning/gaming progress that he/she has already managed.

The conception of flow theory is not completely new (this is also annotated e.g., by [12]). Similar theoretical approaches are e.g., the Yerkes-Dodson-law [14] [15]) or the theoretical conception of Berlyne [16] regarding stimulus complexity (see also [17] regarding intrinsically motivating instruction). Roughly spoken, the common denominator of these approaches is, that pleasure as well as learning is optimized in an medium level (of arousal, challenge, complexity) that neither bores nor overwhelmed the subject whereby the absolute level (what is "medium") depends on the person's individual abilities/characteristics. It is worth noting that these approaches partly regard to physiological, partly to affective and partly to cognitive aspects. Thus, it could be suggested that not only cognitive adaptivity but also socio-affective and physiological adaptivity will be crucial [5]. This threefold-conception of flow addresses not only the cognitive adaptivity on the available cognitive skills but also the players need for a reasonable parasocial interaction and mood-management via games.

2.4 Resume

To sum up, for enhancing the joy as well as the educational impact of educational games, it is essential to monitor the actual behaviour and communication of the player. This ongoing data-assessment provides the adequate basis for adapting on the players needs. Thereby the core challenge is an data-assessment without interrupting the gaming process. Thus, new methods of data-tracking are required. In the following two non-invasive methods will be described and their fruitful symbiosis of data will be outlined.

3 DATA-TRACKING BY LOGFILES

3.1 Logfile-information

Logfiles can be defined as automatically written protocols of actions and processes within a computer-based environment (e.g. sequencing of pages, clicking a button, duration …). The concrete actions and processes which are recorded are various. The assessed parameters are nearly unlimited and have to be defined in advance. The accordingly logfile is written in the background and the user is (normally) not aware of the data tracking. In principle, logfiles could be used in every computer-based environment including conventional e-learning methods as well as (educational) games.

3.2 Pros of Logfiles

The major advantage of data assessment by logfiles lies in the objective measurement of user behaviour. The method allows an immediate data-tracking without time delay. Moreover, the learning and/or playing process is not disrupted and thus, the gameplay and flow-experience is not tackled by data-assessment. Data-assessment by logfiles is independent from the compliance of the user and normally, the user is not even aware of the data-tracking. Thereby, this measurement method can be considered as non-reactive method without demand effects. Since various and multiple parameters could be defined, logfiles can serve as a quick and cost-efficient measurement of a huge amount of information.

3.3 Shortcomings and drawbacks of logfile information

Two of the main advantages of logfiles, the efficient measurement of a huge amount of information and the measurements of objective data, could be also considered as the two main drawbacks. On the one hand the possibility to measure a nearly unlimited number of variables can cause an information overload and sometimes additional analysis tools are needed. The core problem in this context is the definition and careful selection of appropriate behavioural indicators that provide the adequate basis for further interventions. The reasonable selection of appropriate behavioural indicators is very difficult, especially since logfiles comprises only pure objective data without subjective meaning. Even if the parameters are carefully designed, they are often ambiguous and different interpretations of the same behavioural data are possible.

Like pointed out elsewhere [1] the combination with questionnaire data might fill logfiles with psychological meaning. This could be useful for purposes of evaluation, however in the case of an ongoing monitoring and adaptation this approach can not be applied in (educational) games since questionnaires would destroy the game play during the game (and after the game its to late). Thus, there is the need for other non-invasive assessment method that might provide more information on the subjective meaning of the assessed behavioural data and reduce the ambiguousness of logfile information. One possibility could lie in the combination with NLP. Even though NLP-data are not always distinct, they deliver a rather different qualitative informational impact and thereby could be seen as a supplementary extension of logfile information.

4 DATA-TRACKING BY NATURAL LANGUAGE PROCESSING (NLP)

4.1 Characterization of Natural Language Processing

Historically, the work on Natural Language Processing (NLP) was inspired by the claim that there is a connection between the language an individual uses and its mind [18][19]. Based on this claim, the first algorithmic efforts in NLP aimed at an automatic full *understanding of the subjective meaning of natural language text*. This *understanding* of the subjective meaning is done by interpreting the text an individual produces, working autonomously once developed and deployed.

Last decades have shown that it is difficult to establish the connection between language and mind algorithmically and consequently a full understanding of natural language is not yet feasible. Therefore, recent NLP research has focused on smaller subfields of NLP that cover only aspects of natural language (e.g. Information Extraction, Word Sense Disambiguation or Sentiment Detection). Nowadays the term NLP is used as an umbrella term whenever there is an attempt to algorithmically analyse or generate natural language text.

4.2 Pros of NLP

NLP brings the same benefit as logfile analysis, namely additional information on what the individual is doing. NLP can be used to provide an automated abstraction of language concepts that goes beyond the counting of words. The fully automatic approach is especially suitable for applications scenarios we are interested in our work, namely at game scenarios where the user should not be disturbed during game play.

4.3 Shortcomings and drawbacks of NLP

The most efficient NLP methods [20][21][22] currently rely on supervised machine learning. On the one hand the supervised machine learning approach provides a good performance in means of precision and recall of the extracted meanings, but does also introduce a big drawback – the need of a fair amount of training data to be produced in advance to train classifiers. Although semi-supervised methods [23] lower the amount of training data needed, still some effort has to be put into the creation of training data.

Also, it is currently quite clear that the initial aim of NLP, the understanding of the meaning of natural language, will probably stay a hard to solve computational problem.

5 COMBINATION OF LOGFILE INFORMATION AND NLP

Data tracking by logfiles as well as NLP can be considered as non-reactive methods without demand effects. They are not dependent of the compliance of the player and in most cases the player is not even aware of the data-assessment.

Logfile-data delivers quantitative objective data regarding the activity of the user (e.g., assessed by click-rate) or his/her direction of behaviour (e.g., interaction with a specific NPC). Information gathered by NLP could enrich the quantitative logfile-data by qualitative information, e.g., if the interaction/dialogue with another player is friendly or hostile. Thus, logfile-data and NLP-analyses deliver different information about the player. The combination of both methods provides not only an extended view of the players but can also add a surplus meaning by reducing ambitiousness of the single data pools.

An example is the analysis of the interaction between two players: The pure activity level and interaction rate indicates if and how extensive the players are in contact. The NLP information could give evidence if the two players like or dislike each other. The combination of both information provides information if they have an intensive useful collaborative or a destructive quarrel. Accordingly, the combination of logfiles and NLP could also lead to different interventions of the system. If there are only a few negative comments there might be no need for an intervention (of the system). However in the case of a quarrel the system should lead the players to another scenario which might help to resolve the conflict. Contrariwise, in the case of positive comments the system should rather react if the interaction rate is to low.

This example illustrates also, that the rules for combining the information of logfiles and NLP as well as the accordingly adaptive functionalities of the system has to be carefully defined. The rules could be based on psychological and pedagogical theories and/or on existing research findings on games and game-based learning.

Analogous to the example described, the combination of logfiles and NLP could also lead to a more precise and sophisticated assessment of player's motivational state, the player's interest for a specific topic or tool, the (para)social interaction of a player or the shared interests of a subgroup of users within a multiplayer scenario.

The described methodological framework can be used for optimizing user-support and adaptive gaming not only in the context of serious games but also in other e-learning environments or pure entertaining video games. The described indices and cases are only exemplarily and are open for further extensions.

6 OPEN FRAMEWORK FOR A COMBINED ASSESSMENT METHODOLOGY

6.1 Overview

The following framework on a combined non-invasive assessment approach is conceptualized as an open framework that could be applied to different context of educational games as well as for conventional games or conventional e-learning environments. The core idea is to use the diversity of pure objective information from two non-invasive measurement methods to reduce the ambiguousness of both isolated sources. Thus, the combined approach analyses the intersections, the differences and similarities between the two data sources in order to come to distinct and more holistic information about the player/user and the ongoing learning/gaming process. Fig. 1 provides an overview on the architecture of the framework.

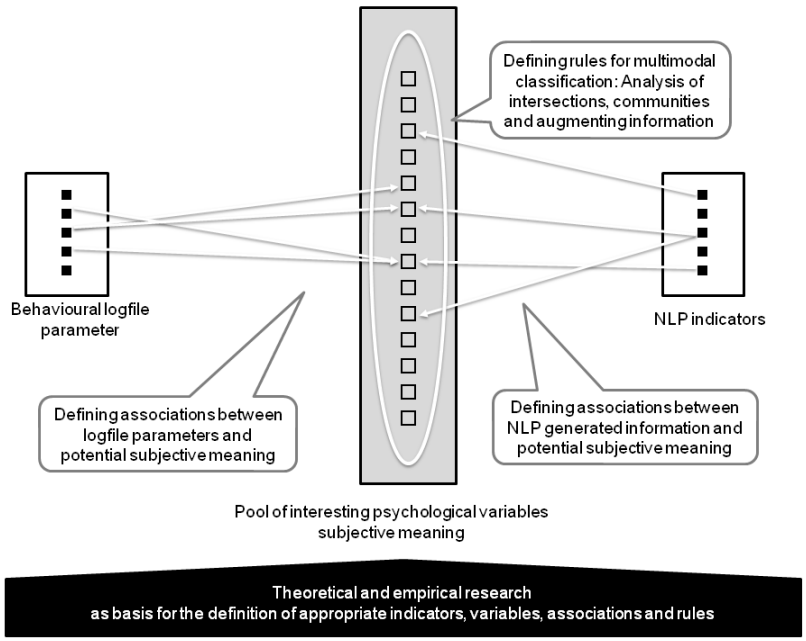

Fig.1: Overview on the Framework for combining logfiles and NLP information

Fig. 1 comprises different steps or phases which will be explained in more detail in the subsequent subchapters which reflect also the general workflow of the framework.

6.2 Defining the interesting psychological variables

First of all, the interesting subjective psychological variables have to be defined. These indicators strongly depend on the aim and the circumstances of the educational game (or another environment) as well as of the player characteristics. Like pointed out in the beginning of this paper, the core variables of an enjoyable educational game are related to a meaningful parasocial interaction, mood-management and adaptivity on the learner characteristics. However, these very general psychological factors have to be refined for the actual purpose. A meaningful parasocial interaction could be a clever interaction with a hostile NPC for a learning game on rhetorics or it could be a romantic flirt with another player in an educational role playing game when learning about physics. The accordingly psychological subjective variables could be the interest in competition or the sympathy for the other player. (Accordingly consequences of the game might be to change the NPC or strengthen the argumentation or to bring the user to another group of players.) Besides these very specific variables, normally also general psychological variables are of interest, like e.g., the curiousness of the player, confusion, positive versus negative mood or the actual available cognitive resources. Depending on the concrete (educational) game, some of the indicators are more important than others. In principle, the variables have to be defined with respect to the concrete purpose and intervention/adaptation possibilities.

After defining the interesting subjective variables, the associated objective behavioural correlates have to be established for logfiles as well as for NLP. Thereby, it is advantageously to have for the same subjective variable multiple behavioural indicators because this in turn provides a broader data base for the combined analysis of logfile and NLP information.

6.3 Defining associated logfile parameters (including the possibility of ambiguousness and multiconnectedness)

Like pointed out above, within logfiles various and multiple parameters could be recorded. In most cases the crucial variables are hard to identify and the concrete subjective meaning has to be interfered. On the other hand, also the appropriate behavioural indicators are partly hard to define. Sometimes, it is obvious, e.g. the clicking-rate indicates the activity of the user. However, even in this simple case, the question arises what this activity-index concretely means: Is it a sign for confusion or enthusiasm? In such ambiguous constellations some other behavioural data, e.g. the frequency of help-seeking behaviour or number of using help-tools might be insightful. However, in most cases the meaning of the pure behavioural data remains unclear. Furthermore, different educational games have different functionalities and thus, at least some of the assessed parameters have to be designed with respect to the concrete game scenario.

Thus, it is advantageously to have some rough information about the subjective meaning of the behavioural data in advance. For this purpose, pilot-testings with thinking-aloud protocols might be useful. Depending on the concrete case, also former findings on similar games could be used. Furthermore, (based on the assumptions of the media equation theory, [24]) also psychological models on behavioural indicators of subjective experiences could be applied. However, even when being very carefully, it is not always possible to correlate subjective variables with behavioural indicators assessed by logfiles. In this case, the ambiguousness and multiconnectedness should be incorporated in the definition. Table 1 provides some examples.

Table 1: Examples for the potential meaning of behavioural indicators recorded by logfiles

Logfile-Data / Probs	**(Potential) Meaning/interpretation**
Mouse-clicking rate (number of clicks/time unit)	Activity of the user: confusion/nervousness vs. Ambition
Frequency of tool-usage (of the different available tools)	Activity and direction/purpose of user-behavior
Frequency help-seeking behaviour	Confusion / overload of the learner
Sequencing of material (in relation to the task)	Approximation to the solution vs. disorientation
Interaction frequency/rate with NPC or another player	Affiliation/Intensity of parasocial interaction with an NPC / another player: positive (sympathy) vs negative (aversion)
Length of Interaction: Time spent on interaction with an NPC or another player	Affiliation/Intensity of parasocial interaction with an NPC / another player: positive (sympathy) vs negative (aversion)
Eye-contact between player' avatar and the NPC/another player's avatar	Affiliation/Intensity of parasocial interaction with an NPC / another player: positive (sympathy) vs negative (aversion)
Number of characters (NPCs and/or other player's avatars) the player is interacting with	Integration in the (virtual) community/group (Integration-Index)

The technical realisation of the logfile analysis will depend on the level of ambiguousness of logfile data. Depending on the ambiguosness of the interpretable behaviour the methods used will differ.

6.4 Defining classes of NLP concepts

In the following we will list three related NLP usage scenarios that are of relevance for our work. We chose these three scenarios because we think they might be useful to show what we can learn about the player by doing an analysis of what was written during the gameplay.

The first scenario [25] discussed the usefulness of a certain class of words, namely function words, that might help to *identify depression, reactions to individual life stressors, sex, reactions to socially-shared stressors, deception, status and culture*. Function words were defined as a class of words that can be derived from a lexical analysis of text - pronouns, articles, prepositions, conjunctives, and auxiliary verbs. The developed computer program, called LWIC, was used to simply count function words and give the interpretation to the expert user (an implementation based on this work is available on http://www.liwc.net/).

The second scenario [26] used a slightly different approach dealing with the semantic interpretation of words and affect. In this scenario the authors *define affect as the semantic orientation of a subjective sentence*, *directed at an object, having intensity*. Based on this definition it was then tried to detect affect in text automatically, by interpreting the previously defined semantic orientation. The NLP interpretation used in this scenario was much more complex than the simple counting of function words presented in [25].

In the third scenario [27] it was tried to build a training corpus for chatbot communication by identifying *fascinating, acceptable and unsuitable entries a chatbot might give.* The categories of conversations replies were chosen to have potential for high relevance data for the training of autonomous chatbots. The language features most useful turned out to be a combination of features of about the structure of the conversation as well as the content and the reputation of the author.

The scenarios described previously point at some possible important features and classes that might be useful interpretations of text in our application domain. However, due to how machine learning methods work it is hard to foresee which language features and classes are most suitable for the interpretation of user behaviour. Every interpretation depends on the context (the game) the player is in.

6.5 Multimodal Classification of logfile and NLP information

When combining logfile and NLP information several constellations are possible. In the following two main cases will be exemplarily described. First, analysing intersections and communities, and second, augment necessary additional information.

Analysing intersections and communities

If logfile-data as well as NLP delivers ambiguous information regarding the psychological subjective meaning of the behavioural data, the analysis of intersections and communities might be helpful to reduce ambiguousness. Imagine a situation, when the user has a high clicking rate that could be a sign for enthusiasm vs. confusion/nervousness. At the same time, his communication behaviour towards an NPC uses negative words and incomplete sentences that could be a sign for anger and dislike of the NPC vs. confusion. Thus, combining these pieces of information, the probability, that the gamer is confused (and not enthusiastic towards the game or hostile towards the NPC) is rather high. However, it could still be the case, that the gamer only dislikes the NPC and his anger about the game character is the source for confusion/nervousness or the high clicking rate. Thereby the history of communication behaviour with the NPC could provide the necessary clarification. If the negative and incomplete communication style of the user is new (i.e. before the concrete situation the communication with the NPC was rather positive), then one can conclude, that the user is confused by the game. However, if the communication with the NPC was of such negative style from the very beginning, then it can be concluded that the NPC causes confusion and nervousness (via anger).

Adding necessary information

Another constellation might be that logfiles as well as NLP-data provide an only incomplete view of the situation. Imagine a game-situation, in that the user communicates within a group of game characters (NPCs or other gamers within a multiplayer game). The purpose of the (non-invasive) data analysis is to find out which character the user likes and dislikes most. On the one hand via logfiles one can assess the frequency and length of communication/dialogues and the non-verbal interaction with the single other game-characters. But with this information alone one can not decide if this is an indicator of strong liking (e.g. flirting) or disliking (e.g. struggling). On the other hand, NLP information delivers information about the use of negative or positive word. Thereby one can conclude if a single game-character is liked or disliked. However since NLP information delivers only insights about the verbal interaction this could lead to wrong conclusions about the quantitative aspects of liking/disliking. For example, the user might exchange only a few negative words with a strongly disliked character, but nevertheless has a long nonverbal interaction time with him (e.g. fighting). Thus, neither logfiles no NLP alone could lead to an accurate estimation of the liking/disliking of specific game characters. However, the combination of both data-sources allows a more accurate estimation of the sympathy toward other game characters.

These two examples illustrate that the multimodal classification of logfiles and NLP has to be designed with respect to the concrete game and the purpose of data assessment. Additionally, it has to be considered at which level the multimodal classification works best – on a feature level, a high semantic level or on something in between. Are simple rules enough to combine both data sources or are more advanced fusion methods necessary. Besides these two main cases also other constellations are thinkable, depending on the concrete environment/game and the aim of data-assessment.

7 OUTLOOK

Like pointed out above, the multimodal classification of logfiles and NLP will lead to a more holistic view of the user/player and allows a more precise understanding of his/her socio-affective, cognitive and physiological state. Thereby, the described cases are only some rough examples. The concrete implementation of the combined assessment approach depends strongly on the game/computer-based environment and on the purpose of data-assessment (e.g. evaluation versus adaptive interventions or modifications during the gaming process). The framework is open not only in the sense, that multiple indicators for logfiles and NLP could be defined, but also about the combinations of information from both data-sources. It can be assumed that the future research on games and game-based learning might lead to new insight on the association of user-behaviour and subjective meaning. Thereby it is also thinkable to define different rules and indicators for different subgroups of users. For example, there is evidence for gender-dependent communication behaviour. Thus, the inclusion of the user's gender in the definition of indicators and rules could augment the value of data-assessment. The framework is also open in the sense, that other additional data-sources could be incorporated and connected. For example, it is quite easy to assess some user characteristics by a pre-questionnaire before playing and to use these data (e.g. gender, age, usual game-preferences) for more fine grained rules of multimodal classification of different data sources.

AKNOWLEDGEMENT

This paper is part of the EC-Project TARGET funded by the 7th Framework Programme of the European Commission's IST-Programme (contract no. 231717). The authors are solely responsible for the content of this paper. It does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of data appearing therein.

References

[1] Linek, S. B., Marte, B., & Albert, D. (2008). The differential use and effective combination of questionnaires and logfiles. *Proceedings of the International Conference on Interactive Computer Aided Learning (ICL), Special Track "Computer-based Knowledge & Skill Assessment and Feedback in Learning settings" (CAF), 24th to 26th September, 2008, Villach, Austria.*

[2] Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience.* New York: Harper & Collins.

[3] Dondi, C., Edvinsson, B., & Moretti, M. (2004). Chapter 2: Why choosing a game for improving learning and teaching processes? In M. Pivec, A. Koubek & C. Dondi (Eds.). *Guidelines for game-based learning* (pp. 20-76). Lengrich, Germany: Pabst Science Publisher.

[4] Grodal, T. (2000). Video games and the pleasure of control. In D. Zillmann & P. Vorderer (Eds*.), Media entertainment: The psychology of its appeal* (pp. 197-213). Mahwah, N.J.: Erlbaum.

[5] Linek, S. B. (2007a). Creating flow in game-based learning: Threefold conception of challenges and skills. *Proceedings of the International Technology, Education and Development Conference (INTED), 7th-9th March, 2007 Valencia, Spain.*

[6] Vorderer, P., Hartmann, T. & Klimmt, C. (2003). Explaining the enjoyment of playing video games: The role of competition. In D. Marinelli (Ed.), *Proceedings of the 2nd International Conference of Entertainment Computing (ICEC 2003)*, Pittsburg (1-8). New York: ACM.

[7] Vorderer, P., Klimmt, C. & Ritterfeld, U. (2004). Enjoyment: At the heart of media entertainment. *Communication Theory, 14*, 388-408.

[8] Zillmann, D. (1988). Mood management through communication choices. *American Behavioral Scientist, 31*, 327-340.

[9] Zillmann, D. (2006). Dramaturgy for emotions from fictional narration. In J. Bryant & P. Vorderer (Eds.), *Psychology of entertainment* (215-238). Mahwah, New Jersey: Lawrence Erlbaum Associates.

[10] Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety.* San Francisco, California: Jossey-bass.

[11] Asakawa, K. (2004). Flow experience and autotelic personality in Japanese college students: how do they experience challenges in daily life? *Journal of Happiness Studies*, 5(2), 123-154.

[12] Moneta, G. B., & Csikszentmihalyi, M. (1996). The effect of perceived challenges and skills on the quality of subjective experience. *Journal of Personality, 64*, 275-310.

[13] Schweinle A, Meyer D. K., & Turner J. C .(2006). Striking the right balance: Students' motivation and affect in elementary mathematics. *Journal of Educational Research,* 99 (5), 271-293.

[14] Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*, 459-482.

[15] Hebb, D. O. (1955). Drives and the C.N.S. (conceptual nervous system). *Psychological Review, 62*, 243-254.

[16] Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception and Psychophysics, 8*, 279-286.

[17] Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. Cognitive *Science, 5*, 333-369.

[18] Jackendoff, R. S. (1985). Semantics and Cognition (Current Studies in Linguistics) (p. 304). The MIT Press.

[19] Noam Chomsky. (1969). Aspects of the theory of syntax. MIT Press. Schank, R. C. (1972). Conceptual Dependency: A Theory of Natural Language Understanding. Cognitive Psychology, 631(3), 552-631.

[20] McCallum, A. (2005). Information extraction: Distilling structured data from unstructured text. ACM Queue, 3(9), 57. ACM. doi: http://doi.acm.org/10.1145/1105664.1105679.

[21] Navigli, R. (2009). Word sense disambiguation: A survey. ACM Comput. Surv., 41(2), 1-69. New York, NY, USA: ACM. doi: http://doi.acm.org/10.1145/1459352.1459355.

[22] Tang, H., Tan, S., & Cheng, X. (2009). A survey on sentiment detection of reviews. Expert Systems with Applications, 36(7), 10760-10773. doi: DOI: 10.1016/j.eswa.2009.02.063.

[23] Yarowsky, D. (1995). Unsupervised word sense disambiguation rivaling supervised methods. In Proceedings of the 33rd annual meeting on Association for Computational Linguistics (p. 189–196). Association for Computational Linguistics Morristown, NJ, USA. doi: 10.3115/981658.981684.

[24] Reeves, B., & Nass, C. (1996). *The media equation*. Stanford, California: CSLI Publications.

[25] Chung, C. K., & Pennebaker, J. W. (2007). The psychological functions of function words. Social communication: Frontiers of social psychology, 343-359.

[26] Hatzivassiloglou, V., & McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. Proceedings of the 35th annual meeting on Association for Computational Linguistics - (pp. 174-181). Morristown, NJ, USA: Association for Computational Linguistics. doi: 10.3115/976909.979640.

[27] Huang, J., Zhou, M., & Yang, D. (2007). Extracting chatbot knowledge from online discussion forums. In Proc. of IJCAI (pp. 423-428).