

## FLAIRS-29 Poster Abstracts

Zdravko Markov and Ingrid Russell

*Editors*

### **Evaluation Technique for Argumentation Architectures from the Perspective of Human Cognition**

Abdulrahman Alqahtani, Marius Silaghi  
(Florida Institute of Technology, USA)

DirectDemocracyP2P (ddp2p.net) provides an environment for electronic debates, moderated in a decentralized way by ordering justifications (aka comments) based on collaborative filtering. The intuition was that comment threading models for such systems ought to have the peculiarity that each user should be allowed to submit a single justification when voting for a given motion, unlike common fora. To verify such intuitions, it is important to evaluate and compare the qualities of different available threading models with respect to specific applications. We propose to use a set of interviews to evaluate the "Understandability" of different threading models. We ask participants to navigate the same debates on three systems with similar graphical features and where the only difference consists of the alternative threading models being analyzed. Then participants list the arguments they remember. The debate systems with the desired threading models are obtained using the flexible user interface of DebateDecide (debatedecide.org) and the main challenge remained to port the same debates into the corresponding threading models. The solution undertaken in our study consists of selecting two debates from a well-organized televised session rich in arguments and interaction, but an unusual topic, such as to provide a level playground to the interviewed participants. The debates are further recast into the compared models by replaying them in the way expected from advanced users of the corresponding systems.

### **Activity Transition Detection by Relative Density Ratio Estimation**

Samaneh Aminikhanghahi, Diane Cook  
(Washington State University, USA)

Detecting activity breakpoints or transitions based on characteristics of observed sensor data from smart homes or mobile devices can be formulated as change point detec-

tion. Change points are abrupt variations in time series data and they are useful in segmenting activities, interacting with humans while minimizing interruptions, providing activity-aware services, and detecting changes in behavior that provide insights on health status. In this work, we consider the problem of detecting activity transitions using a non-parametric unsupervised machine learning technique, Relative unconstrained Least-Squares Importance Fitting (RuLSIF) and test the algorithm on an unscripted smart home data. RuLSIF is an algorithm to directly estimate the relative density ratio between two consecutive intervals of data using Pearson (PE) divergence as a dissimilarity measure. The ratio is used to detect changes in the underlying data distribution and the higher the ratio measure is, the more likely the point is a change point. The experiments conducted to evaluate RuLSIF on real-world smart home datasets show promising results with a true positive rate of greater than 70%. Our future study is to integrate change point detection algorithm with a smart home activity recognition and prompting system to test the effectiveness of the algorithm.

### **Practical 3D Tracking Using Low-Cost Cameras**

Roman Barták, Michal Koutný, David Odrzalek  
(Charles University in Prague, Czech Republic)

Tracking of objects in 3D space in general is not an easy task. There exist solutions involving hi-tech cameras and powerful computer systems capable of tracking many objects simultaneously (and changing the number of them) in a large dynamic space in real time. On the other hand, there are situations where such functionality is not necessary and the conditions may be specified in more detail, which makes the task significantly easier. This work falls in this second category. It shows the possibility to track a single object using low-cost cameras on an ordinary laptop in a small-scale and mostly static environment. This solution is useful for tracking a single object of interest in mobile robotics and particularly in the debugging phases, where the user needs to judge the robot movement system independently on what the robot claims.

## Application of Paraconsistent Logic to Technology

Jean-Yves Beziau (Tel Aviv University, Israel)

Paraconsistent logics are logics in which from a proposition  $P$  and its negation  $\neg P$ , it is not possible, contrary to what happens in classical logic, to deduce everything. Paraconsistent logic was mainly promoted by the Brazilian logician Newton da Costa who, together with Miro Quesada, coined the name. There are many applications of paraconsistent logic to science and technology, developed for example by V.S. Subrahmanian. In Brazil a robot called Emmy has been constructed who is able to cross a room with obstacles despite opposite information she is receiving. This can be illustrated with the following typical example. There is a transparent glass. According to a sensor based on visual information, there is for Emmy no obstacle. But according to another sensor based on tactile information there is an obstacle. If Emmy uses a program based on classical logic, she doesn't know what to do because everything is possible. But if Emmy uses a program based on paraconsistent logic, she can bypass the obstacle. The two sources of information are opposites, not contradictory, as it can be understood using the theory of the square of opposition. In this theory, initiated by Aristotle, three notions of opposition are distinguished: contradictories, contraries and subcontraries. In the present case, the two sources of information are contrary, not contradictory.

## Robustness in Abstract Argumentation Frameworks

Stefano Bistarelli, Francesco Faloci, Francesco Santini, Carlo Taticchi (Università di Perugia, Italy)

We introduce the notion of "robustness" in Abstract Argumentation Frameworks. We consider it as the property of an argumentation graph to withstand changes in terms of classical extension-based semantics. The robustness of a graph, with respect to a given semantics, is measured by computing the number of changes, for instance, addition and subtraction of arguments or attacks needed to change the corresponding extension. The graphs with the same semantics are considered part of the same equivalence class. In this way, we obtain a lattice of partially ordered graphs with the same semantics; in addition, a normal form is introduced to uniquely represent all the frameworks in the same equivalence class. The resistance to change is then analyzed and compared with the resistance to changes in belief revision, and the "unchangeable" sub-graphs of the frameworks are characterized with the purpose to identify the strong "clusters" of a debate. We then consider the classical semantics by restricting us to the unchangeable parts of a given framework. The relation between the re-

stricted and the extended argumentation problem is analyzed to better understand their difference and likeness. From this comparison, we outline some guidelines related to the computation of skeptically and credulously accepted arguments, depending on their frequency in such robust sub-graphs. Clearly, the link with persuasion (in Argumentation) is very thin, and all the proposed approaches can be reversed in order to better understand how easy is to persuade a (part of a) debate.

## Method to Validate Word Sense Disambiguation for Big Data

Alan Black (Drexel University, USA), Rosina Weber (Drexel iSchool, USA), Asta Zelenkauskaitė (Drexel University, USA)

The vast popularity of micro blogging platforms like twitter has created a need for big data approaches that are amenable to the analysis of short, often informal textual passages (microtext). Word sense disambiguation (WSD) can be a valuable tool for identifying relevant language within large microtext collections. The validation of WSD approaches has typically relied on a sense-tagged corpus created by human judges, the so-called Gold Standard. Unfortunately, creating gold standard corpora is an expensive and time-consuming undertaking, thereby limiting the practical size of gold standard corpora for use in WSD research. The methodology described in this poster leverages minimal human input in order to create sense tagged data on a large scale. The key insight upon which the methodology relies is the well-established notion of one-sense-per collocation. For example, the bigram "bee hive" is strongly indicative of the word "bee" referring to a flying insect and not a spelling competition. We use this finding to construct a corpus of sense-tagged tweets comprising Twitter data representing 58 bigrams mapped to 5 target words. The total run-time for data collection (which is rate limited by Twitter) and post-processing was under three days. The resultant corpus includes approximately a half million tagged tweets spanning one week. The creation of large-scale sense-tagged corpora enables the study of WSD approaches for big data generating platforms that never sleep.

## Evaluating and Enhancing Keystroke Dynamics Authentication

Youssef Bou Issa, Taline Boyajian, Maroun Kousseify (Antonine University, Lebanon)

Several biometric techniques such as face recognition, eye recognition are used nowadays in authentication systems. Keystroke dynamics is a recently developed biometric method that is based on the assumption that different people type in uniquely specific manners. In this work, our aim is to evaluate the authentication based on existing key-

stroke dynamics parameters while adding new parameters. To achieve this goal we created a mobile application in which we created a keystroke profile for each user. The user profile is based on specific parameters like the finger choice, the key hold time, the time between two consecutive keystrokes, the total time and the overall pressure on the screen. We designed three scenarios for the evaluation protocol in which we simulated the authentication procedures. Results show that the security based on keystroke dynamics authentication is very efficient. However, we noticed that with time, users tend to forget the way they type their password. For this reason we intend to add a learning/remembering phase in which we give users the possibility to get familiar with this new technique.

### **Distributed and Collaborative Sensing for Providing Situation and Option Awareness**

Tina Erlandsson (University of Skövde, Sweden),  
Carina Marcus, Per Boström (Saab AB, Sweden)

Teams of manned and unmanned aircraft have the opportunity to gather a huge amount of information regarding the situation at hand. Together they can cover larger areas as well as sensing the same objects from different directions enabling triangulation using combination of active and/or passive sensors as well as image recognition. Combining this data with intelligence, maps and weather information has the potential of providing both the teams and the commanders with situation awareness. This information should aid the decision makers to identify possible options and anticipate their consequences, which is also known as option awareness. A number of challenges need to be resolved in order to actualize this scenario. The commander who is interested in the entire battlespace needs to decide how to distribute the teams in air. Each team needs to collaborate to perform its tasks, taking into account the members' capabilities and resources, such as sensor and communication ranges. The teams should also request support from each other when needed. At all levels, there is a trade-off between gathering valuable information and avoiding risk exposure. In this initial work, we elaborate on the opportunities and challenges with distributed and collaborative sensing with the aim of providing situation and option awareness for both commanders and pilots.

### **Approximation Algorithms for Real-Time Homeomorphic and Isomorphic Tree Matching**

D. Michael Franklin (Kennesaw State University, USA),  
Xiaolin Hu (Georgia State University, USA)

In many common gaming and real-world scenarios players are trying to predict the behavior of the other player or

players. This assumes that there is some underlying strategy that these players are following, that such strategies can be inferred, and that a reasonable player can counter such strategies in real-time. These strategies may be modeled as various graph structures such as trees, finite state machines, or probabilistic graphical models. To infer which strategy another agent is following it is necessary to match prospective trees representing observed behaviors or policies with known trees representing previously learned strategies and policies. While matching two graphs can be done in super-linear time in the small scale to NP for the more complicated cases, isomorphic graph matching is much more complex. This is a well-known NP-Complete problem, especially when one considers homeomorphic and isomorphic tree matching. At this scale, the only way to accomplish such an algorithmic matching is to utilize highly-parallel processes. To accomplish strategy inference in real-time multi-agent systems there must be improved algorithms for tree matching. It has been shown that strategies offer significant performance enhancement to artificially intelligent agents, strategies can be recognized in real-time when complexity is limited, and AI's utilizing strategy inference will outperform their originally superior opponents, so discovering better approximation algorithms for this kind of tree matching is crucial to their success.

### **Flexible Machine Learning (ML-Flex) in the Veterans Affairs Clinical Personalized Predictions of Outcomes (Clinical3PO) System**

Lewis Frey, Leslie Lenert (Medical University of South Carolina, USA), Scott Duvall, Augie Turano (Veterans Affairs, USA), Brian Davis (MindFull Informatics LLC, USA), Michael Matheny (Vanderbilt University, USA), Jonathan Nebeker (Veterans Affairs, USA)

The flexible machine learning environment (ML-Flex) is deployed within the Clinical Personalized Pragmatic Predictions of Outcomes (Clinical3PO), which is an open source medical big data platform for predictive analytics funded by the National Institutes of Health. The ML-Flex and Clini-cal3PO platform enables secure large-scale, reproducible machine learning analysis of medical data. To support interoperability and reproducibility, Clinical3PO uses the open Observational Medical Outcomes Partnership (OMOP) data model. It is deployed and tested behind the firewall of the U.S. Department of Veterans Affairs (VA) Informatics and Computing Infrastructure (VINCI) using data from the VA's Corporate Data Warehouse and in a test bed platform deployed to Amazon Web Services (AWS) using non-VA publicly available data. The platform includes a collection of tools developed in such companies as Google, Facebook and Amazon such as Hadoop, Yarn, and Accumulo. Clinical3PO's scalable and secure framework can smooth transitions from small data solu-

tions to comprehensive internet-scale solutions that have the potential to create a paradigm shift to personalized and individualized care based on the information gathered from large cohorts of patients. A key feature of Clinical3PO is its ability to analyze patients based on sequences of clinical processes. This feature helps meet healthcare organizations' analytical needs for lean management and near-real-time point-of-care decision support for complex and unusual patients.

## **Hybrid Approaches to Community Detection for Recommendation**

Robert Frye, David Wilson (UNC Charlotte, USA)

Detecting clusters within network graphs is an important task in the development of recommender systems. Finding tightly connected clusters within user-to-user and item-to-item graphs is the foundation for a variety of approaches to improve the accuracy of item recommendations. In this research we examine hybrid approaches to cluster and community detection by evaluating techniques that combine the Infomap and Louvain algorithms with the newest iteration of OSLOM community detection. We investigated combinations of Infomap and Louvain approaches with OSLOM detection on a substantial real-world citation dataset. Results show that hybrid approaches outperformed baseline clustering performance, and that the hybrid OSLOM-Infomap approach provided superior community detection. Based on our results, we present a novel recommendation approach that takes advantage of the hybrid techniques.

## **Identifying Condition-Action Statements in Medical Guidelines Using Domain-Independent Features**

Hossein Hematiam, Wlodek Zadrozny  
(University of North Carolina at Charlotte, USA)

Medical guidelines establish criteria regarding diagnosis, management, and treatment in specific areas of healthcare. The most important parts of the guidelines describe actions of medical personnel to be performed under certain conditions. We propose and evaluate a supervised machine learning model extracting condition-action sentences. We establish baselines for hypertension and rhinosinusitis guidelines, as well as compare our method against previously done work on asthma guidelines. We use part of speech (POS) tags as features in our classification models. Using POS tags instead of semantic pattern rules makes our model more domain-independent, and therefore more suitable for establishing baselines, not only for text mining of medical guidelines but also in other domains, such as text mining of business rules. We annotated part of asthma

guidelines used previously other researchers, hypertension and rhinosinusitis guidelines to create gold standards to measure the performance of our condition-action extracting models. We classified them into three classes: condition-action, condition-consequence (effect, intention, and event) and action. We used ZeroR, Naive Bayes, J48, and random forest classifiers were applied in our project. The results show that generally random forest classifier seems to work best in extracting Condition-Action statements. Our work established baselines for automated extraction of condition-action rules from medical guidelines, but its performance is still inferior to a collection of manually created extraction rules. We are currently augmenting our model with semantic information and evaluating improvements based on "events" vs. "entities".

## **Discovering Students' Outcomes Based on Their Interactions in Online Learning Platform**

Fazel Keshtkar, Jordan Cowart, Andrew Crutcher, Ben Kingen  
(Southeast Missouri State University, USA)

Online learning platforms such as Moodle are popular in higher education. Moodle provides information that are useful in obtaining student learning models. One available type of information is in the form of Moodle logs of students' interaction with one another and their instructor. These interactions contain various activities such as: participation in forum discussion, frequency of account login, and other useful activities. In this paper, we apply statistical techniques to discover the students' trails in Moodle based on their interaction and their final grades. We use data from undergraduate computer science and computer information systems courses which contained an online component. At the beginning, we group students by their interactions with Moodle and their final grades obtained in the course. We then, conduct statistical analysis with students and their interactions separately in order to obtain the relation with their interaction and their final outcomes. We assume an interaction is any login to course website in Moodle and perform some type of activity such as: browsing the grade book, posting or reading a message, observing new material, submitting a homework, taking a test or quiz, and other similar activities provided by the platform to the learner. Our goal in is to provide a learning model to help online learning platforms and educators to assist their learners more effectively, potentially helping detect students that are falling behind. Our results shows that there is some degree of relation between interaction and final outcomes.

## **Analysis of Word Order in Multiple Treebanks**

Vladislav Kubon, Marketa Lopatkova, Jiří Mirovský  
(Charles University in Prague, Czech Republic)

This poster gives an overview of the results of automatic analysis of word order in 23 dependency treebanks. These treebanks have been collected in the frame of in the HamleDT project, whose main goal is to provide universal annotation for dependency corpora and thus it also makes it possible to use identical queries for all corpora. The queries are written in a PML-TQ query language. The analysis concentrates on the basic characteristic of the word order, the order of three main constituents, a predicate, a subject and an object. The quantitative analysis is performed separately for main clauses, subordinated clauses and subordinated clauses containing active verbs. This detailed analysis makes it possible to capture the differences between the word order of main and subordinated clauses and between subordinated clauses with active and passive verbs. The analysis also makes it possible to discover inconsistencies or errors in the annotation if the actual results differ from the expected ones. A measure of word order freedom based upon the (un)even distribution of individual word order types is suggested. It exploits the natural idea that if a language has absolutely free word order, the individual types of word order will be evenly distributed, while for a language with fixed word order one type (e.g. SVO) will strongly prevail.

## **Personal Name Recognition in Sumerian Texts Using a Sign-Based HMM**

Yudong Liu (Western Washington University, USA)

In this work we described a named-entity recognition system to identify personal names in Sumerian cuneiform documents from the Ur III period. As Sumerian has a syllabic, sign-based written form, we adopted a sign-based HMM. Our motivation is to investigate the potential of sign-based  $n$ -th order HMMs to accurately identify personal names in this corpus with minimal contextual information and no prior knowledge regarding the language for feature selection as would generally require human domain experts, and attempt to determine empirically the order of Markov model giving the optimal results. In our experiment, we have access to a well-annotated source corpus made available by CDLI and ORACC. In the HMM we constructed, each state corresponds to a part-of-speech tag (such as personal name, profession name, geographic name, etc., with 12 part-of-speech tags in all) provided in the corpus, and each emission corresponds to an observable sign occurred in the training data. We train the HMM by constructing transition and emission matrices from this

data, observing the frequency with which a sign  $w_i$  is emitted by state  $t_i$  (where  $0 < i < N$ ,  $N$  is the number of signs in the training corpus), and identify the most likely state sequence for the testing sentence using the standard Viterbi decoding. We trained both a bigram model and several higher order  $n$ -gram ( $2 \leq n \leq 7$ ) models, and applied different smoothing techniques to compensate for data sparsity. Results show that bigram model achieves F1 score of 84%, which is the most effective among all.

## **Establishing Extensible Evaluation Metrics from Crowdsourced Data**

Sarah Luger (University of Edinburgh, UK)

Online education resources are now plentiful as educators connect the opportunities associated with augmenting conventional teaching approaches with the technologies used in online community building, gaming, and crowdsourcing. Cutting edge web-based tools and platforms produce valuable student performance data that is anonymously re-used to improve the quality, evaluation, and efficacy of Multiple Choice Question tests. New metrics combine established methodologies developed in Educational Theory's Item Response Analysis (IRT) with human decision making strategies tested in crowdsourced data collection. While there are new perspectives on developing open source educational tools, the underlying questions that motivate good exam development and curriculum coverage are unchanged: What makes a Multiple Choice Question difficult? What questions best differentiate students into performance groups? Further, automatically identifying the question answering behaviors of good students has value in improving the performance of all students. The new evaluation methodology uses IRT to automatically identify the best performing students and the questions that best sort students into performance cohorts. Related to a similar IRT-based crowdsourcing methodology proposed by Christoforaki and Ipeirotis (2014), this work extends those findings to new types of questions and analyses. This methodology aids "intelligent textbook" development by refining question scope and tailoring the questions to student aptitude. This research contributes to the public domain knowledge base for evaluation (an area dominated by proprietary algorithms) and also is useful for improved testing of crowdsourced worker quality in online labor markets. Future work includes automating test performance analysis in real time and developing active learning-based approaches.

## **Designing a Fitness Function for a Human-Like Pac-Man Controller Based on Neuroevolution**

Maximiliano Miranda, Antonio A. Sánchez-Ruiz,  
Federico Peinado (Complutense University of Madrid, Spain)

As part of our research in Artificial Intelligence for simulating the behavior of human players we have found that a Neuroevolution approach performs well both in terms of the instrumental similarity measure and the phenomenological evaluation by human spectators. Neuroevolution is a form of machine learning that employs Evolutionary Algorithms (EAs) to train Artificial Neural Networks (ANNs) by considering the weights of the ANN as the chromosomes of the EA. A fitness function should be used for this purpose, which is determinant for the final results of the algorithm. This function tries to characterize the human likeness of the automatic game controller. Designing this fitness function is a complex task and depends on the specific game. In this work we use Ms. Pac-Man, a version of the classic arcade game. We have compared two different approaches: in the first one we use raw data extracted directly from human traces (the set of directions chosen by the human player in the maze and their timestamps); in the second one we complement this information with more elaborated features such as recklessness (distance to the closest ghost when Pac-Man escapes) or aggressivity (when Pac-Man chases the ghosts). Our goal is to measure the importance of these abstract features to replicate the behavior of a human player even in simple games like Pac-Man.

## **A Real-Time N-Gram Approach to Choosing Synonyms Based on Context**

Brian Moore, Robert Mercer  
(The University of Western Ontario, Canada)

Synonymy is an important part of all natural language but not all synonyms are created equal. Because two words are synonymous doesn't necessarily mean that they can always be interchanged. The problem that we address is that of near-synonymy and choosing the appropriate near-synonyms based purely on the surrounding words. Our computational method, unlike previous methods developed for this problem, is capable of making multiple word suggestions (which more accurately models human choice), it contains millions of contexts, it does not require training, and it runs in real-time (system is currently capable of running in real-time up to approximately 120 words a minute on a mid-tier commercial computer). This poster also presents new synonym sets and human annotated test data that we believe better represents the problem that we are studying.

## **Evolutionary Strategies for Rays Cycle Mining**

John Ross, James Buckley  
(University of Dayton, USA)

Cycles are of use in identifying particularly interesting patterns in data. A cycle is a set of mined rules where the antecedent of one rule is the consequent of another one. The entire chain forms a cycle. RAY cycles are standard cycles where one non-cycle rule's consequent "enters" the cycle and the antecedent of one of the cycle rules is the consequent ("spawn") of a non-cycle rule. Genetic Algorithms, and more specifically, evolutionary strategies are well-suited for data mining applications in general. They are especially useful in the detection and exploitation of cycles in large data sets and their multidimensional genome representations could accommodate most any data type. Effective mutation and recombination strategies allow them to adapt to emerging data cycles. Furthermore, they could provide a temporal component or predictive aspect to otherwise static data.

## **Improvement of Emotion Detection Based on Hidden Markov Models**

Romeo Saad, Elio Gebrayel, Taline Boyajian, Youssef Bou Issa  
(Université Antonine, Lebanon)

The science of body language has increasingly become an evaluation method for social and emotional conditions of a human; it is also called Kinesics. This means that the interpretation of a nonverbal communication shows a person's psychological state. Finger pointing, nail biting, head shaking, blinking and many other facial expressions and body gestures determine persuasiveness of a speaker undergoing an experimental test. In this work, our aim is to enhance the possibilities to automate the analysis of the body language in order to have a reliable extended analysis. Our previous study was performed in this issue and based on an analytical study using the conditional probability calculus with posture detection using Kinect. Thus, our aim was to enhance the performance of the probabilistic results by using the Hidden Markov Models (HMM) for more accurate modeling. The Markov process is represented by unobserved (hidden) states which are the person's psychological states, and with observed states called "observations" collected from the body language. Our prototype is able to register these "observations", along with their duration in order to calculate the probability of the person's state(s) undergoing the experimental test. During the total duration of the test, the specific hidden emotional states related to the detected observations are given with their probabilities obtained by our system. The extension of the study consists of including the image processing functionalities provided

by “Kinect” in order to automatically detect the observations of the body movements and analyze the body language of the tested subject.

### **Generating Word Problems Similar to a Given Math/Physics Word Problem**

Savitha Sam Abraham, Sowmya S. Sundaram  
(Indian Institute of Technology, Madras, India)

Given a word problem, we present a method to automatically generate other problems similar to it. Our domains comprise of physics word problems involving projectile motion and math word problems involving addition/subtraction. We use a knowledge based approach to solve the problem. While solving, we extract the problem’s characteristics such as type and difficulty. Here, ‘type’ refers to the model that has been used. For projectile problems, we first generate a qualitative model describing the problem scenario. We then add the given quantitative facts and solve. In this process, we can identify features like the model used, whether implicit information was required and whether an intermediate calculation, such as finding the value for an unasked quantity, was needed. Similarly, for math word problems, we identify the schema(s) which is/are applicable and which quantity is to be solved. By difficulty, we refer to the average sentence length, presence of extraneous information and the kind of numbers (decimal/non-decimal, 3-digit/2-digit, etc.) present. With this information at hand and some natural language templates, we can generate word problems with the same type and difficulty. Here, we need to check the correctness of the problem that has randomly-generated numerical values within the domain of the quantities. We can present the generated word problems in natural language using an existing Natural Language Generation (NLG) toolkit. We can also provide an option to change the level of difficulty. With this, we hope to help students in mastering a concept.

### **Representing Adaptive Course Navigation in the Generalized Intelligent Framework for Tutoring**

Robert Sottolare, Keith Brawner  
(US Army Research Laboratory, USA)

This poster explores the use of Markov Decision Processes (MDPs) in support of adaptive course navigation in the Generalized Intelligent Framework for Tutoring (GIFT). GIFT is an open source architecture for authoring and evaluating Intelligent Tutoring Systems (ITSs) and adaptive course navigation is an AI-based technique which considers attributes of the learner and the instructional context to select actions which will optimize learning, retention, and transfer of training to work environments. GIFT’s cur-

rent adaptive course navigation model is decision tree-based. The adaptive course navigation model put forth in this poster aligns closely with the principles of MDPs where a user’s current state, possible actions and a reward function determine movement to a future state. Unlike decision trees used which are currently used in GIFT, MDPs also account for uncertainty in the assessment of learner states and thereby support more adaptive options than regimented decision trees. We put forth a model for the use of MDPs in lieu of decision trees in GIFT-based ITSs and provide an exemplar to illustrate the differences and benefits that MDPs might offer over the current course navigation model.

### **Inconsistent Knowledge Integration with Bayesian Network**

Yi Sun, Yun Peng  
(University of Maryland Baltimore County, USA)

Given a Bayesian network (BN) representing a probabilistic knowledge base of a domain, and a set of low-dimensional probability distributions (also called constraints) representing pieces of new knowledge coming from more up-to-date or more specific observations for a certain perspective of the domain, we present a theoretical framework and related methods for integrating the constraints into the BN, even when these constraints are inconsistent with the structure of the BN due to different dependencies among relevant variables. Our framework allows for identifying the structural inconsistencies between the constraints and the existing BN. We use the disjunction method to find out dependency relations in the BN, and use the dependency test to find out dependency relations in each constraint. The dependency relations that exist in the constraint but are missing in the BN are the cause of the structural inconsistencies. Our framework also allows for overcoming the identified structural inconsistencies during the knowledge integration by changing the structure of the BN. This is done by adding one node for each structural inconsistent constraint in a way similar to the use of the virtual evidence node in Pearl’s virtual evidence method. Other more computationally efficient variations of adding node methods are also considered. Compared to the existing methods for probabilistic knowledge integration, our methods have the advantage of being able to integrate new dependencies into the existing knowledge base as constraints become available from more reliable sources.

## **Combining Hand-Crafted Rules and Machine Learning Methods in Opinion Target Identification**

Kateřina Veselovská (Institute of Formal and Applied Linguistics,  
Charles University in Prague, Czech Republic)

Aleř Tamchyna (Charles University in Prague, Czech Republic)

Opinion target identification, i.e. the task in which the evaluated entities need to be identified in natural language texts, is one of the main subtasks in sentiment analysis. In addition to the polarity classification task where we try to identify polarity of the evaluation, it is important to be able to identify which entities the evaluation concerns. Both in academia and industry, this task is usually addressed using statistical methods. However, for morphologically rich languages, such as Czech, it can still be advantageous to combine machine learning methods with hand-crafted rules, taking into account linguistic features of the evaluative structures. This poster shows how we created the rules for opinion target identification based on evaluative structures automatically found in Prague Dependency Treebank and combined them with the machine learning methods, namely linear-chain conditional random fields. The results show that employing the rules increased the performance significantly. Also, we compare the results for Czech data with the state-of-the-art results achieved in aspect-based sentiment analysis in English, using the data provided by SemEval.