Logic and Commonsense-Guided Temporal Knowledge Graph Completion

Guanglin Niu[1], **Bo Li**[1,2]
[1] Institute of Artificial Intelligence, Beihang University, Beijing, China
[2] Hangzhou Innovation Institute, Beihang University, Hangzhou, China
{beihanggngl, boli}@buaa.edu.cn

A Appendix

A.1 Description of the Temporal Rule Patterns

(1) $p_2(x, y, t+t_1) \Leftarrow p_1(x, y, t)$. With regard to a length-1 rule, the body atom $p_1(x, y, t)$ and the head atom $p_2(x, y, t+t_1)$ occur at different times. This temporal rule pattern signifies that two events containing predicates p_1 and p_2 would occur one after the other if both the subject and object are the same, respectively. For instance, given a temporal rule $Consult\,(x,\ y,\ t+t_1) \ \Leftarrow\ Express\ intent\ to\ meet\ or\ negotiate\,(x,\ y,\ t)$ and an event $(China, Express\ intent\ to\,meet\ \ or\ negotiate, U.S.A., TS_1)$ occurring at the timestamp TS_1, we could easily deduce that an event $(China,$ $Consult, U.S.A., TS_1 + T)$ will happen after TS_1.

(2) $p_2(x, y, t) \Leftarrow p_1(x, y, t)$. For a length-1 rule, the body atom $p_1(x, y, t)$ and the head atom $p_2(x, y, t)$ occur simultaneously. This pattern is equivalent to the length-1 static rule.

(3) $p_3(x, y, t + t_1 + t_2) \ \Leftarrow\ p_1(x, z, t) \wedge p_2(z, y, t + t_1)$. The three atoms in this pattern of temporal rule occur at different time from each other. This pattern implies the three causally related events occur one after the other. For instance, the temporal rule $Express\,intent\,to\,meet\,or\,negotiate(x,\ y,\ t{+}t_1{+}t_2) \Leftarrow Engage\ in\ negotiation^{-1}\,(x,\ z,\ t) \wedge Express\ intent\ to\ cooperate\,(z,\ y,\ \ t+t_1)$ belongs to this temporal rule pattern, where the predicate $Engage\ in\ negotiation^{-1}$ is the inverse version of the predicate $Engage\ in\ negotiation$.

(4) $p_3(x, y, t + t_1) \ \Leftarrow\ p_1(x, z, t) \wedge p_2(z, y, t)$. Specific to this rule pattern, the two events corresponding to the body atoms $p_1(x, z, t)$ and $p_2(z, y, t)$ such as (s, p_1, e, t) and (e, p_2, o, t) occur at the timestamp t, and then the event $(s, p_3, o, t + t_1)$ associated with the atom $p_3(x, y, t + t_1)$ would happen.

(5) $p_3(x, y, t) \ \Leftarrow\ p_1(x, z, t) \wedge p_2(z, y, t)$. Similar to the pattern (2), this pattern of the temporal rule indicates that all the events corresponding to the atoms occur at the same time, which is equivalent to the length-2 static rule.

A.2 Evaluation Criteria of Length-2 Temporal Rules

For the length-2 candidate temporal rule in the general form of $p_3\,(x,\ y, t+T_1+T_2) \Leftarrow p_1\,(x,\ z,\ t) \wedge p_2\,(z,\ y, t+T_1)$, there are three diverse patterns with the different values of T_1 and T_2. In allusion to each timestamp t, the SD, SC and HC are defined as:

$$SD = \begin{cases} \sum_{t_{1i}\in[1,\ W_T]} \sum_{t_{2i}\in[1,\ W_T]} \#(a,c) : p_1\,(a,b,t) \\ \quad \wedge\, p_2\,(b,c,t+t_{1i}) \wedge p_3\,(a,c,t+T_{1i}+T_{2i})\,, \\ \hfill T_1 > 0, T_2 > 0 \\ \sum_{t_{2i}\in[1,\ W_T]} \#\,(a,c) : p_1\,(a,b,t) \wedge p_2\,(b,c,t) \\ \quad \wedge\, p_3\,(a,c,t+t_{2i})\,, \qquad T_1 = 0, T_2 > 0 \\ \#\,(a,c) : p_1\,(a,b,t) \wedge p_2\,(b,c,t) \wedge p_3\,(a,c,t)\,, \\ \hfill T_1 = T_2 = 0 \end{cases} \tag{18}$$

$$SC = \begin{cases} \frac{1}{|W_T|} \cdot \frac{SD}{\sum_{t_{1i}\in[1,\ W_T]} \#(a,c):p_1(a,b,t)\wedge p_2(b,c,t+t_{1i})}, \\ \hfill T_1 > 0 \\ \frac{SD}{\#(a,c):p_1(a,b,t)\wedge p_2(b,c,t)}, \qquad T_1 = 0 \end{cases} \tag{19}$$

$$HC = \begin{cases} \frac{SD}{\sum_{t_{1i}\in[1,W_T]} \sum_{t_{2i}\in[1,W_T]} \#(a,c):p_3(a,c,t+t_{1i}+t_{2i})}, \\ \hfill T_1 > 0, T_2 > 0 \\ \frac{SD}{\sum_{t_{1i}\in[1,W_T]} \#(a,c):p_3(a,c,t+t_{1i}+t_{2i})}, \\ \hfill T_1 = 0, T_2 > 0 \\ \frac{SD}{\#(a,c):p_3(a,c,t)}, \qquad T_1 = T_2 = 0 \end{cases} \tag{20}$$

It is noteworthy that the temporal rule with the settings (1) $T_1 > 0, T_2 > 0$, (2) $T_1 > 0, T_2 = 0$, (3) $T_1 = T_2 = 0$ correspond to the temporal rule patterns (3), (4) and (5), respectively.

A.3 The Algorithm of Our Proposed Temporal Rule Learning Module

For better understanding, Algorithm 1 summarizes the whole procedure of our temporal rule learning procedure.

A.4 Proof of Representing Causality Among Events via RGPR Mechanism

In this paper, we take the temporal rule patterns (1) and (3) as examples and provide the lemma proofs that the temporal rule-guided predicate embedding regularization $\mathbf{G}$ could represent the causality among events.

Lemma 1. On account of the temporal rule pattern (1) $p_2(x, y, t{+}t_1) \Leftarrow p_1(x, y, t)$, the temporal rule-guided predicate embedding regularization $\mathbf{G}$ specific to Eq. 5 could represent the causality between two events.

Proof. Given an event (s, p_1, o, t), we could obtain the causality term of the time-sensitive score specific to this event that is transformed at the timestamp $t + t_1$ according to Eq. 10 and the time transfer operator $\mathbf{T}$, which can be written as

$$\mathrm{Re}\left(\mathbf{s}^{\top}\, diag(\mathbf{T} \circ \mathbf{p}_{r1})\, \bar{\mathbf{o}}\right) = 1 \tag{21}$$

According to $\mathbf{G}$ defined in Eq. 5, the association between $\mathbf{p}_{r1}$ and $\mathbf{p}_{r2}$ satisfies:

$$\mathbf{T} \circ \mathbf{p}_{r1} = \mathbf{p}_{r2} \tag{22}$$

Substituting Eq. 22 into Eq. 21, we can derive that

$$\mathrm{Re}\left(\mathbf{s}^{\top}\, diag(\mathbf{p}_{r2})\, \bar{\mathbf{o}}\right) = 1 \tag{23}$$

It can be discovered that the event $(s, p_2, o, t + t_1)$ holds if (s, p_1, o, t) occurs. Therefore, the causality between two events occurring at different timestamps could be represented via the temporal rule-guided predicate embedding regularization $\mathbf{G}$ of the temporal rule pattern (1) defined in Eq. 5. □

Algorithm 1: our temporal rule learning module.

Input: $\mathcal{G}_s$, $\mathcal{G}_t$: the GSKG and the TKG
  T, W_T: the number of timestamps and the size of time window
  T_{sc}, T_{hc}: the thresholds of SC and HC of temporal rules
Output: $\mathcal{R}_t$: the set of temporal rules satisfying T_{sc} and T_{hc}

Mine static rules by a rule mining tool such as AMIE+ from $\mathcal{G}_s$;
Convert each static rule into candidate temporal rules via temporal rule patterns;
while *select a candidate temporal rule* Ru **do**
  for *t=1,2,…, T* **do**
    Generate two sub-graphs $\mathcal{G}_{s1}$ and $\mathcal{G}_{s2}$ via fusing the events in the time intervals [t+1, t+W_T] and [t+W_T+1, t+2W_T] from $\mathcal{G}_t$;
    if *the length of* Ru *is 1* **then**
      Calculate the SC and HC of Ru by Eqs. 2-4;
    else
      Calculate the SC and HC of Ru by Eqs. 18-20;
  if *the mean value of SC and HC at all the timestamps are both higher than* T_{sc} *and* T_{hc} **then**
    Add Ru into $\mathcal{R}_t$;

Lemma 2. Based on the temporal rule pattern namely (3) $p_3(x, y, t{+}t_1{+}t_2) \Leftarrow p_1(x, z, t) \wedge p_2(z, y, t{+}t_1)$, the temporal rule-guided predicate embedding regularization specific to Eq. 7 could represent the causality among three events at different timestamps.

Proof. If there are two events (s, p_1, e, t) and $(e, p_2, o, t + t_1)$ holds, the causality term of the time-sensitive scores specific to these two events transformed to the timestamp $t + t_1 + t_2$ are achieved according to Eq. 10 as follows:

$$\mathrm{Re}\left(\mathbf{s}^{\top}\, diag(\mathbf{T} \circ \mathbf{T} \circ \mathbf{p}_{r1})\, \bar{\mathbf{e}}\right) = 1 \tag{24}$$

$$\mathrm{Re}\left(\mathbf{e}^{\top}\, diag(\mathbf{T} \circ \mathbf{p}_{r2})\bar{\mathbf{o}}\right) = 1 \tag{25}$$

In terms of the following equation

$$\begin{aligned} &\mathbf{s}^{\top}\, diag(\mathbf{T} \circ \mathbf{T} \circ \mathbf{p}_{r1})\, \bar{\mathbf{e}}\mathbf{e}^{\top}\, diag(\mathbf{T} \circ \mathbf{p}_{r2})\bar{\mathbf{o}} \\ &= \mathbf{s}^{\top}\, diag(\mathbf{T} \circ \mathbf{T} \circ \mathbf{p}_{r1} \circ \mathbf{T} \circ \mathbf{p}_{r2})\, \bar{\mathbf{o}} \end{aligned} \tag{26}$$

and Eqs. 24 and 25, we could derive that

$$\begin{aligned} &\mathrm{Re}\left(\mathbf{s}^{\top}\, diag(\mathbf{T} \circ \mathbf{T} \circ \mathbf{p}_{r1} \circ \mathbf{T} \circ \mathbf{p}_{r2})\, \bar{\mathbf{o}}\right) = 1 - \\ &\mathrm{Im}\left(\mathbf{s}^{\top}\, diag(\mathbf{T} \circ \mathbf{T} \circ \mathbf{p}_{r1})\bar{\mathbf{e}}\right) \cdot \mathrm{Im}(\mathbf{e}^{\top}\, diag(\mathbf{T} \circ \mathbf{p}_{r2})\bar{\mathbf{o}}) \end{aligned} \tag{27}$$

Furthermore, if the following equation

$$\mathrm{Im}\left(\mathbf{s}^{\top}\, diag(\mathbf{T} \circ \mathbf{T} \circ \mathbf{p}_{r1})\bar{\mathbf{e}}\right) \cdot \mathrm{Im}(\mathbf{e}^{\top}\, diag(\mathbf{T} \circ \mathbf{p}_{r2})\bar{\mathbf{o}}) = 0 \tag{28}$$

is satisfied, Eq. 27 can be rewritten as:

$$\mathrm{Re}\left(\mathbf{s}^{\top}\, diag(\mathbf{T} \circ \mathbf{T} \circ \mathbf{p}_{r1} \circ \mathbf{T} \circ \mathbf{p}_{r2})\, \bar{\mathbf{o}}\right) = 1 \tag{29}$$

Substituting $\mathbf{G}$ defined in Eq. 7 into Eq. 29, we can achieve that

$$\mathrm{Re}\left(\mathbf{s}^{\top}\, diag(\mathbf{p}_{r3})\, \bar{\mathbf{o}}\right) = 1 \tag{30}$$

Therefore, we prove that the event $(s, p_3, o, t + t_1 + t_2)$ is valid if (s, p_1, e, t) and $(e, p_2, o, t + t_1)$ successively occur via the temporal rule-guided predicate embedding regularization $\mathbf{G}$ of the temporal rule pattern (3), which indicates the causality among three events occurring at different timestamps. □

The proofs of the causality among events corresponding to the other three temporal rule patterns modeled by our designed temporal rule-guided predicate embedding regularization $\mathbf{G}$ can be obtained similarly to the above proofs.

From the proofs corresponding to Lemma 1 and Lemma 2, the causality term in the time-sensitive score of a candidate event causally related to the other already occurred events would be a higher value. More interestingly, it can explain the effectiveness of representing the causality of events via our temporal rule-guided predicate embedding regularization mechanism and facilitate higher accuracy of TKGC.