Game Abstraction (GA) Using Temporal Event Similarity With Uncertainty

Vinayak Jagtap
Department of Computer Engineering
College of Engineering Pune Maharashtra India
Parag Kulkarni
Chief Scientist
Knowlation Research Labs Pune Maharashtra India
Paper Number: 240007

Abstract
While describing a game, details of game moves in a commentary are not considered, which loses the data. Due to lack of expertise in writing or narrating the game, surprises cannot be revealed completely which loses the interest of game abstract description. The actual spectator has more information about surprises, gameplay, and game conditions in his/her observation of the game. The surprise-based abstraction of the game represents more information than other descriptions. The writer's perspective, reader's perspective, and viewer's perspectives are different which has more information; each one having its perspective for the game surprises. The unrevealed surprises can be mined from each perspective which might create interest in the description. Existing statistical models do not cover these perspectives and surprises. An abstraction of these multi-perspective surprises is proposed in this paper. The surprises are calculated with the help of event uncertainty in the game. The surprising similarity can be measured for various games which can help to decide the game plan, strategy, player style, team psychology. The surprising similarity can be measured with the help of uncertain events in times series of surprises. The surprise-based similarity is compared with the time series similarity measures. The surprise-based similarity game comparison has improved. The surprises can also be revealed in other time-series data like stock market data, network traffic data, etc.
Keywords: *Game Abstraction, Uncertainty, Surprise, Time Series*

1 Introduction

In the current era of multimedia and computing, more and more games are played. They are designed, strategized in such a way that people especially the gaming community should get attracted as well as the new generation too. In recent development, most games are getting attention based on internet advancements like online gaming, multiplayer gaming, and social gaming. These types of new gameplay approaches are evolved and got accepted by gaming communities open-heartedly in computer games.

If anyone can observe computer game development, it has a direct co-relationship with computational evolve. Earlier games are dos games and having less computing related to graphics. As gaming improved in graphics and designing part more complex games were developed which has the computer as an opponent. As computer movements are strategically fixed one having less uncertainty, these games lose interest in gaming communities regardless of their visual effects. Hence, there is a need fora person to person or human to human gameplay and multimedia games are evolved, which got the support of the internet revolution to get more and more interest in online multiplayer gaming. In the recent era, online gaming is most famous and played daily by the young generation.

Gaming strategies are also used in different business development under a computational domain called Gamification. These strategies are incorporated in business for various applications like promotional strategies, social media, engagement, etc. All these strategies make an impact differently based on domain and utilities. In the case of gaming considering brand loyalty if the offer is given such strategies can be adapted in businesses. Some strategies are used to sustain in the market, in such cases new updates or free features are provided for the product to add interest in the product. Like feature phone came with music as utility added which attracted many customers. Hence to stay in the market other feature phone producing companies were also adding these features.

As discussed, there are various ways to apply gaming strategies, but these strategies can be only compared if moves are compared in between them. Most of the time while watching football, or cricket game, anyone can relate special series of moves which can be referred to as event is similar to another match event like Rohit Sharma hit six to Hussain Ali in world cup 2019 is exactly similar to Sachin Tendulkar's six to Shoaib Akhtar. Here, Sachin Tendulkar didn't hit uppercut once only neither Rohit Sharma but why only these two shots are considered similar? The reason is simple these two events are having many similarities and other uppercut didn't. Human nature tries to measure maximum similarities in an event to say both are similar or not. Here similarities are on the different levels like cricket World Cup event, against Pakistan, against one of best bowler of that era, top player, opener, the direction of bowl hit, wide short pitch delivery, etc. Such all similarities are measured in the sequence of events lead to result in these two shots are similar.

Human nature also always tries to find uniqueness in events to say similarities, many cover drives are played in the world cup which is not compared for similarities but events having moves like straight→ cover→pull such sequence comes then analyst rate it for similarities. Hence uniqueness in game-play and sequence matters to find out similarities. This is observed in the same way because other statistical data do not provide actual play similarities in matches. Hence sequences needed to be mapped for similarity measures. In this paper, uncertainty-based unique events identification and their association to find abstract event similarities in a sequence of events are computed to find similarities in the player, player style, and match similarities.

2 Literature Survey

The main stage of discovery of knowledge discovery from time-series data requires sequence analysis. The input sequence could be modelized as time series sequences even they are text or multimedia inputs. In general, numeric data provided in sequence of series for medical and other domains like moves in case of games or sports can be modeled in sequence input with time associated i.e. time series inputs. Many researchers used an integrated approach to discover the knowledge. For this, collaborative and distributed approaches to computing are used [1-3]. Some of the researcher used social media dataset [4], review dataset [5], pre-knowledge of Wikipedia to extract useful information [6-8].

These sequences are compared and analyzed in a different way to derive different inferences by just comparison. Based on information gain in world sequencing and uncertainty associated, sequence comparison is performed by Vigna et al. which represents Euclidian distance between the sequences. They provide results in different formats angle between words and information, covariance, correlation, information theory, Kolmogorov measures, and Universal sequence map [9]. Different alignment and algorithmic techniques are given for subsequence comparison which leads to knowledge discovery in the form of different patterns [10]. Similarly, biological sub-sequences are compared for pattern determination using AI computational methods [11-14].

Once sequences are analyzed and compared, based on inferences, these are modeled in different knowledge models. F. M¨orchen et al. created a framework that finds the muscle active information extraction mechanism. They extracted time-series data acquired from kinesiological EMG to find the pattern for multivariant input sequences [15-17]. To build knowledge there is a necessity of evidence of activities and their documentation in the physical system. Here activities could be drop-down to events for time series input sequences. The cases are mapped to each event to build knowledge in various forms, it may be relationships, graphs, or any other mathematical representation available [18, 19].

Sequences are categorized in events by some researchers to analyze it in case of time series analysis. LeadLine is one of event identification interactive visual analytical situations, which integrate topic modeling, event detection (identification) and name data entity relationship based on 4W's (Who, What, When, Where). The knowledge is mined using time sequence referencing and visually represented to extract more knowledge for decision making. Then they are represented as entity graphs and geo-mapping based on the position of the context in the text [20]. A mathematical representation of event detection for time-series inputs is given by Kom [21]. Similarly, in a nuclear powerplant, to find abnormal event neural network is used [22]. Also, for medical domain event discovery from time-series data [23] EEG signals are processed to find an event for knowledge discovery [24]. In the domain for the social network, based on the post-event is tried to determined using text or other multimedia inputs [25-27]. For finance, sector events are determined based on transaction, notification, and news as inputs [28]. There are various approaches discussed to identify the event from the different domains for knowledge discovery which will help in decision making [29, 30].

These events are clustered and classified for decision-making. By applying fuzzy c-means on the knowledge extracted by using time series frameworks and sub sequencing using Hidden Markov Model to improve the predictive power [31]. Different clustering approaches are applied for time series subsequences inputs [32-35]. Similarly, Deep Neural Network is applied for different Supervised learning approaches i.e. for classification [36, 37]. Table 1 shows the literature review.

Table 1:Literature Survey

ID	Author	Findings
[1].	Gurparkash Singh, Louise Hawkins, and Greg Whymark	• Collaborative Knowledge Building (CKB) based on cultural-historical activity theory
[2].	SannaPekkola, Petri Niemi & Juhani Ukko	• Knowledge maturity model: OSSIC review
[3].	Jim Hewitt and Marlene Scardamalia	• Computer-Supported Intentional Learning Environments

		(CSILE)
[4].	Jun Oshima, RitsukoOshima, Yoshiaki Matsuzawa	• Knowledge Building Discourse Explorer (KBDeX).
[5].	Deborah Finfgeld-Connett	• Content analysis and synthesis • Quantified review of knowledge building
[6].	Ulrike Cress & Joachim Kimmerle	• CKB on wiki knowledge using equilibrium activities
[7].	J. Moskaliuk, J. Kimmerle& U. Cress	
[8].	Joachim Kimmerle, Ulrike Cress, and Christoph Held	• Three technologies introduced using the wiki, Social tagging, pattern-based and use of wiki
[9].	Susana Vinga and Jonas Almeida	• Information theory-based, word sequence-based, word angle based, and word distance-based sequence mapping
[10].	Tak-Chung Fu	• Whole sequence matching • Subsequence matching
[11].	Kun-Mao Chao, Louxin Zhang	• Different alignment techniques • Different algorithms for pairwise sequence comparison
[12].	William R. Pearson	• Comparison between algorithm FASTP and FASTA

[13].	William R. Pearson and David J. Lipman	• Three Algorithms to compare biological protein sequences: FASTA, RDF2, LFASTA
[14].	Michael Levitt, and Mark Gerstein	• Comparison between FASTA and BLAST Sequence comparison algorithms
[15].	Guy St C Slater and Ewan Birney	• Sequence comparison using heuristic bounded sparse programming
[16].	Fabian M¨orchena, Alfred Ultscha and Olaf Hoosb	• TSKM Framework • Primitive pattern finding
[17].	Fabian Mörchen& Alfred Ultsch	• Time series knowledge representation (TSKR) • Fast interval pattern finding using depth-first and marginal pattern approach
[18].	Fabian Moerchen	
[19].	Donald Philip	• Teaching Knowledge building framework
[20].	William Cerbin and Bryan Kopp	
[21].	Wenwen Dou, Xiaoyu Wang, Drew Skau, William Ribarsky, and Michelle X. Zhou	• Event identification using LeadLine and 4Ws
[22].	Edward L. Kom, Barry I. Graubard, and Douglas Midthune	• Mathematical analysis to identify task from time series
[23].	YukiharuOhga and Hiroshi Seki	• Neural Network-based abnormal event identification
[24].	Christine L. Tsien	• Medical event

		discovery using medical time series data analysis
[25].	Vernon Lawhern, W. David Hairston, Kay Robbins	• Artifact detection in EEG based on event identified and classify
[26].	M´ario Cordeiro(B) and Jo˜ao Gama	• Event detection and discovery: RED, NED • Supervised and unsupervised event detection
[27].	Anuradha Goswami, Ajey Kumar	• Social text stream based event detection • Detailed survey
[28].	Farzindar Atefeh And Wael Khreich	• Different event detection for Twitter text dataset
[29].	Richard E. A. Escher	• Neural network-based financial event detection
[30].	H. Langer, S. Falsaperla, T. Powell, G. Thompson	• Five fundamental classes: Volcano-Tectonic Events, Regional Events, Long-Period Events, Hybrid Events, and Rockfalls. Identification and classification of these transients, which have been hitherto carried out manually
[31].	SoheilaMehrmolaei and Mohammad Reza Keyvanpourr	• Statistical and ARIMA based classification for events
[32].	Vasile Georgescu	• Performed Time-

		series clustering especially fuzzy c-means clustering ● Improvement in predictive power using HMM
[33].	Saeed Aghabozorgi, Ali SeyedShirkhorshidi, The Ying Wah	● Different time series clustering approaches: wholistic, time subsequence, time point clustering
[34].	T. Warren Liao	● Different clustering of time series survey
[35].	Geeta Shikka	
[36].	Sylvia Frühwirth-Schnatter	● Markov Chain, Bayesian, and Monto Carlo approach for clustering
[37].	Hassan Ismail Fawaz · Germain Forestier, · Jonathan Weber, · LhassaneIdoumghar, · Pierre-Alain Muller	● Time Series Classification (TSC) ● Use of CNN-based deep learning.
[38].	Eamonn Keogh, Shruti Kasetty	● Use of classification, clustering, indexing, and segmentation for the importance of time-series data

3 Proposed Game Abstraction (GA) System

In the case of similarity measurement, statistically, similarity only provides results of move, the sequence of moves, or match results. Ex Australia Vs South Africa cricket match held in 2006 at Johannesburg can be mapped both innings are similar on key events that happened in a match like the run rate at 10 or 20 overs but the match is different with opening partnership and in other constraints. These small changes are not articulated when events are mapped statistically. Here uncertainty in the form of surprises played a very important role. Sequences are needed to be mapped along with statistics in events. Hence, the system is proposed based on uncertainty available in sequences as well as uncertainty in statistics. Figure 1 represents an uncertainty-based abstraction of game events. In the proposed system sequences uncertainty is computed with statistical uncertainty is computed. These with event similarities are used to determine similarities for different similarities.Context-Li [39] Model is used to limit uncertainty which is out of bound. Like in cricket beamer ball batsman doesn't need to play the delivery hence uncertainty related to this can be omitted while operating.

Cricket commentary is used as game inputs. For text commentary bag of words,theapproach is applied to find surprises in the text, and features are extracted, represented in FM matrix with weights in WMmatrix as shown in figure 2. Events from the game filtered for surprises and relevant abstract events are filtered. Like in cricket; single and double run doesn't provide as a surprise as six or four runs. Then, Surprise Index(SI) in the domain of [0,1] is calculated on filtered data as a function based on FM and WMas:

$$SI = \{BoW, n - gram, Wordnet\} \quad (1)$$

Figure 1: A Game Abstraction System using Surprises

Based on SIcalculated for each commentary used abstraction. Usually, human tries to an abstract situation based on his/ her perspective. Here without considering such biased perspective abstraction is provided for Cricket commentary. It improves accuracy while considering surprises in events while abstracting.

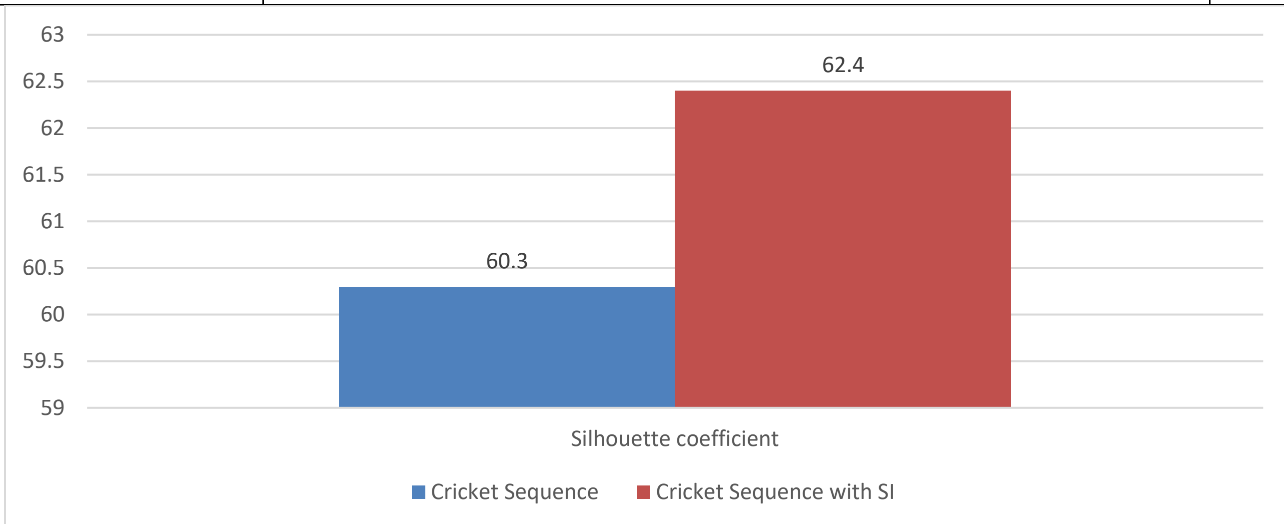
4 Results and Discussion

IPL Cricket commentary is used from cricbuzzAPI. The dataset is created has **54*240** (one season data with each delivery commentary of 54 matches and each match has 240 deliveries)

Table2: A Sample Commentary and SI

Over	Commentary	SI
0.5	Nehra to Mandeep, FOUR, first boundary for Mandeep and RCB. Full and on the pads, needed to be put away and Mandeep did just that, picked it up and dispatched it over mid-wicket, couple of bounces and into the fence	0.33
1	Nehra to Mandeep, FOUR, back-to-back boundaries to end the first over. Again, Nehra is a tad short in his	0.18

	length, Mandeep had the width to cut and he didn't try to hit it hard, just placed it behind point and Bhuvi at third man gave up the chase pretty quickly	



Abstraction Sequence Clustering comparison

Dataset is filtered and *SI* is calculated. Table 2 represents from commentary based on *SI* matches events are represented. It is seen that *SI* is different but statistics is four runs and can have more meaning in abstraction. Figure 3 represents sequence clustering results of all events. Based on *SI* and statistical data as shown in table 2 sequence clustering was performed to measure abstract level similarity. The results improved due to surprises in the match is considered, by clustering in abstract level two players or two matches are similar or not.

5 Conclusion

Automatic abstraction or summarization misses important events or it is biased with the person. This reduces the visibility of news or another document. It happens more with fast dynamic events like gaming etc. where there are each event is unique. Finding out the most suitable event in summary or abstract is mostly computed manually. Using surprises from inputs, abstraction might provide more meaning. Adding uncertainty based computation again filter and refine surprises. Here significant improvement for similarity measurement as well as in abreaction will help in multiple domain applications like share market, online transaction, etc.

References

Singh, G., Hawkins, L. and Whymark, G., 2007. An integrated model of collaborative knowledge building. Interdisciplinary Journal of E-Learning and Learning Objects, 3(1), pp.85-105. available: www.learntechlib.org

Pekkola, S., Niemi, P. and Ukko, J., 2013. Building understanding of the development of performance management for collaborative networks with a knowledge maturity model. International journal of networking and virtual organisations, 12(3), pp.179-200.

Hewitt, J. and Scardamalia, M., 1998. Design principles for distributed knowledge building processes. Educational psychology review, 10(1), pp.75-96.

Oshima, J., Oshima, R. and Matsuzawa, Y., 2012. Knowledge Building Discourse Explorer: a social network analysis application for knowledge building discourse. Educational technology research and development, 60(5), pp.903-921.

Finfgeld-Connett, D., 2014. Use of content analysis to conduct knowledge-building and theory-generating qualitative systematic reviews. Qualitative research, 14(3), pp.341-352.

Cress, U. and Kimmerle, J., 2008. A systemic and cognitive view on collaborative knowledge building with wikis. International Journal of Computer-Supported Collaborative Learning, 3(2), pp.105-122.

Moskaliuk, J., Kimmerle, J. and Cress, U., 2009. Wiki-supported learning and knowledge building: effects of incongruity between knowledge and information. Journal of Computer Assisted Learning, 25(6), pp.549-561.2729.2009.00331.x

Kimmerle, J., Cress, U. and Held, C., 2010. The interplay between individual and collective knowledge: technologies for organisational learning and knowledge building. Knowledge Management Research & Practice, 8(1), pp.33-44.

Vinga, S. and Almeida, J., 2003. Alignment-free sequence comparison—a review. Bioinformatics, 19(4), pp.513-523.

Fu, T.C., 2011. A review on time series data mining. Engineering Applications of Artificial Intelligence, 24(1), pp.164-181.

Chao, K.M. and Zhang, L., 2008. Sequence comparison: theory and methods (Vol. 7). Springer Science & Business MediaPearson, W.R., 1990. [5] Rapid and sensitive sequence comparison with FASTP and FASTA.

Pearson, W.R. and Lipman, D.J., 1988. Improved tools for biological sequence comparison. Proceedings of the National Academy of Sciences, 85(8), pp.2444-

Levitt, M. and Gerstein, M., 1998. A unified statistical framework for sequence comparison and structure comparison. Proceedings of the National Academy of sciences, 95(11), pp.5913-

Slater, G.S.C. and Birney, E., 2005. Automated generation of heuristics for biological sequence comparison. BMC bioinformatics, 6(1), pp.1-11.
Mörchen, F., Ultsch, A. and Hoos, O., 2005. Extracting interpretable muscle activation patterns with time series knowledge mining. International Journal of Knowledge-based and Intelligent Engineering Systems, 9(3), pp.197-208.
Mörchen, F. and Ultsch, A., 2007. Efficient mining of understandable patterns from multivariate interval time series. Data mining and knowledge discovery, 15(2), pp.181-215.
Moerchen, F., 2006, August. Algorithms for time series knowledge mining. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 668-673).
Philip, D., 2007. The knowledge building paradigm: A model of learning for net generation students. Innovate: Journal of Online Education, 3(5).available: www.learntechlib.org
Cerbin, W. and Kopp, B., 2006. Lesson study as a model for building pedagogical knowledge and improving teaching. International journal of teaching and learning in higher education, 18(3), pp.250-257.
Dou, W., Wang, X., Skau, D., Ribarsky, W. and Zhou, M.X., 2012, October. Leadline: Interactive visual analysis of text data through event identification and exploration. In 2012 IEEE Conference on Visual Analytics Science and Technology (VAST) (pp. 93-102). IEEE. 10.1109/VAST.2012.6400485
Kom, E.L., Graubard, B.I. and Midthune, D., 1997. Time-to-event analysis of longitudinal follow-up of a survey: choice of the time-scale. American journal of epidemiology, 145(1), pp.72-80.
Ohga, Y. and Seki, H., 1993. Abnormal event identification in nuclear power plants using a neural network and knowledge processing. Nuclear Technology, 101(2), pp.159-167.
Tsien, C.L., 2000. Event discovery in medical time-series data. In Proceedings of the AMIA Symposium (p. 858). American Medical Informatics Association.
Lawhern, V., Hairston, W.D. and Robbins, K., 2013. DETECT: A MATLAB toolbox for event detection and identification in time series, with applications to artifact detection in EEG signals. PloS one, 8(4), p.e62944.
Cordeiro, M. and Gama, J., 2016. Online social networks event detection: a survey. In Solving Large Scale Learning Tasks. Challenges and Algorithms (pp. 1-41). Springer, Cham. 6_1
Goswami, A. and Kumar, A., 2016. A survey of event detection techniques in online social networks. Social Network Analysis and Mining, 6(1), pp.1-25.

Atefeh, F. and Khreich, W., 2015. A survey of techniques for event detection in twitter. Computational Intelligence, 31(1), pp.132-164.
Escher, R.E., Recognia Inc, 2008. Method of providing a financial event identification service. U.S. Patent 7,469,226.
Langer, H., Falsaperla, S., Powell, T. and Thompson, G., 2006. Automatic classification and a-posteriori analysis of seismic event identification at Soufriere Hills volcano, Montserrat. Journal of volcanology and geothermal research, 153(1-2), pp.1-10.
Mehrmolaei, S. and Keyvanpourr, M.R., 2015. A brief survey on event prediction methods in time series. In Artificial Intelligence Perspectives and Applications (pp. 235-246). Springer, Cham.
Georgescu, V., 2009. A Time Series Knowledge Mining Framework Exploiting the Synergy between Subsequence Clustering and Predictive Markovian Models. Fuzzy Economic Review, 14(1), p.41.
Aghabozorgi, S., Shirkhorshidi, A.S. and Wah, T.Y., 2015. Time-series clustering–a decade review. Information Systems, 53, pp.16-38.
Liao, T.W., 2005. Clustering of time series data—a survey. Pattern recognition, 38(11), pp.1857-1874.
Rani, S. and Sikka, G., 2012. Recent techniques of clustering of time series data: a survey. International Journal of Computer Applications, 52(15).
Frühwirth-Schnatter, S., 2011. Panel data analysis: a survey on model-based clustering of time series. Advances in Data Analysis and Classification, 5(4), pp.251-280
Fawaz, H.I., Forestier, G., Weber, J., Idoumghar, L. and Muller, P.A., 2019. Deep learning for time series classification: a review. Data Mining and Knowledge Discovery, 33(4), pp.917-963.
Keogh, E. and Kasetty, S., 2003. On the need for time series data mining benchmarks: a survey and empirical demonstration. Data Mining and knowledge discovery, 7(4), pp.349-371.
Jagtap, V. and Kulkarni, P., 2019, February. Contextual High-Level Uncertainty Modeling Reducing Surprises in Decision Making. In 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT) (pp. 1-4). IEEE.