Chapter 18
Identifying Latent Semantics in Action Games for Player Modeling

Katia Lida Kermanidis
https://orcid.org/0000-0002-3270-5078
Ionian University, Department of Informatics, Corfu, Greece

ABSTRACT

Machine learning approaches to player modeling traditionally employ a high-level game-knowledge-based feature for representing game sessions, and often player behavioral features as well. The present work makes use of generic low-level features and latent semantic analysis for unsupervised player modeling, but mostly for revealing underlying hidden information regarding game semantics that is not easily detectable beforehand.

INTRODUCTION

Player modeling has been attracting the interest of game design and development experts for several years, as a means to increase player satisfaction and immersion. According to the inclusive reviews in (Smith et al., 2011) and (Hooshyar et al., 2018), modeling techniques vary from empirical (data-driven) (Thue et al., 2007; Thurau & Bauckhage, 2010; Roberts et al., 2007; Geisler, 2002; Drachen et al., 2013), where the application of machine learning or statistical analysis to gaming data enables predictions of playing styles, to theoretical (i.e. analytical), mostly applicable to board-like games, where search and optimization techniques are used to determine the moves towards the best outcome (Bellman, 1965). The term 'play style' indicates the manner in which each player behaves while playing, i.e. the choices he makes, his reactions, his response time etc.

Regarding empirical approaches to player modeling, various learning techniques have been experimented with; supervised, like support vector machines for predicting difficulty adjustment (Missura & Gaertner, 2009), Bayesian networks for classification (He et al., 2008), statistical analysis of the distribution of player actions (Thawonmas & Ho, 2007), and unsupervised, like self-organizing maps

DOI: 10.4018/978-1-6684-7589-8.ch018

(Drachen et al., 2009), reinforcement learning (Kang & Tan, 2010), transfer learning (Shahine & Banerjee, 2007) and preference learning (Yannakakis et al., 2009). Supervised techniques (stand-alone or in combination with unsupervised approaches) have been gaining in popularity (Bisson et al., 2015; Luo et al., 2016; Min et al., 2016; Tamassia et al., 2016; Falakmasir et al., 2016; Gao et al., 2016) during the last three-four years, compared to purely unsupervised approaches (Drachen et al., 2009; Anagnostou & Maragoudakis, 2009; Cowley et al., 2014), mostly due to their improved performance. Dimensionality reduction techniques, other than self-organizing maps, have been experimented with for unsupervised modeling: Linear Discriminant Analysis has been applied to arcade-style as well as combat-style games (Gow et al., 2012), where match data are annotated with the players' identity to enable the supervised application of Linear Discriminant Analysis, and then *k*-means clustering groups together players of the same gaming style.

All previous approaches use a limited number of high-level game and player features to perform modeling, that are game-dependent (vary from game to game, and a game expert is required to define them) and whose impact on the player model is to some extent a-priori sensed. High level features indicate directly and almost explicitly the game status. High-level features in combat-style games may, for instance, include the number of weapons obtained, the number of shots performed, the number of spare lives accumulated. The high-level features pose significant demands on knowledge resources, while they minimize expectations to extract new knowledge and unforeseen relations and dependencies between game and player features. Low level features, on the other hand, are features that describe the morphology of the game terrain, at specific time intervals, and the game status needs to be indirectly deducted.

Latent Semantic Analysis (LSA) has been applied with significant success to several domains, other than IR, like essay assessment in language learning (Haley et al., 2005), intelligent tutoring (Graesser et al., 2007), text cohesion measurement (McCarthy et al., 2007), summary evaluation (Steinberger & Jezek, 2004), text categorization (Nakov et al., 2003). Although all previously mentioned LSA applications have been performed on text corpora, some approaches have proposed its use in different non-textual knowledge domains like board game player modeling (Zampa & Lemaire, 2002), complex problem solving (Quesada et al., 2001), gene function prediction (Done et al., 2010; Dong et al., 2006; Ganapathitaju et al., 2005), web navigation behavior prediction (van Oostendorp & Juvina, 2007), collaborative filtering (Hofmann, 2004), semantic description of images (Basili et al., 2007).

Latent Dirichlet Allocation (LDA) has been applied to several distinct modeling applications, like topic detection (Griffiths & Steyvers, 2004; Weinshall et al., 2013), tag recommendation (Krestel et al., 2009), entity resolution (Bhattacharya & Getoor, 2006), human action recognition in video streams (Deepak et al., 2013), spam recognition (Bro et al., 2009), satellite image clustering (Tang et al., 2013), flood event modeling (Aubert et al., 2013), source code analysis (Grant et al., 2013).

The present work proposes grouping similar playing styles together by modeling the semantics of the game domain. There are two possible ways for supplying domain knowledge (Lemaire, 1998): by hand, making use of domain experts' know-how, and automatically, by deriving the semantics from large corpora of "word" sequences, i.e. sequences of concepts that carry units of meaning related to the domain. The first approach is more accurate, but domain-dependent, while the second (adopted in the present work) is useful when no hand-crafted knowledge is available.

In an attempt to disassociate the player modeling process from the necessity of an already known high-level knowledge-based feature set, the present paper proposes:

Related Content

Design Factors for Effective Science Simulations: Representation of Information
Jan L. Plass, Bruce D. Homer, Catherine Milne, Trace Jordan, Slava Kalyuga, Minchi Kimand Hyunjeong Lee (2009). *International Journal of Gaming and Computer-Mediated Simulations (pp. 16-35).*
www.irma-international.org/article/design-factors-effective-science-simulations/2159

Computer-Generated Three-Dimensional Training Environments: The Simulation, User, and Problem-Based Learning (SUPL) Approach
Michael Garrettand Mark McMahon (2010). *International Journal of Gaming and Computer-Mediated Simulations (pp. 43-60).*
www.irma-international.org/article/computer-generated-three-dimensional-training/47085

World of Race War: Race and Learning in World of Warcraft
Alfred Weissand Sharon Tettegah (2012). *International Journal of Gaming and Computer-Mediated Simulations (pp. 33-44).*
www.irma-international.org/article/world-race-war/74833

Flip-Game Engineering and Technology Methodology
Milagros Huerta Gómez de Merodio, Juan Manuel Dodero, Nestor Mora Núñezand José Mª Portela Núñez (2023). *Research Anthology on Game Design, Development, Usage, and Social Impact (pp. 839-871).*
www.irma-international.org/chapter/flip-game-engineering-and-technology-methodology/315519

Affordances and Constraints of Analog Games for Ethics Education: Dilemmas and Dragons
Spencer P. Greenhalgh (2016). *Examining the Evolution of Gaming and Its Impact on Social, Cultural, and Political Perspectives (pp. 219-237).*
www.irma-international.org/chapter/affordances-and-constraints-of-analog-games-for-ethics-education/157623