Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

IEEE Int Conf Healthc Inform. Author manuscript; available in PMC 2019 June 01.

Published in final edited form as:

IEEE Int Conf Healthc Inform. 2018 June ; 2018: 36–43. doi:10.1109/ICHI.2018.00012.

Intention Mining in Medical Process: A Case Study in Trauma Resuscitation

Sen Yang,
Electrical and Computer Engineering Department Rutgers University Piscataway, NJ 08854, USA

Weiqing Ni,
Electrical and Computer Engineering Department Rutgers University Piscataway, NJ 08854, USA

Xin Dong,
Electrical and Computer Engineering Department Rutgers University Piscataway, NJ 08854, USA

Shuhong Chen,
Electrical and Computer Engineering Department Rutgers University Piscataway, NJ 08854, USA

Richard A. Farneth,
Division of Trauma and Burn Surgery, Children's Nat'l Medical Center Washington, DC 20010, USA

Aleksandra Sarcevic,
College of Computing and Informatics Drexel University Philadelphia, PA 19104, USA

Ivan Marsic, and
Electrical and Computer Engineering Department Rutgers University Piscataway, NJ 08854, USA

Randall S. Burd
Division of Trauma and Burn Surgery, Children's Nat'l Medical Center Washington, DC 20010, USA

Abstract

In medical processes such as surgical procedures and trauma resuscitations, medical teams perform treatment activities according to underlying invisible goals or intentions. In this study, we present an approach to uncover these intentions from observed treatment activities. Developed on top of a hierarchical hidden Markov model (H-HMM), our approach can identify multi-level intentions. To accurately infer the H-HMM, we used state splitting method with maximum a posteriori probability (MAP) as the scoring function. We evaluated our approach in both qualitative and quantitative ways, using a case study of the trauma resuscitation process. This dataset includes 123 trauma resuscitation cases collected at a level 1 trauma center. Our results show our intention mining achieved an accuracy of 86.6% in classifying medical teams' intentions. This work is an exploration of unsupervised intention mining of complex real-world medical processes.

Keywords

Intention Mining; Process Mining; Trauma Resuscitation; Hierarchical Hidden Markov Model

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

I. Introduction

Process mining techniques [1] use activity logs to discover process models and analyze process deviations. The process models are often designed as a step-by-step workflow, specifying what executions to perform at each step. The limitation of existing studies in the process mining domain is that they only focused on the analysis of low-level activities. Most medical processes, however, are problem-solving processes, rather than a workflow execution process like issuing a driver's license or submitting an insurance claim in which the exact workflow with all steps is precisely specified. Research has found that adverse outcomes are most commonly due to process errors that occur when an incorrect intention is formed by the team, leading to incorrect actions [2]. In these cases, the input data are correctly perceived, but an incorrect intention is formed, and the wrong activity is performed. Only monitoring and analyzing low-level process executions can be insufficient for understanding medical processes. In this study, we present an approach to mine the intentions from low-level medical team executions. We tested our approach on a real-world trauma resuscitation process dataset.

Intentions are the thoughts directed towards achieving process goals. We can often infer people's intentions from their activities [3]. For example, during the trauma resuscitation, "maintain oxygenation" is a goal. Activities for addressing this goal may include "placing oxygen" or "placing an oxygen saturation monitor." The team may have the goal of "maintain oxygenation," but not have the intention of achieving this goal for several minutes. When they do intend to satisfy this goal, their intention may be identified by observing the two oxygen-related activities. Previous workflow analyses in process mining have focused on the shallow mining of the patterns of observed activities. They did not attempt to understand the hidden (or unobservable) intentions behind the observed activities. In addition, most previous research has focused on simple processes that include only a limited number of activities. Performing these analyses in medical settings is more challenging because of the concurrency of associated activities.

Although it is possible to manually perform intention mining, this approach is vulnerable to bias. Experienced observers may provide useful domain insights, but may also be prone to identifying only familiar or expected patterns while neglecting others. Mining intentions by reviewing activity lists based on patient attributes could mitigate this bias, but is an approach that is labor intensive. For this reason, a data-driven intention discovery approach is attractive for complex processes.

We perform data-driven unsupervised intention discovery by finding patterns in the observed activities. our intention discovery is based on the hypothesis that observed activities are frequently associated with the underlying intentions. We also assume that activities related to the same intention will be temporally associated with each other. We used an HMM inference algorithm to extract these associations automatically. In addition, because no general criteria are available to assess the quality of trauma resuscitations across patients with various injuries and conditions, our intention mining method mainly focuses on common process executions rather than the rare pathways.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Our contributions are:

- *A novel intention mining approach* using a hierarchical hidden Markov model (H-HMM) to represent intentions at multiple levels of granularity. We introduced a state-splitting approach that avoids subjective and labor-intensive initialization of model parameters (e.g., number of hidden states, transition probabilities).
- *A case study on a real-world medical process.* We applied our intention mining approach on the trauma resuscitation process and extracted a multi-level intention model. Our evaluation includes both qualitative reviews from medical experts, and a quantitative accuracy of 86.6% with an F_1-score of 0.87.

II. Trauma Resuscitation Data and Formalization

The collection and use of the data for this study were approved by the Institutional Review Board at our hospital. One hundred and twenty-three trauma resuscitation cases were coded manually from video recordings collected at the trauma bay of Children's National Medical Center, Washington DC between Aug. 2014 and Apr. 2016. The dataset includes 3058 resuscitation activities of 17 different types in the secondary survey of the trauma resuscitation (a procedure for assessing a patient head to toe for injury) (TABLE I).

The process log $\boldsymbol{O} = [\boldsymbol{O}_1, \ldots, \boldsymbol{O}_n]^T$ is a vector of n execution traces. A **execution trace** is $\boldsymbol{O}_i = [a_1, \ldots, a_k]^T$. where k total activities a of the i^{th} trace are ordered by activity start time. Traces of different process executions for trauma resuscitation may have varying lengths due to different patient conditions and the treatment strategies adopted.

An **intention** $I_i^{(\ell, p)}$ is defined as a hidden (or unobservable) goal, objective, or motivation that can be achieved by a group of activities. Intentions can exist on multiple levels. Low-level intentions are called subintentions or sub-subintentions. We use l to denote the l-th level of intention, i to denote the intention's index at the l-th level, and p to denote the index of a parent intention at the (l–1)-th level. At the highest level (l = l), p = *null*.

A **hidden Markov model** $\lambda = (\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi}, \boldsymbol{Q}, \sum)$ models the temporal sequences through hidden states. $\boldsymbol{A}$ is the state transition probability matrix, so $t_{ij} \in \boldsymbol{A}$ represents the transition probability between states i and j, $\boldsymbol{B}$ represents each state's observation probability distribution, $\boldsymbol{\pi}$ is a vector of the initialized state distribution, $\boldsymbol{Q} = [q_1, \ldots, q_{|\boldsymbol{Q}|}]^T$ is the vector of hidden states, and $\sum = [e_1, \ldots, e_{|\sum|}]^T$ is the emission alphabet (i.e., observed activities). During trauma resuscitations, we can directly observe the team activities, but not their intentions. The unobservable intentions need to be inferred from the observed activities. We used hidden states of an HMM to model these intentions. To model the hierarchical intention structure, we adopted a hierarchical HMM. Higher-level intentions (or states) are composed of lower-level intentions, and the lowest-levels of intentions are carried out by observable activities (Figure 1).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

III. Mining Intentions

We mined intentions by inferring a HMM. To model intentions in a scenario as complex as trauma resuscitation, we used an HMM inference model with several enhancements. First, to avoid the random guessing of the model topology, we used an advanced model inference approach, state-splitting algorithm [4][5][6][7], instead of the Baum-Welch method. Second, we used maximum a posteriori probability (MAP) scoring [8] to guide the state-splitting and control model complexity. Third, to model intentions at different levels of granularity, we introduced a recursive algorithm for hierarchical HMM discovery.

A. MAP State-Splitting HMM

State splitting works by initializing a generic HMM (e.g., a single state) and successively splitting states (Eq. 1) until any further splits cannot increase the score (Eq.2). All candidate states for splitting are split and the best split found is used in the next iteration i of the model:

$$\lambda_i = \underset{\lambda_{i-1}^q}{\text{argmax}} \text{Score}(\lambda_{i-1}^q, \boldsymbol{O}) \quad s.t. \quad q \in \boldsymbol{Q} \quad (1)$$

where λ_{i-1}^q is the candidate model obtained by splitting state q in the model λ_{i-1}. The scoring function Score(λ, $\boldsymbol{O}$) quantifies how well the model λ fits the observations $\boldsymbol{O}$, balanced against the model complexity penalty [4][5][6]. The algorithm terminates when further splitting does not increase the score:

$$\text{Score}(\lambda_i, \boldsymbol{O}) \leq \text{Score}(\lambda_{i-1}, \boldsymbol{O}) \quad (2)$$

Too many split states may lead to model overfitting, trading representativeness for accuracy. Splitting too little may lead to underfitting and less accurate models. We used the maximum a posteriori probability [8] scoring function to guide splitting. This method has several advantages over existing complexity metrics such as BIC [5], MDL [6], and heuristic approaches [6]. In our experiments, other metrics led to either over-penalizing (preventing any splitting) or imbalanced splitting (where only a few states retain most emissions). MAP does not have these problems as it has a comprehensive penalty over the model complexity, penalizing structural complexity and biases in parameter distribution (Eq. 3).

The model prior $P(\lambda)$ is composed of structure priors $P(\lambda_S)$ (i.e., prior probability distribution over all possible model topologies specified as a set of states, transitions and emissions) and parameter priors $P(\theta_\lambda)$ (i.e., prior probability distribution of transitions and emissions).

$$P(\lambda) = P(\lambda_s)P(\theta_\lambda \mid \lambda_s) \quad (4)$$

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Let $P(\lambda_G)$ denote the prior for the global aspect (the number of states) of the model structure (i.e., prior probability distribution of a topology with a given number of states). $P(\lambda_G)$ is assumed to unbiased and can be ignored in the maximization.

$$P(\lambda) = P(\lambda_G) \prod_{q \in \boldsymbol{Q}} P\left(\lambda_s^{(q)} \middle| \lambda_G \right) P\left(\theta_\lambda^{(q)} \mid \lambda_G, \lambda_s^{(q)} \right) \quad (5)$$

where $P(\lambda_s^{(q)})$ is a prior for the structure associated with state q. and $P(\theta_\lambda^{(q)} \mid \lambda_G)$ is a prior for the parameters of state q. The structure here includes the non-zero paths (transitions) and outputs (emissions) that are associated with state q. The parameters here refer to the probability values of transitions and emissions that are associated with state q. We adopted the narrow parameter priors defined in [8] for structure (Eq.6) and parameter (Eq.7) of state q.

$$\begin{aligned} & P\left(\lambda_s^{(q)} \middle| \lambda_G \right) \\ & = p_t^{n_t^{(q)}} (1 - p_t)^{|\boldsymbol{Q}| - n_t^{(q)}} p_e^{n_e^{(q)}} (1 - p_e)^{|\boldsymbol{\Sigma}| - n_e^{(q)}} \end{aligned} \quad (6)$$

where $n_t^{(q)}$ is the estimated number of outgoing transitions from state q and $n_e^{(q)}$ is the number of its emissions. $\overline{n_t}$, $\overline{n_e}$ are the expected (i.e., average) number of transitions and emissions per state. $p_t = \overline{n_t} / |\boldsymbol{Q}|$ is the probability of a potential transition's existence, and $p_e = \overline{n_e} / |\boldsymbol{Q}|$ is the probability of a possible emission.

$$\begin{aligned} & P\left(\theta_\lambda^{(q)} \middle| \lambda_G, \lambda_s^{(q)} \right) \\ & = \frac{1}{B(\alpha_t, \ldots, \alpha_t)} \prod_{i=1}^{n_t^{(q)}} \theta_{qi}^{(\alpha_t - 1)} \frac{1}{B(\alpha_e, \ldots, \alpha_e)} \prod_{j=1}^{n_e^{(q)}} \theta_{qj}^{(\alpha_e - 1)} \end{aligned} \quad (7)$$

where we used *Dirichlet distribution* as the prior distribution over the model parameters (transition probabilities $\boldsymbol{\theta}_t^{(q)} = \left[\theta_{q_1}, \ldots, \theta_{qn} \right]^T$ and emission probabilities $\boldsymbol{\theta}_e^{(q)} = \left[\theta_{q_1}, \ldots, \theta_{qn} \right]^T$) in a model structure $\lambda_s^{(q)}$. We chose the Dirichlet distribution because it is a standard prior used in multinomial models and because it is mathematical convenient [10]. The normalizing constant beta function $B(\alpha, \ldots, \alpha)$ [11] produces more balanced states (i.e., states containing similar number of non-zero emissions) and avoids imbalanced states (i.e., states that greatly differ in the number of non-zero emissions) in the intention model. We set the prior weights α_t and α_e to 2 in the beta function to guide the model parameters towards 0.5.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

B. Hierarchical HMM Inference

The hierarchical HMM (H-HMM) can be considered as a tree of HMMs. Each node of the tree is a HMM inferred recursively using the MAP-SS inference method, in a top-down fashion (Alg.1). The initial HMM topology is set to three states (one for start, one for end, and one split-able state in between) (Step 1). At the first level, we infer the first intention model $\hat{\lambda}^1$ using MAP-SS (Step 2). We then find a subset of observations $\boldsymbol{O}_i^s \subseteq \boldsymbol{O}$ that can be emitted from hidden state q_i in the intention model $\hat{\lambda}^1$. This discovery uses the Viterbi algorithm [9], which finds the optimal state sequence associated with the observed activity sequence. A lower level intention model $\hat{\lambda}_i^{\ell+1}$ is then recursively inferred based on $\boldsymbol{O}_i^s$ (Step 3). The recursion terminates at levels where the complexity does not allow any more splitting. During post-processing, the inferred model $\boldsymbol{\lambda}^H$ is smoothed to flatten the emission probability distribution so that all traces can occur with some probability (Step 4).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Algorithm 1. Hierarchical HMM Inference

Input: $\boldsymbol{O}$, ω_1, ω_2

Output: $\boldsymbol{\lambda}^H$

/* Initialization */

Step 1. Initialize HMM topology λ_0 as three states (a single state of observations and two functional states, start and end);

/* When intention level $l = 1$ (top level) */

Step 2. Infer top level of intention model: $\hat{\lambda}^1$ = MAPSS ($\boldsymbol{O}$, ω_1, ω_2);

/* When intention level $l > 1$, do recursive inference */

Step 3. $\boldsymbol{\lambda}^H$ = RecursiveInference($\hat{\lambda}^1$, $\boldsymbol{O}$);

Step 4. Smooth the model: $\boldsymbol{\lambda}^H$ = Smoothing($\boldsymbol{\lambda}^H$);

Step 5. **return** $\boldsymbol{\lambda}^H$;

Function: RecursiveInference (inferred λ^l, observation $\boldsymbol{O}^s$)

Step 1. Find subsequence $\boldsymbol{O}_i^s \subseteq \boldsymbol{O}^s$ that can be observed in hidden state q_i, $\left\{\boldsymbol{O}_i^s\right\} = \text{Viterbi}\left(\boldsymbol{O}^s, \lambda^\ell\right)$;

Step 2. **for** each state q_i in λ^ℓ, **do**

Step 3.

Step 4. Infer subintention model: $\hat{\lambda}_i^{\ell+1} = \text{MAPSS}(\boldsymbol{O}_i^s, \omega_1, \omega_2)$;

Step 5. **If** number of states in subintention model $\left|\boldsymbol{Q}_i^{\ell+1}\right| > 3$

Step 6.

Step 7. RecursiveInference($\hat{\lambda}_i^{\ell+1}$, $\boldsymbol{O}_i^s$);

end if

Step 8. **end for**

$\lambda^H = \lambda^H \cup \left\{\lambda^\ell\right\}$;

end Function

Function: Smoothing (inferred model $\boldsymbol{\lambda}^H$)

Step 1. **for** each λ in $\boldsymbol{\lambda}^H$, **do**

Step 2. **for** each state q_i in λ, **do**

Step 3. Let $|\boldsymbol{\Sigma}|$ denote the number of emissions in q_i;

Step 4. **for** each emission e in q_i, **do**

Step 5. **if** $\boldsymbol{B}(e) \neq 0$, **do**

Step 6. $\boldsymbol{B}(e) = \boldsymbol{B}(e) \cdot (1 - 0.003) + 0.003 \,/\, |\boldsymbol{\Sigma}|$ / / 0.003 is selected based on three-sigma rule of thumb;

Step 7. **else**

Step 8. $|\boldsymbol{B}(e) = 0.003 \,/\, |\boldsymbol{\Sigma}|$;

Step 9. **end if**

Step10. **end for**

Step 11. **end for**

Step12. return λ^h;

end Function

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

the source code is available at https://github.com/allen9408/IntentionMining

C. Parameter Selection

$\omega_1$is the weight (or hyperparameter) that controls the inferred model's complexity and topology. ω_1 is 1 by default. A smaller ω_1 leads to larger, deeper models (Figure 2). As ω_1 increases, the depth of the intention model drops, the total number of states and transitions decreases, and the average number of activities per lowest-level state increases. We favor a hierarchical model that is smaller, simpler, and easily-interpretable for labeling purposes.

IV. Empirical Study

We evaluated our intention mining system from two aspects. First we evaluated the discovered intention model, checking whether our data-driven approach can discover meaningful intentions. We also assessed the accuracy of the intentions classified by our method versus the ground truth labelled by medical experts.

A. Experimental Metrics

We adopted accuracy and multi-class weighted-average F_1 score to measure the compliance between algorithm-derived intention and human-labelled intention. Specifically, the accuracy measures the percentage of correct intention estimation based on all estimations made. Another metric, the multi-class weighted-average F_1 score, calculates metrics (precession and recall) for each class and finds their average, weighted by the support $|C_i|/n$ (the percentage of true instances in class c_i out of all instances).

$$F_1 = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision+recall}} \quad (8)$$

with multi-class weighted-average precision defined as:

$$\text{precision} = \sum_{i=1}^{l} \frac{|c_i|}{n} \cdot \frac{tp_i}{(tp_i + fp_i)} \quad (9)$$

and multi-class weighted-average recall as:

$$\text{recall} = \sum_{i=1}^{l} \frac{|c_i|}{n} \cdot \frac{tp_i}{(tp_i + fn_i)} \quad (10)$$

where l is the number of multi-classes c_i $(1 \le i \le l)$. tp_i are true positives for c_i, tn_i are true negatives, fn_i are false negatives, and fp_i are false positives. Compared with accuracy, the weighted-average F_1 score takes class imbalance into account.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

B. Intention Model Evaluation

We discovered the intention model in an unsupervised way, under the assumption that sequential relationships between the observed activities are correlated with medical team intentions. To validate our assumption and evaluate our intention model, we conducted two different experiments. First, we asked medical experts whether they were able or not to manually assign the discovered intentions (Figure 3, Figure 5) with meaningful labels. Second, we provided medical experts with 40 randomly selected cases that include 1074 activities and asked them to manually label the activities with top-level (level-1) intentions (Figure 5). We then compared the manually-derived labels to algorithm-derived labels.

1) Qualitative Evaluation—From an engineering perspective, the emission distribution of low-level intentions is sparse (matrix in Figure 5). This phenomenon is a result of the nature of our trauma resuscitation data. Many resuscitation activities have a strict or strong association with certain intentions. For example, activity “left-spine-bk” (Figure 5) is only associated with intention $I_{13}^{(2,5)}$ “assessment of lumbar spine.” For other activities, a strict or strong association with a specific intention is not observed. For example, “palpation-head” is not only correlated with level-1 intention “assessment of head and face” but also the level-1 intention “assessment of back and posterior aspect of the head and extremities.” This happens because the back of the head is exposed once, when the medical team turn the patient to assess the back. Assessing the back of the head allows the examining provider to complete the head exam without excessive force to the neck or cervical spine. The medical experts commented that the intention model (Figure 6) correctly captured the intentions associated with medical tasks. Associated activities grouped by the model were largely in line with the expectations of medical experts. In addition, the model correctly identifies the sequential correlations of related intentions. For example, head → neck → chest → abdomen, which reflects the same resuscitation execution order that medical experts usually perform.

One limitation is that the model only provided a limited description of provider intentions. It is unable to distinguish between deviations and novel process executions. For example, medical experts were unable to assign some level 2 and level 3 intentions under the level 1 intention of “assessment of back and posterior aspect of the head and extremities.” Most of the activities that occur under this intention take place while the patient is rolled to the side exposing the back. Given this context, medical experts were unable to determine why the ear examination using otoscope was included because it is typically performed while the patient is supine. The lack of an available explanation, however, should not be considered evidence of no association. Future work is needed to clarify data-driven results.

2) Quantitative Evaluation—The intention labeling (algorithm vs. manual) results (Figure 4) show high intention mining accuracy of 86.6% and F_1-score of 0.87, indicating the feasibility of data-driven algorithms for intention mining. The misclassification mainly occurs at the cells (1), (2) and (3). Cells (1) and (2) represent activities of assessment of head & face (“H”) and assessment of extremities (“Ex”), which were misclassified into intention assessment of neck, chest & ears (“N”). The most likely cause of this misclassification is the proximity of one body region to another. The head, neck and shoulders (upper extremities)

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

are located close to each other and providers will frequently move between these regions during an assessment. When multiple intentions occur simultaneously or interact with each other, it can be challenging for the algorithm to accurately identify the current intention. Cell (3) shows the activity of “H” misclassified as “Bk” (assessment of back). This confusion can be understood because the head and back exams as the back of head are commonly assessed during the back exam.

V. Related Work

Our study aims to design a decision support system to monitor and analyze the trauma resuscitation process in real time. We were able to automatically collect low-level activities using sensors. Process mining techniques [1][3] can discover process errors by checking the compliance between the activity sequence and expert-based workflow models. While some process errors arose from erroneous performance, others were found to be triggered by incorrect intentions formed by the medical team [2][12][13][14].

Research in intention mining is still in its infancy and mainly lies in the domain of information retrieval [15][16][17][22]. Jethava et al [15] presented a generic framework for learning the multi-dimensional representation of user intent from the query words. Ricardo et al and Markus et al [16][17] developed some approaches for the identification of user’s interest in an automatic way, based on the analysis of query logs. Other applications have also been proposed, e.g. human-computer interaction [18][20][21] and content analysis [19]. Kim et al [18] proposed a deep intention estimation and recommendation system by understanding human attention from gestures. Mei et al [19] presented a novel method to model and mine the captured intentions of camcorder users based on digital video recorders and home video data. Oliver et al [21] developed a generic model including connections between intentions, actions, and sensor measurements through a Hybrid Dynamic Bayesian Network. In contrast to these studies, we attempted to mine the intention from activity logs collected from a medical process. To our best knowledge, intention mining has not been applied in medical process analysis. Most research based on activity logs has been from a research single group [23][24]. This group has presented a supervised and an unsupervised intention mining approach, termed the Miner Method (MMM). They applied their methods on software usage activity logs to uncover user intentions. Compared to our hierarchical model, their work models only a single level of attention, and requires a predetermined number of intentions.

VI. Conclusion

We present a process intention mining method and applied it on the real-world process. We used a hierarchical hidden Markov model to mine multi-level intentions from low-level process executions performed by the medical team performing the process of trauma resuscitation. To accurately infer the hierarchical hidden Markov model, we adopted state-splitting inference and selected maximum a posteriori probability as the inference-guiding objective function. The inferred intention model is simple and interpretable. The hierarchical structure with small models at each node allows easier bottom-up interpretation and analysis. We evaluated our method in two different ways, requesting feedback from medical

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

experts and quantifying the intention mining accuracy. We have shown that the discovered intention groupings aligned with medical expert knowledge.

Acknowledgement

This paper is based on research supported by National Institutes of Health under grant number 1R01LM011834-01A1.

References

[1]. Van Der Aalst Wil M.P. Process mining: discovery, conformance and enhancement of business processes. Springer Science & Business Media, 2011.

[2]. Gruen RL, Jurkovich GJ, McIntyre LK, Foy HM, Maier RV. Patterns of errors contributing to trauma mortality: lessons learned from 2,594 deaths. Ann Surg. 2006 9;244(3):371–80. PubMed PMID: ; PubMed Central PMCID: PMC1856538. [PubMed: 16926563]

[3]. Deneckere Rebecca, et al. "Intentional process mining: discovering and modeling the goals behind processes using supervised learning." International Journal of Information System Modeling and Design 5.4 (2014): 22–47.

[4]. Ostendorf Mari, and Singer Harald. "HMM topology design using maximum likelihood successive state-splitting." Computer Speech & Language 11.1 (1997): 17–41.

[5]. Siddiqi Sajid M., Gordon Geoffrey J., and Moore Andrew W.. "Fast state discovery for HMM model selection and learning." International Conference on Artificial Intelligence and Statistics. 2007.

[6]. Mavromatis Panayotis. "Minimum description length modelling of musical structure." Journal of Mathematics and Music 3.3 (2009): 117–136.

[7]. Yang Sen, et al. "Medical Workflow Modeling Using Alignment-Guided State-Splitting HMM." Healthcare Informatics (ICHI), 2017 IEEE International Conference on. IEEE, 2017.

[8]. Stolcke Andreas, and Omohundro Stephen M.. "Best-first model merging for hidden Markov model induction." arXiv preprint cmp-lg/9405017 (1994).

[9]. Rabiner LR. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE, vol.77, no.2, pp. 257–286, 2 1989.

[10]. Buntine Wray. "Theory refinement on Bayesian networks." Uncertainty Proceedings 1991. 1991 52–60.

[11]. Ferguson Thomas S. "A Bayesian analysis of some nonparametric problems." The annals of statistics (1973): 209–230.

[12]. Clarke JR, Spejewski B, Gertner AS, Webber BL, Hayward CZ, Santora TA, Wagner DK, Baker CC, Champion HR, Fabian TC, Lewis FR, Jr, Moore EE, Weigelt JA, Eastman AB, Blank-Reid C. An objective analysis of process errors in trauma resuscitations. Acad Emerg Med. 2000 11;7(11):1303–10. PubMed PMID: . [PubMed: 11073483]

[13]. Pucher Philip H., et al. "Identifying and addressing preventable process errors in trauma care." World journal of surgery 37.4 (2013): 752–758. [PubMed: 23340709]

[14]. Webman Rachel, et al. "Classification and team response to non-routine events occurring during pediatric trauma resuscitation." The journal of trauma and acute care surgery 81.4 (2016): 666. [PubMed: 27648769]

[15]. Jethava Vinay, et al. "Scalable multi-dimensional user intent identification using tree structured distributions." Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval. ACM, 2011.

[16]. Baeza-Yates R, Calderon-Benavides R & González-Caro C (2006). The intention behind web queries In Crestani F, Ferragina P & Sanderson M (Eds.), Proceedings of the 13th International Conference on String Processing and Information Retrieval (pp. 98–109). Berlin, Heidelberg: Springer.

[17]. Strohmaier Markus, and Kröll Mark. "Acquiring knowledge about human goals from search query logs." Information Processing & Management 48.1 (2012): 63–82.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[18]. Kim Sangwook, et al. "Intention estimation and recommendation system based on attention sharing" International Conference on Neural Information Processing. Springer, Berlin, Heidelberg, 2013.

[19]. Mei Tao, et al. "Modeling and mining of users' capture intention for home videos." IEEE transactions on multimedia 9.1 (2007): 66–77.

[20]. Howard Newton, and Cambria Erik. "Intention awareness: improving upon situation awareness in human-centric environments." Humancentric Computing and Information Sciences 3.1 (2013): 9

[21]. Schrempf Oliver C., and Hanebeck Uwe D.. "A generic model for estimating user intentions in human-robot cooperation." ICINCO. 2005.

[22]. Sadikov Eldar, et al. "Clustering query refinements by user intent." Proceedings of the 19th international conference on World wide web. ACM, 2010.

[23]. Khodabandelou G, Hug C, Deneckère R, & Salinesi C (2014, 5). Unsupervised discovery of intentional process models from event logs. In Proceedings of the 11th Working Conference on Mining Software Repositories (pp. 282–291). ACM.

[24]. Khodabandelou G, Hug C, Deneckere R, & Salinesi C (2013, 5). Supervised intentional process models discovery using hidden markov models. In Research Challenges in Information Science (RCIS), 2013 IEEE Seventh International Conference on (pp. 1-11). IEEE.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The MAP State-splitting (MAP-SS) HMM Inference: Given: A set of observed process traces $\boldsymbol{O} = [\boldsymbol{O}_1, \ldots, \boldsymbol{O}_n]$. **Objective:** Successively splitting states to find an HMM topology λ that maximizes the posterior probability $P(\lambda|\boldsymbol{O})$:

$$\hat{\lambda}_{MAP}(\boldsymbol{O}) = \underset{\lambda}{\operatorname{argmax}} P(\lambda \mid \boldsymbol{O}) = \frac{P(\lambda)^{\omega_1} P(\boldsymbol{O} \mid \lambda)}{P(\boldsymbol{O})} \quad (3)$$

where $P(\boldsymbol{O}|\lambda)$ is observation sequence probability, solved with the forward algorithm [9]. The model prior $P(\lambda)$ can be considered as model complexity penalty. The weight (hyperparameter) ω_1 is used to control model complexity. Observation $P(\boldsymbol{O})$ is fixed for a given data point, so it can be ignored in the maximization.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

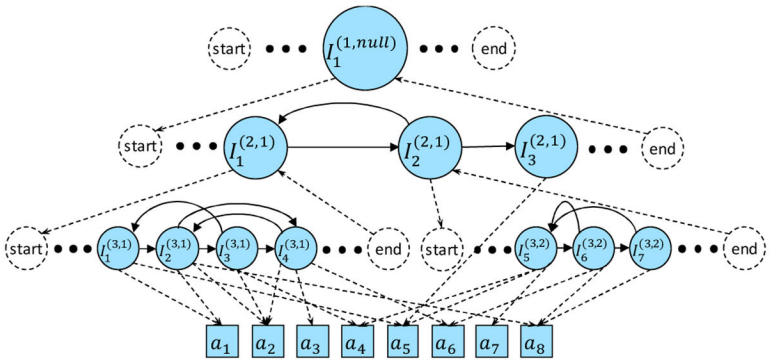

Figure 1.

An example of hierarchical intention model with three levels. Higher-level intentions are composed of lower-level intentions. The lowest levels of intentions are carried out by observable activities.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

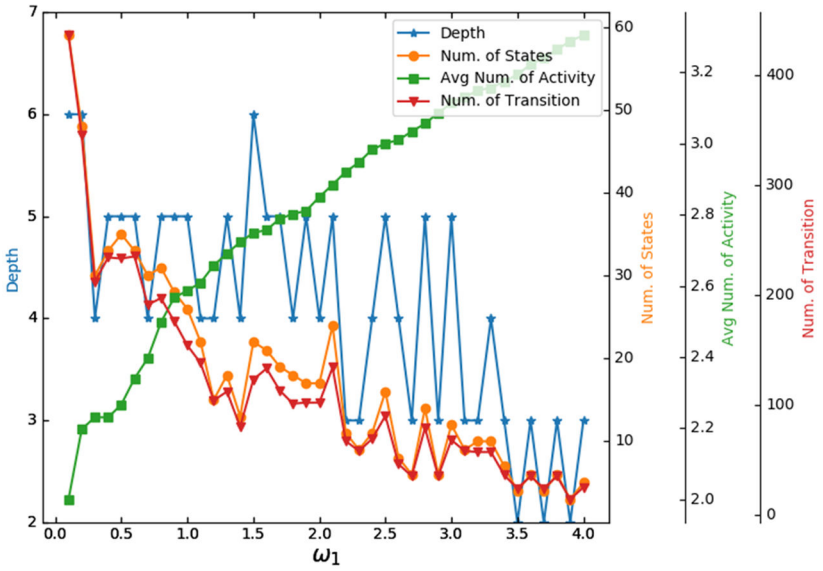

Figure 2.

Model topology and complexity changes according to hyperparameter ω_1. The features of model toplogy include the depth of the inferred hierarchical HMM (blue), the number of states (orange) in the model, the avg. num. of activities (green) per lowest-level intentions before smoothing, and the number of transitions in the model.

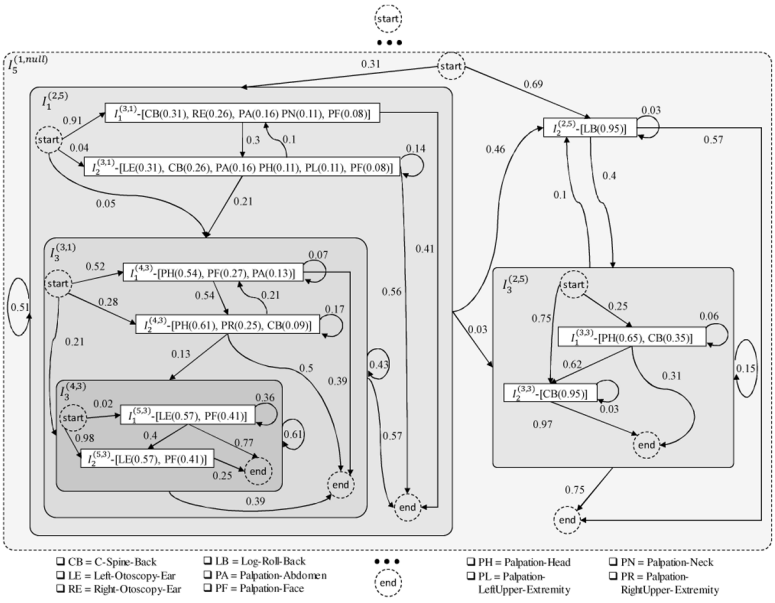

Figure 3.

Part of a discovered multilevel intention model. Each box represents an unnamed intention as a group of subintentions or observable activities.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

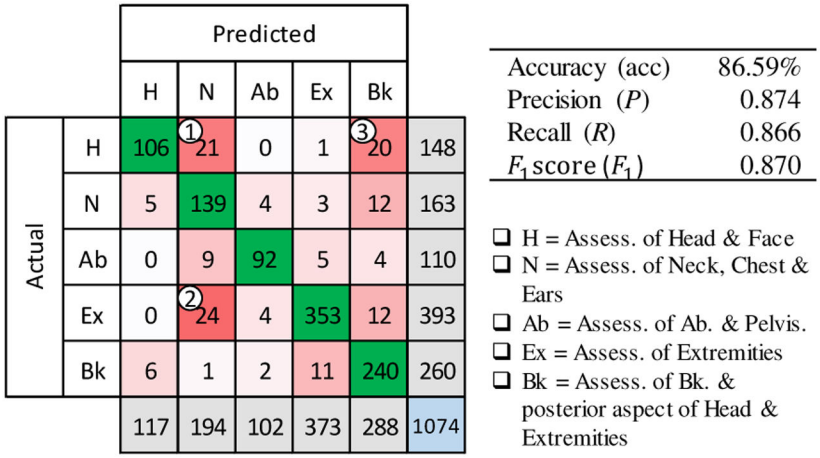

Figure 4.

Confusion matrix for algorithm-derived (predicted) intentions vs. hand-labelled (actual) intentions. The five intentions are the level 1 intentions in (Figure 5).

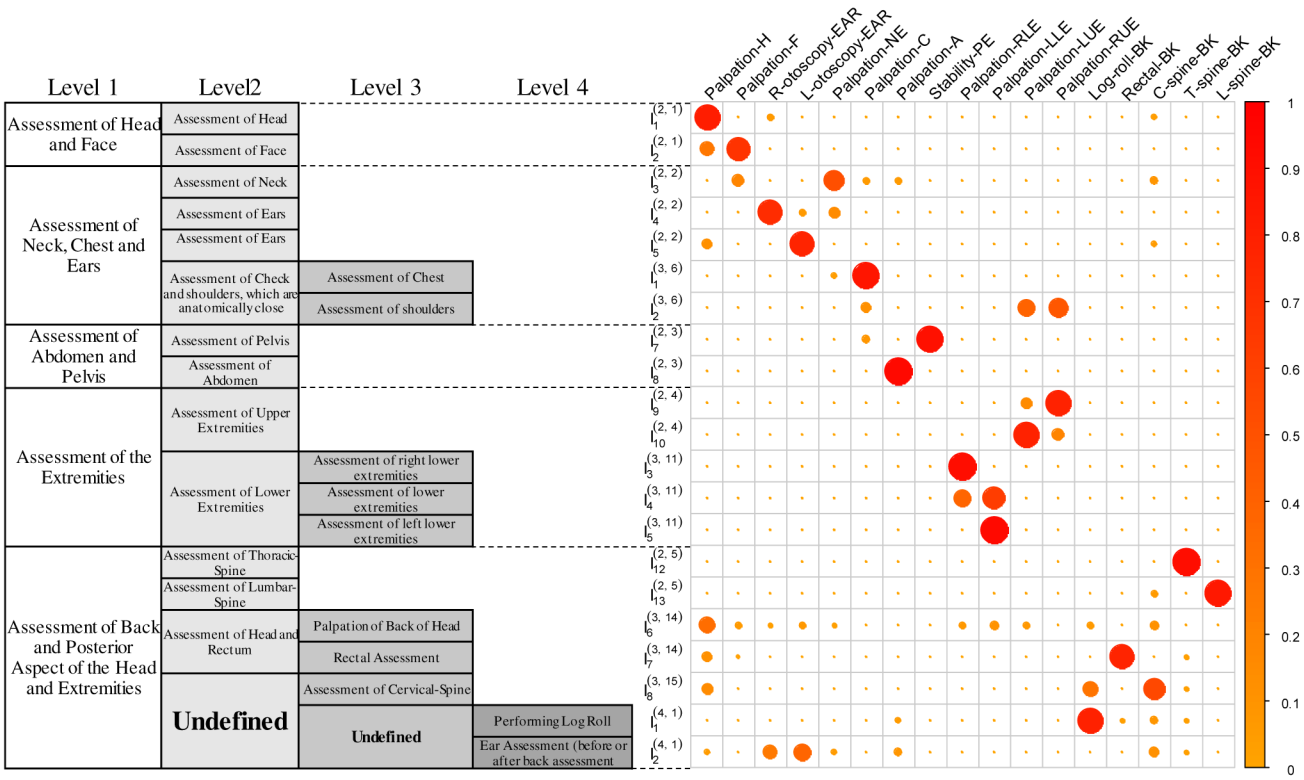

Figure 5.

Multi-level intentions (left) labeled by medical experts and the corresponding emission matrix (right) of the intention model (Figure 6). The column of the matrix represents activity type and the row represents the lowest level of intentions. The size and color of the dots in the matrix represents the value of emission probability.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

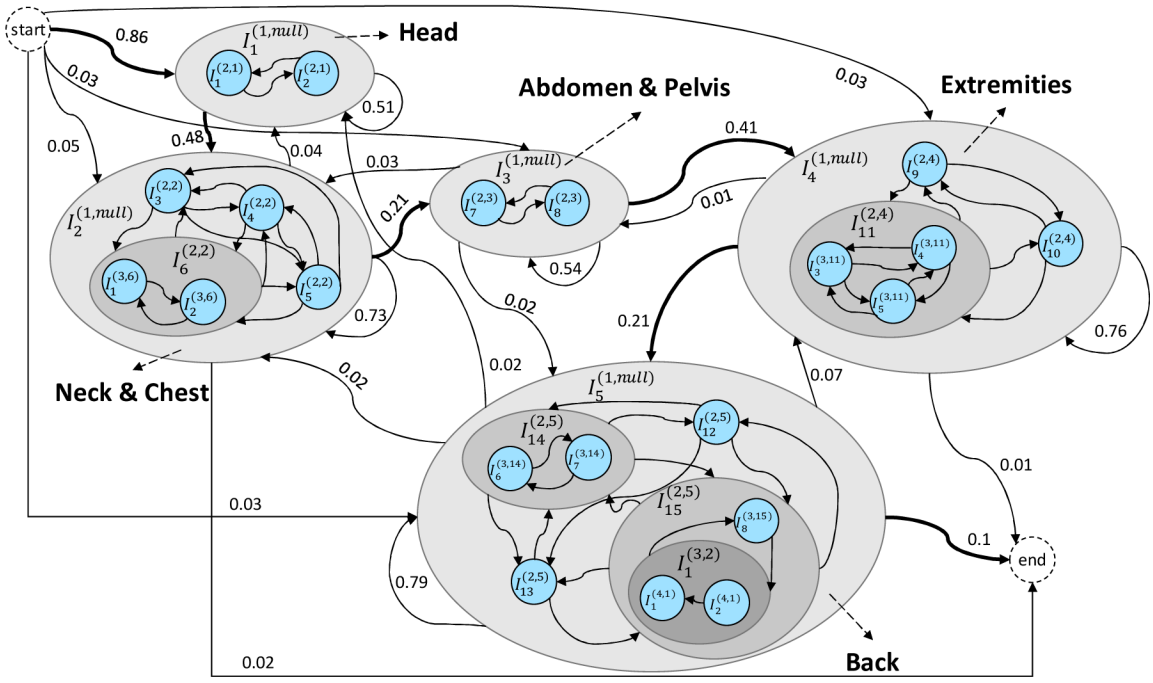

Figure 6.

Discovered intention model. The model is trained based on sample data with 17 secondary survey activities in 123 resuscitations. The transition probabilities for low-level intentions are not shown. The emission probabilities for low-level intentions (blue circles) are shown in (Figure 5). Thick transitions highlighted are the most dominating outgoing transitions (excluding self-transitions) between first level intentions.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Algorithm 1. Hierarchical HMM Inference

Input: $\boldsymbol{O}$, ω_1, ω_2

Output: $\boldsymbol{\lambda}^H$

/* Initialization */

Step 1. Initialize HMM topology λ_0 as three states (a single state of observations and two functional states, start and end);

/* When intention level $l = 1$ (top level) */

Step 2. Infer top level of intention model: $\hat{\lambda}^1$ = MAPSS ($\boldsymbol{O}$, ω_1, ω_2);

/* When intention level $l > 1$, do recursive inference */

Step 3. $\boldsymbol{\lambda}^H$ = RecursiveInference($\hat{\lambda}^1$, $\boldsymbol{O}$);

Step 4. Smooth the model: $\boldsymbol{\lambda}^H$ = Smoothing($\boldsymbol{\lambda}^H$);

Step 5. **return** $\boldsymbol{\lambda}^H$;

Function: RecursiveInference (inferred λ^l, observation $\boldsymbol{O}^s$)

Step 1. Find subsequence $\boldsymbol{O}_i^s \subseteq \boldsymbol{O}^s$ that can be observed in hidden state q_i, $\left\{\boldsymbol{O}_i^s\right\} = \text{Viterbi}\left(\boldsymbol{O}^s, \lambda^\ell\right)$;

Step 2. **for** each state q_i in λ^ℓ, **do**

Step 3.

Step 4. Infer subintention model: $\hat{\lambda}_i^{\ell+1} = \text{MAPSS}(\boldsymbol{O}_i^s, \omega_1, \omega_2)$;

Step 5. **If** number of states in subintention model $\left|\boldsymbol{Q}_i^{\ell+1}\right| > 3$

Step 6.

Step 7. RecursiveInference($\hat{\lambda}_i^{\ell+1}, \boldsymbol{O}_i^s$);

end if

Step 8. **end for**

$\lambda^H = \lambda^H \cup \left\{\lambda^\ell\right\}$;

end Function

Function: Smoothing (inferred model $\boldsymbol{\lambda}^H$)

Step 1. **for** each λ in λ^H, **do**

Step 2. **for** each state q_i in λ, **do**

Step 3.

Step 4. Let $|\boldsymbol{\Sigma}|$ denote the number of emissions in q_i;

Step 5. **for** each emission e in q_i, **do**

Step 6. **if** $\boldsymbol{B}(e) \neq 0$, **do**

$\boldsymbol{B}(e) = \boldsymbol{B}(e) \cdot (1 - 0.003) + 0.003 \,/\, |\boldsymbol{\Sigma}|$ / / 0.003 is selected based on three-sigma rule of thumb;

Step 7. **else**

Step 8. $\boldsymbol{B}(e) = 0.003 \,/\, |\boldsymbol{\Sigma}|$;

Step 9. **end if**

Step10. **end for**

Step 11. **end for**

Step12. return λ^h;

end Function

the source code is available at https://github.com/allen9408/IntentionMining

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

TABLE I.

DATA SAMPLE (A) AND STATISTICS (B)

Case ID	Activity	Start Time	End Time
xx1	Palpation_Head	0:03:50	0:03:56
xx1	Palpation_Face	0:04:15	0:04:24
xx1	Palpation_Neck	0:04:22	0:04:29
xx1	C_spine_Back	0:04:37	0:04:40
xx1	Palpation_Chest	0:05:11	0:05:16
xx1	Palpation_Abdomen	0:05:14	0:05:16
xx1	Stability_Pelvis	0:05:28	0:05:30
xx1	Palpation_Abdomen	0:05:49	0:06:09

(a) Trauma resuscitation trace

Properties	Stats
Num. Cases (or Patients)	123
Num. Total Activities	3058
Num. Activity Types	17
Data Time Period	2014.08 – 2016.04
Size of Medical Team	[7, 12]
Longest Trace (Num. Acts.)	51
Shortest Trace (Num. Acts.)	7
Avg. Num. Acts. in Traces	25

(b) Data statistics