
DAG-GNN: DAG Structure Learning with Graph Neural Networks (Supplementary Material)

Yue Yu^{*1} **Jie Chen**^{*23} **Tian Gao**³ **Mo Yu**³

1. Proofs

Proof of Theorem 1. Let $B = A \circ A$. Clearly, B is nonnegative. The binomial expansion reads

$$(I + \alpha B)^m = I + \sum_{k=1}^m \binom{m}{k} \alpha^k B^k.$$

It is known that there is a cycle of length k if and only if $\text{tr}(B^k) > 0$ when $B \geq 0$. Because if there is a cycle then there is a cycle of length at most m , we conclude that there is no cycle if and only if $\text{tr}[(I + \alpha B)^m] = \text{tr}(I) = m$. \square

Proof of Theorem 2. Write

$$(1 + \alpha|\lambda|)^m = \left(1 + \frac{c|\lambda|}{m}\right)^m.$$

For given c and $|\lambda|$, the right-hand side of the equality is a function of m . This function monotonically increases for positive m and has a limit $e^{c|\lambda|}$. Hence, for any finite $m > 0$, $(1 + \alpha|\lambda|)^m \leq e^{c|\lambda|}$. \square

2. Structure Learning over KB Relations

We construct the data set from triples in FB15K-237 (Toutanova et al., 2015), which is a subset of FreeBase with approximately 15k entities and 237 relations. Each sample corresponds to an entity and each variable corresponds to a relation in this knowledge base. Each sample has on average 7.36 relations (i.e. 7.36 non-zero entries in each row).

Table 1 gives additional examples learned by our model with highest confidence scores. For each target relation on the right-hand side, we show the highest ranked relations within the same domain (i.e. the contents in the field before “/” such as “film” and “tvProgram”). On the left-hand side, we omit the relations that are common to the associated entity

^{*}Equal contribution ¹Lehigh University ²MIT-IBM Watson AI Lab ³IBM Research. Correspondence to: Yue Yu <yuy214@lehigh.edu>, Jie Chen <chenjie@us.ibm.com>.

types, e.g. “profession” and “gender” to persons and “genre” to films, because almost all entities with these types will contain such a relation.

References

- Toutanova, K., Chen, D., Pantel, P., Poon, H., Choudhury, P., and Gamon, M. Representing text for joint embedding of text and knowledge bases. In *EMNLP*, 2015.

Table 1. (Continued from Table 3) Examples of extracted edges with high confidence. The dot · appearing in $R_1 \cdot R_2$ means that the sample entity is connected to a virtual node (i.e. compound value types in FreeBase) via relation R_1 , followed by a relation R_2 to a real entity.

film/ProducedBy	\Rightarrow	film/Country
film/ProductionCompanies	\Rightarrow	film/Country
tvProgram/CountryOfOriginal	\Rightarrow	tvProgram/Language
tvProgram/RegularCast.regularTv/AppearanceActor	\Rightarrow	tvProgram/Language
person/Nationality	\Rightarrow	person/Languages
person/PlaceOfBirth	\Rightarrow	person/Languages
person/PlaceOfBirth	\Rightarrow	person/Nationality
person/PlaceLivedLocation	\Rightarrow	person/Nationality
organization/Headquarters.mailingAddress/Citytown	\Rightarrow	organization/PlaceFounded
organization/Headquarters.mailingAddress/StateProvince	\Rightarrow	organization/PlaceFounded