Exploiting Temporal Relations of Events for Classification Tasks – Proof of Concept Study with Synthetic Data Sets*

Mark Gromowski[1][0000-0003-3376-6566], Ute Schmid[1,2][0000-0002-1301-0326], and Dennis Müller[2][0000-0002-4482-4912]

[1] Cognitive Systems, University of Bamberg, Bamberg, Germany
{mark.gromowski,ute.schmid}@uni-bamberg.de
[2] Fraunhofer Institute for Integrated Circuits IIS, Project Group Comprehensible AI
dennis.mueller@iis.fraunhofer.de

Abstract. We report on results of an ongoing investigation into the suitability of the ILP system Aleph for time series represented using an interval algebra. More specifically, we synthetically generate data sets consisting of suitable representations of sequences of facial expressions indicating emotional states such as pain, using multiple sets of generating rules of differing complexity, and analyse Aleph's capability to recover the original rule sets for different parameter settings. We conclude that Aleph, despite being a generic ILP system not optimized for time series, indeed manages to accurately discover complex patterns in our data.

Keywords: Event Sequences · Temporal Reasoning · Pain Classification

1 Introduction

Reasoning about actions and change in dynamic systems is crucial for many application domains, such as the classification of action sequences of human or robot agents [4] or in genetics [7]. Inductive logic programming (ILP) has been shown to be well suited to learn event calculus theories [6]. In many event calculi, events are represented as points, defining event sequences as successions of elements out of some event alphabet. Alternatively, events can be represented as intervals, as proposed in the Allen calculus for temporal reasoning [1].

In an interdisciplinary project, we explore how facial expressions indicate emotions and other internal states such as pain [8]. Diagnosing pain from facial expressions can be helpful in clinical contexts where patients are unable to express themselves verbally, but has to achieve high accuracy results to ensure correct treatment [3]. Facial expressions are represented as sets of action units (AUs) describing facial muscle movements [2], an approach well established in the field of pain classification [3,8]. Taking into account not only the occurrence of AUs but also their succession within a sequence can increase diagnostic accuracy [5] and results in more realistic facial expressions for avatars [10]. While

* Funded by the DFG (German Research Foundation) – 405630557 (PainFaceReader).

earlier work represented AUs as points in time, we use a representation of AUs as time intervals with start and end points that can be identified in video frames with or without some ϵ-noise being allowed. Relations between these intervals can be expressed via the Allen calculus. Furthermore, negation (e.g. *AU10 does not occur in the sequence*) might be taken into account. AUs might be grouped according to the facial region they appear in, and additional personal details – e.g. age, sex, or a specific illness – could be included as background knowledge. We plan to explore these variants in order to achieve accuracy values which are acceptable for clinical applications, while ensuring a reasonable trade-off between accuracy and effort of model induction. The complexity of the learned models, that is, the number of rules and predicates within the rule bodies, should allow model decisions to be comprehensible for clinical experts.

Due to the restricted amount of available data, we generate synthetic data as additional data source for our investigations. In this paper, we present first results of an application of the ILP system Aleph [9] on synthetic data sets of AU sequences with and without underlying temporal patterns of varying complexity. We expect the learned theories to recover the original patterns used for sequence creation, and test the influence of different parameter settings on the results.

2 Methods

For the artificial data sets, we use an alphabet of AUs that are associated with the emotional state of pain (AU4, AU6, AU7, AU9, AU10, AU43). Every sequence is a set of six events, each being a different AU from the alphabet associated with a start and end point relative to the start of the sequence. For data sets without underlying temporal regularities, these points in time are chosen randomly within given boundaries concerning earliest start, latest end and minimum and maximum duration of an event[3]. For the data sets with temporal regularities, the positive examples also have to satisfy a rule that consists of one or more randomly chosen temporal relations between two randomly chosen AUs. The start and end points of the corresponding events are chosen such that all the temporal relations within the given rule are present in each positive example (updating the boundaries for event duration and position if necessary). The set of temporal relations used in the rules for the sequence creation and as candidate literals for Aleph's learning process is a simplified version of the original set, treating relations that rely on an exact matching of two or more points in time as instances of the most similar relation without this feature (see figure 1).

Data sets without temporal regularities have no rules (**R0**). The temporal regularities of the other data sets are represented as a theory consisting of 1, 2 or 3 rules with a complexity (= number of relations that have to be included in the positive examples) of 1, 2 or 3. All combinations {**R1-1**, **R1-2**, ..., **R3-3**} are tested. For each of them, 500 Sequences are created (200 positive and 200 negative training examples, 50 positive and 50 negative test examples). Each

[3] In order to approximate realistic values, these parameters were based on real AU events present in an unpublished database of manually decoded facial expressions.

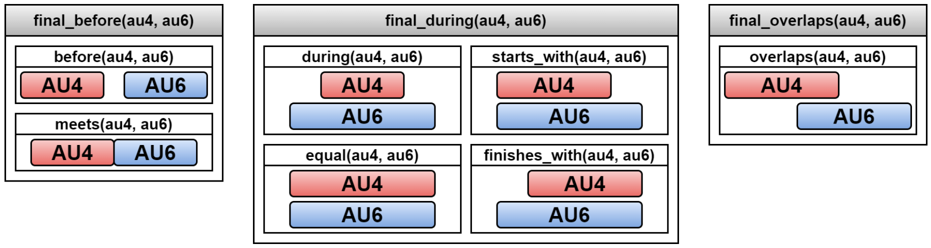

Fig. 1. Temporal relations between AUs. All relations grouped in one box are treated as the same type of relation (as labeled in the box heading) in the course of this study.

of the 10 training sets is used as input for Aleph's learning procedure, once with the Aleph parameter settings `minpos=2`, `noise=0` and once for `minpos=10`, `noise=10`. In each case, the procedure is repeated for five different random seeds in order to reduce any advantageous or disadvantageous random effects. All results presented in the following report the mean value over all five seeds.

3 Results and Discussion

Figure 2 shows the results for applying Aleph to the data sets. Considering the test accuracy and the number of learned rules, we encounter differing results that express a certain pattern: while the data sets without temporal regularities lead to the lowest test accuracy values and the highest number of learned rules, an increasing complexity of the rules defining the positive examples corresponds with higher accuracy values and fewer learned rules. Concurrently, an increasing number of underlying rules leads to lower accuracy values and a higher number of learned rules. With higher values for the parameters `minpos` and `noise`, the number of learned rules decreases considerably, while the test accuracy gets slightly better (the only exception being data sets without any underlying rules).

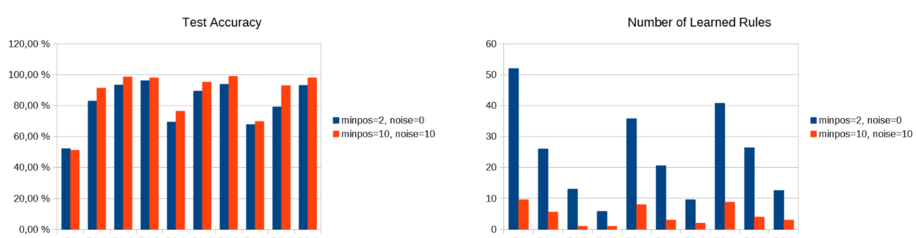

Fig. 2. Test accuracy and number of learned rules for artificial data sets. Underlying temporal regularities range from R0 (no rules) to R3-3 (3 rules each consisting of 3 required relations). All results show the average values over 5 different random seeds.

For the data sets with underlying temporal regularities, the learned rules resemble the original rules used for sequence creation, but often extend them with several variations of one or more additional relations that are actually noise which prevents false positives from being covered. This leads to an inflated number of learned rules where each original rule is represented in different alterations. The noise within the learned theory is reduced considerably in most cases when a certain number of false positives is admitted, allowing Aleph to recreate the exact set of rules that was used for the sequence creation in several cases. This is more likely to happen for rule sets with higher-complexity original rules.

4 Conclusion and Future Work

In general, Aleph seems to be a valid approach for exploiting temporal relations when represented using an interval-based algebra. In particular, this shows Aleph to be a valid option in principle when dealing with time series for purposes such as medical diagnoses, where comprehensibility and accuracy are of the utmost importance. In future work, we aim to integrate additional information from our specific use case into our representations, such as the facial regions that the individual action units belong to, or background information about the person a sequence is obtained from (e.g. age, sex, medical history), in order to determine the sweet spot regarding the accuracy-complexity-trade off.

References

1. Allen, J. F.: An Interval-Based Representation of Temporal Knowledge. IJCAI **81**, 221–226 (1981)
2. Ekman, P., Friesen, W. V.: Facial action coding system: Investigator's guide. Consulting Psychologists Press (2016)
3. Hassan, T., Seuß, D., Wollenberg, J., Weitz, K., Kunz, M., Lautenbacher, S., Garbas, J.-U., Schmid, U.: Automatic detection of pain from facial expressions: a survey. IEEE transactions on pattern analysis and machine intelligence **43**(6), 1815–1831 (2019)
4. Katzouris, N., Artikis, A., Paliouras, G.: Online learning of event definitions. Theory and Practice of Logic Programming **16**(5-6), 817–833 (2016)
5. Kunz, M., Lautenbacher, S.: The faces of pain: a cluster analysis of individual differences in facial activity patterns of pain. European Journal of Pain **18**(6), 813–823 (2014)
6. Moyle, S., Muggleton, S.: Learning programs in the event calculus. International conference on inductive logic programming, 205–212 (1997)
7. Ribeiro, T., Folschette, M., Magnin, M., Inoue, K.: Learning any semantics for dynamical systems represented by logic programs (2021). hal-02925942v4
8. Siebers, M., Schmid, U., Seuß, D., Kunz, M., Lautenbacher, S.: Characterizing facial expressions by grammars of action unit sequences–A first investigation using ABL. Information Sciences **329**, 866–875 (2016)
9. Srinivasan, A.: The Aleph manual (2001)
10. Tessier, M.-H., Gingras, C., Robitaille, N., Jackson, P. L.: Toward dynamic pain expressions in avatars: perceived realism and pain level of different action unit orders. Computers in Human Behavior **96**, 95–109 (2019)