

Structuring Ordered Nominal Data for Event Sequence Discovery

Chreston A. Miller
Center for Human-Computer
Interaction, Virginia Tech
2202 Kraft Drive,
Blacksburg, Va. 24060, USA
chmille3@cs.vt.edu

Francis Quek
Center for Human-Computer
Interaction, Virginia Tech
2202 Kraft Drive,
Blacksburg, Va. 24060, USA
quek@cs.vt.edu

Naren Ramakrishnan
Dept. of Computer Science,
Virginia Tech
114 McBryde Hall (0106),
Blacksburg, Va. 24061, USA
naren@vt.edu

ABSTRACT

This work investigates using n-gram processing and a temporal relation encoding to providing relational information about events extracted from media streams. The event information is temporal and nominal in nature being categorized by a descriptive label or symbolic means and can be difficult to relationally compare and give ranking metrics. Given a parsed sequence of events, relational information pertinent to comparison between events can be obtained through the application of n-grams techniques borrowed from speech processing and temporal relation logic. The procedure is discussed along with results computed using a representative data set characterized by nominal event data.

Categories and Subject Descriptors

I.2.4 [Artificial Intelligence]: Knowledge Representation Formalisms and Methods—*relation systems, temporal logic*

General Terms

Design

Keywords

Temporal Events, Temporal Patterns, Nominal Data, Event Sequences, N-grams, Probabilistic Modeling, Sequence Modeling, Temporal Relations

1. INTRODUCTION

Nominal data is descriptive in nature but difficult to assign any ordering or interval relational structure of one piece of data with respect to another. Typical of such data are those that are conceptually and categorically labeled where any numerical ordering of the data is difficult. For ordinal values, this question is easily solved since the data are numerical in nature or can be relationally ranked or ordered (e.g., high,

medium, low). Relational information about nominal data is dependent on the context and the interpreter of the data.

This work focuses on providing a relational understanding between events from media streams. Providing relational insight to events seen in media streams such as audio, video, and motion data is difficult. In the analysis of an interaction session where college seniors employed a high-resolution tabletop [16] for cooperative analysis in a history class [2], the video and audio recordings of the session were coded for (vocal) referential foci, pointing gestures, and gaze fixation using MacVisSTA [23], a multimodal analysis tool. The analysis of relational information between all the coded events was conducted by humans since the nature of the information of interest is nominal and difficult to process otherwise. However, when only the recorded events are available and a model of understanding is desired, building a meaningful model may be difficult since ranking one event with respect to another is dependent on a number of factors that are typically dependent on context.

Media streams of data capture a series of these temporal events. A way to approach the construction of such a model is to start from a small subset of sequential events and “grow” the model one event at a time by incrementally incorporating events of interest. The challenge becomes providing the relational structure of these events that supports addition or removal of events or event substructures to the growing model. How do you compare events that are nominal in nature? One possible answer is by probabilistic modeling of the temporal sequence of events. We base our temporal relationship model on Allen’s temporal relation principles [1]. Many techniques have been created to discover these patterns in given data sets. This work focuses on providing relational information between events based on these patterns so as to better understand how the events interact.

We propose understanding this interaction better by applying n-gram processing and a temporal relation encoding. N-gram processing is a technique for probabilistic modeling of a sequence [3]. Techniques have been developed to use n-grams for speech processing through creating probabilistic models based on sequences of phonemes. Temporal event information can be viewed analogously as a sequence of phonemes in which n-gram probabilistic processing may be applied. Section 2 discusses other approaches relevant to processing temporal event data and nominal event data. Section 3 describes the proposed approach applied to temporal nominal data. Following, Section 4 presents an experiment for the provision of how the proposed system works

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM’10, October 25–29, 2010, Firenze, Italy.

Copyright 2010 ACM 978-1-60558-933-6/10/10 ...\$10.00.

within a data set, after which, conclusions and future work are presented and discussed.

2. RELATED APPROACHES

The ordered relationships that are inherent to temporal event data have been an active area of research. Allen in his work [1] formulated thirteen relationship principles that express all the possible ordering relationships between two events. Since then, much work has been conducted to detect, extract, and represent such temporal information. Examples are [6, 9, 12, 18, 19, 24, 27]. These works represent the trend of discovering events and temporal relationships from temporal data.

Mörchen [20] presents Time Series Knowledge Mining (TSKM) which also is a method to extract and represent temporal knowledge. He further discusses other methods for prediction of events. Such prediction based on temporal patterns is closely related to our proposed work which is more focused on the relationships between events with respect to relational order and history. In [10], prediction is acquired through probabilistic modeling based on detected temporal patterns. The interest of our proposed work is not predicting patterns but understanding the ordered relationships between events, i.e., how do two events relate structurally and contextually.

The authors of [13] report on discovering the relationships between patterns. A number of techniques designed to represent temporal data and discover patterns using unsupervised means are described in [21].

Assigning relational structure to nominal data can be difficult since the meaning can be context dependent and depend on the interpreter and her experiences and expertise. The authors of [14] use a similarity technique in performing unsupervised learning with mixed numeric and nominal data. Through this they were able to cluster the data to provide classification information. This is in contrast to [5] where the authors use dissimilarity to learn classification information for nominal data. Fuzzy logic [29, 30] has also been employed for discriminating between nominal features. However, fuzzy logic requires pre-defined rules that govern how the nominal features relate and hence painstaking consideration is needed before processing can begin.

A number of probabilistic techniques have been employed to describe nominal data. Such techniques include stochastic context-free grammars, Hidden Markov models (HMMs), and class-based n-grams [11, 25, 26]. N-gram processing has been successfully employed as a probabilistic model for speech processing [3, 26]. This kind of processing gives a probabilistic model of successive speech features based on previously seen features, hence providing ordered relational sequence information.

3. PROPOSED APPROACH

The processing of speech phonemes has been successfully done using n-grams for purposes of speech recognition [3]. Such processing defines the probability of one phoneme following others. Then given that k phonemes have occurred, where $k \geq n$, the probability of the next phoneme in the sequence can be given, hence a sequence of phonemes can be built one at a time through probabilistic measures. Initially, phonemes were the basic building block for such process-

ing, but processing has also been conducted with respect to words [3, 11, 22], phrases [8], and word classes [3, 15].

Ordered nominal data within a temporal event stream can be seen as a sequence in which the order gives meaning to the individual parts and the relationships between each piece of data can be defined by this ordering. This is similar to the relational information used in n-gram processing of phonemes. Phonemes within speech have ordered relationships to other phonemes. The n-gram processing of elements of speech provide the relational information necessary for more accurate processing as the previous sequence of phonemes give an idea of what is next. This is possible since n-gram processing is based on conditional probability where $Pr(A|B)$ is the probability of A given B where B is the history and context. For more details on the mathematics of n-grams, the reader is directed to [3].

We would like to use this same processing concept to provide relationship information to temporal event data. Therefore, we would like to answer the question: given a sequence of events, how does one rank events so the next event in the sequence can be wisely chosen? We propose that given a simple mapping, n-gram processing can be used to give meaningful relational information to aid in answering this question. This mapping can be realized through application of Allen's temporal relation principles [1]. An event can be described by its start time and end time. As done in [12], the relationships between events' start and end times can represent all of Allen's principles hence encompassing all relational ordering possibilities. A simple schema similar to that of [12] can be used to represent a sequence of temporal events in a single linear data stream (Figure 1 A).

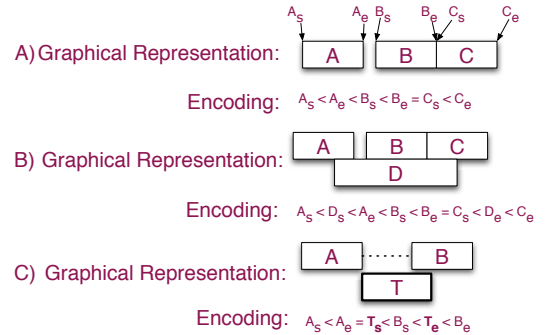


Figure 1: A) Example encoding of a sequence of events. A_s and A_e represent the start and end times of event A , respectively. Relational operators are used to indicate the ordered relations between start/end times. B) Example of events with overlap. C) Example of events with a temporal constraint.

It can easily be seen that all of Allen's principles can be represented through such an encoding, including the many forms of overlap (Figure 1B). In addition, the relational operators between start/end (S/E) times are able to capture how these times interact. An event A might end before B begins ($<$) or at the same time ($=$). The interpreted meaning of these operators with respect to the involved events depends on the events and their context.

The probabilistic expression of a sequential sequence of events without overlaps is analogous to current n-gram processing of speech. Here events are seen as atomic units where

n-gram processing reveals *next/previous* events adjacent to the query. However, with the introduction of overlap, the atomic unit of an event must be broken further down into its *S/E* times. Hence, the probabilities of *next/previous* events are now with respect to *S/E* times. For overlap, another level of n-gram processing is needed to catch overlap adjacencies. For example, in Figure 1 B, the query for event A , $A_s < A_e$, will give information about B_s as an adjacent *S/E* time to A . However, the sequence for A is interrupted by D_s . One form of n-gram processing can provide the adjacent events to A and another can provide overlap information, hence capturing D_s and B_s . These different forms of processing are based on the choice of n and performing selective processing on the event stream.

Another relational constraint to consider that is not part of Allen’s principles is the relational timing between events. A sequence of events may only be of interest if certain timing constraints between events are held. For example, the sequence of $A_s < A_e < B_s < B_e$ may only be of interest if B starts within a certain time window of A ending. To handle this need, a timing event can be introduced to capture the temporal constraint as seen in Figure 1C. A query with a conditional insertion of a timing constraint provides a means to incorporate this constraint information within the context of n-gram processing. Hence, searching for $A < B$ with timing constraint T as in Figure 1C would include a query for $A < B$ with a conditional insertion of T after A . The overlap of T and B will signify a satisfied constraint.

With this linear description of temporal events, n-gram processing can be applied similarly to that of speech processing. The contribution of this approach places emphasis on an applicable ability to give a categorical ranking of data that is normally difficult to do so otherwise. This proposed approach relies on event data extracted from media streams. Earlier mentioned examples are gaze fixations, pointing gestures, and other surface level characteristics of observable behavior. Such examples are within the context of recorded human behavior from various media streams. A description of techniques currently available for said extraction is beyond the scope of this paper but are well studied areas [7, 17, 28].

4. EXPERIMENT

We wrote a library in Python based on the n-gram mathematics described in [3]. The library is initialized with a set of event sequences and queried with a subset of a sequence resulting in probabilistic relationships to potential *previous/next* events. This library provides forward and backward probabilistic viewing since a query cannot be assumed to always describe the beginning of a sequence.

To demonstrate this proposed system, a small set of event sequences were extracted from the well understood domain of baseball. A set of play-by-play event sequences were encoded using the above procedure. The play-by-play information was recorded from the publicly available database at <http://www.hokiesports.com/baseball/stats/2009/>. There are many aspects of a baseball game that can be recorded and represented in a multimedia stream of temporal information: All video and audio of the field and players, the feedback from the crowd, video and audio of both team dugouts, etc. For simplicity of exemplifying the approach, a set of ten plays consisting of current active players on the field during a play were encoded. These plays represent

Table 1: Results - Batter Hits ball and runs to 1st

Previous	Event	Probability	Operator
	Beginning	1.0	N/A
Next			
	R off 2nd _s	0.09	=
	R off 1st _s	0.09	=
	2nd bm to 2nd _s	0.09	=
	R to 3rd _s	0.09	=
	R to 2nd _s	0.09	=
	CF to ball _s	0.09	=
	1st bm to 1st _s	0.18	=
	R to home _s	0.18	=
	1st bm catch ball _s	0.09	<

Table 2: Results - SS catches the ball.

Previous	Event	Probability	Operator
	R off 2nd _s	0.17	<
	R off 1st _s	0.17	<
	B to 1st _s	0.17	<
	B hit ball _e	0.17	<
	CF throw to SS _e	0.17	=
	C throw to SS _e	0.17	=
Next			
	End	0.125	N/A
	R to 2nd _s	0.125	=
	B out _s	0.125	=
	R to 1st _s	0.125	=
	B to 1st _e	0.125	=
	R off 1st _e	0.125	=
	R off 2nd _e	0.125	=
	SS tags R _e	0.125	=

83 events, with 48 being unique. The current version implemented provides the probabilities of the *previous/next* events with some overlap capabilities, however the support of time constraints is not included at this time. The results also include the associated relational operators with respect to the query.

Two queries representing common events in baseball games were submitted to the database of plays. The results of each query give a probabilistic distribution of *previous/next* events where each distribution sums to one. Due to rounding, some of the distributions do not sum to one with an error range of ± 0.02 .

The first query is $[B \text{ hit ball}_s < B \text{ hit ball}_e = B \text{ to 1st}_s]$ where B is the batter. This query looks at the beginning of a play and sees what possible events occur after the batter hits the ball and begins to run to first base. The results can be seen in Table 1 where R, 1st bm, 2nd bm, and CF correspond to runner(at bat team player on base), first and second basemen, and center fielder, respectively. The results for the *previous* events show this query is only seen at the beginning of a sequence. The probabilistic results for the *next* events show a distribution of event start possibilities consisting of runner behavior on base(s) and opponent players positioning themselves according to the ball behavior. The conditional operator is also given to provide insight as to how the events conditionally relate to the query.

The second query is $[SS \text{ catch ball}_s < SS \text{ catch ball}_e]$ where SS denotes the short stop. This query is interested in the actions taken by the SS after catching the ball. The results can be seen in Table 2 where C corresponds to the catcher.

The results of the *previous* events show a probabilistic distribution of events that can explain how the SS potentially came into possession of the ball. The *next* events encompass a batter or runner in the beginning or ending process of acquiring ground, several of which give the SS an opportunity to try and get an opponent out as they are either in transit to a base or have bravely stepped off a base. All the events that began previously to the SS catching the ball appear to end when the SS has possession of the ball.

5. CONCLUSION AND FUTURE WORK

These preliminary results exemplify how combining temporal encodings and n-gram processing results in a strong descriptive measure with which to compare and rank events. Comparison and ranking information is necessary when a model is being built from concurrent media streams. The temporal encoding provides ordered relationships between *S/E* times of events while n-gram processing provides event probabilities based on context and history. Arguably prediction from temporal patterns [10, 20, 21] can be molded for the same probabilistic measures, however doing so essentially results in probabilities based on sequence history which is exactly what n-grams are designed to describe.

Future work includes a further investigation of the relational operator. A *previous/next* event might have several different relational occurrences within the database with respect to the query, giving multiple associated relational operators. Inclusion of time constraints and how n-gram processing behaves for time constraints is also in progress as is how event overlap influences n-gram processing.

Future experiments will include data sets from military exercises and wargame scenarios [4]. The data sets are rich in provided corpus for analysis and have already been partially analyzed providing a known ground truth for comparison.

6. ACKNOWLEDGMENTS

We would like to thank Mark Croushorn for his aid in parsing the baseball plays. This work was partially funded by FODAVA grant CCF-0937133.

7. REFERENCES

- [1] J. F. Allen. Maintaining knowledge about temporal intervals. *Commun. ACM*, 26(11):832–843, 1983.
- [2] C. Andrews, T. Henry, C. Miller, and F. Quek. Cardtable: An embodied tool for analysis of historical information. In *Tabletop 2007*, 2007.
- [3] P. F. Brown, P. V. deSouza, R. L. Mercer, V. J. D. Pietra, and J. C. Lai. Class-based n-gram models of natural language. *Comput. Linguist.*, 18(4):467–479, 1992.
- [4] L. Chen, R. Rose, Y. Qiao, I. Kimbara, F. Parrill, H. Welji, T. Han, J. Tu, Z. Huang, M. Harper, F. Quek, Y. Xiong, D. McNeill, R. Tuttle, and T. Huang. Vace multimodal meeting corpus. *MLMI*, pages 40–51, 2006.
- [5] V. Cheng, C.-H. Li, J. T. Kwok, and C.-K. Li. Dissimilarity learning for nominal data. *Pattern Recognition*, 37(7):1471 – 1477, 2004.
- [6] P. Cohen. Fluent learning: Elucidating the structure of episodes. *Advances in Intelligent Data Analysis*, pages 268–277, 2001.
- [7] D. M. Gavrilu. The visual analysis of human movement: a survey. *Comput. Vis. Image Underst.*, 73(1):82–98, 1999.
- [8] E. Giachin. Phrase bigrams for continuous speech recognition. *IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing*, 1:225–228, 1995.
- [9] G. Guimarães and A. Ultsch. A method for temporal knowledge conversion. *Advances in Intelligent Data Analysis*, pages 369–380, 1999.
- [10] F. Höppner. *Knowledge Discovery from Sequential Data*. PhD thesis, Technischen Universität Braunschweig Genehmigte, 2003.
- [11] Z. Huang, V. Eidelman, and M. Harper. Improving a simple bigram hmm part-of-speech tagger by latent annotation and self-training. In *NAACL '09*, pages 213–216, Morristown, NJ, USA, 2009. Association for Computational Linguistics.
- [12] P. Kam and A. Fu. Discovering temporal patterns for interval-based events. *Data Warehousing and Knowledge Discovery*, pages 317–326, 2000.
- [13] S. Laxman and P. Sastry. A survey of temporal data mining. *Sadhana*, 31(2):173–198, 04 2006.
- [14] C. Li and G. Biswas. Unsupervised learning with mixed numeric and nominal data. *IEEE Trans. on Knowl. and Data Eng.*, 14(4):673–690, 2002.
- [15] S. Martin, J. Liermann, and H. Ney. Algorithms for bigram and trigram word clustering. *Speech Comm.*, 1998.
- [16] C. Miller, A. Robinson, R. Wang, P. Chung, and F. Quek. Interaction techniques for the analysis of complex data on high-resolution displays. In *IMCI '08*, pages 21–28, New York, NY, USA, 2008. ACM.
- [17] T. B. Moeslund, A. Hilton, and V. Krüger. A survey of advances in vision-based human motion capture and analysis. *Comput. Vis. Image Underst.*, 104(2):90–126, 2006.
- [18] C. H. Mooney and J. F. Roddick. Mining relationships between interacting episodes. In *SDM'04*, 2004.
- [19] F. Mörchén. Algorithms for time series knowledge mining. In *KDD '06*, pages 668–673, New York, NY, USA, 2006. ACM.
- [20] F. Mörchén. *Time Series Knowledge Mining*. PhD thesis, Philips-University Marburg, Germany, 2006.
- [21] F. Mörchén. Unsupervised pattern mining from symbolic temporal data. *SIGKDD Explor. Newsl.*, 9(1):41–55, 2007.
- [22] S. Ortmanns, H. Ney, and A. Eiden. Language-model look-ahead for large vocabulary speech recognition. *Intl. Conf. on Spoken Language Processing*, 1996.
- [23] R. T. Rose. Macvissta: A system for multimodal analysis of human communication and interaction. Master's thesis, Virginia Tech, 2007.
- [24] E. Schwalb and L. Vila. Temporal constraints: A survey. *Constraints*, 3(2/3):129–149, 1998.
- [25] A. Stolcke and S. M. Omohundro. Inducing probabilistic grammars by bayesian model merging. In *ICGI '94*, pages 106–118. Springer-Verlag, 1994.
- [26] X. Sun and T. H. Applebaum. Intonational phrase break prediction using decision tree and n-gram model. In *Proc. of Eurospeech2001*, pages 3–7, 2001.
- [27] A. Ultsch. Unification-based temporal grammar. Technical report, Philips-University Marburg, Germany, 2004.
- [28] A. Yilmaz, O. Javed, and M. Shah. Object tracking: A survey. *ACM Comput. Surv.*, 38(4):13, 2006.
- [29] L. Zadeh. Soft computing and fuzzy logic. *Software, IEEE*, 11(6):48–56, Nov 1994.
- [30] L. Zadeh. Fuzzy logic = computing with words. *IEEE TFS*, 4(2):103–111, May 1996.