A Appendices

To better demonstrate the effectiveness of the graph attention mechanism (GAT), and separate its function from LSTM, we performed ablation studies. We build a model called LSTMPath, which removes GAT from AttnPath, so the definition of state vector $\mathbf{s}_{i,t}$ is shortened to $[\mathbf{m}_{i,t}; \mathbf{h}_t]$.

Table 6 shows the success rate (SR) of finding paths of several path-based methods. DeepPath NoPre means directly training DeepPath, without pretraining, i.e, showing several right paths to pretrain the model. This proves that the performance of DeepPath heavily relies on pretraining, while AttnPath model is capable of performing better without pretraining.

Table 7 shows fact prediction MAP of several compared methods. Results of TransE / H / R / D and DeepPath are cited from (Xiong et al., 2017). It proves that introducing GAT into the model is effective.

Table 8 shows link prediction MAP of several compared methods. Results of TransE / R, PRA, DeepPath are cited from (Xiong et al., 2017), while results of MINERVA are cited from (Das et al., 2018).

Table 9 shows link prediction hits@1 of several compared methods. Hits@1 means the proportion of (h, r) pair where the algorithm ranks the ground truth tail entity to the first place. Results of MINERVA are cited from (Das et al., 2018).

Table 6: Ablation study: SR of finding paths of several path-based methods.

Method		FB15K-237		NELL-995	
DeepPath		17.7		31.8	
DeepPath TransD		18.2		32.7	
DeepPath NoPre		7.3		27.8	
LSTMPath	AttnPath	14.7	15.7	33.9	37.3
LSTMPath MS / RR	AttnPath MS / RR	17.4	18.1	35.9	40.4
LSTMPath Force	AttnPath Force	29.0	**30.4**	46.7	**48.1**

Table 7: Ablation study: fact prediction MAP of several compared methods.

Methods		FB15K-237		NELL-995	
TransE		27.7		38.3	
TransH		30.9		38.9	
TransR		30.2		40.6	
TransD		30.3		41.3	
PRA		30.8		57.5	
DeepPath		31.1		49.3	
DeepPath TransD		31.3		53.5	
DeepPath NoPre		23.0		52.1	
LSTMPath	AttnPath	31.0	31.5	57.5	59.8
LSTMPath MS / RR	AttnPath MS / RR	34.4	34.6	63.4	65.4
LSTMPath Force	AttnPath Force	36.0	**37.9**	68.6	**69.3**

Table 8: Ablation study: link prediction MAP of several compared methods.

Method		FB15K-237		NELL-995	
TransE		53.2		73.7	
TransH		36.0		79.7	
TransR		54.0		78.9	
TransD		63.7		81.5	
PRA		54.1		67.5	
MINERVA		29.3		72.5	
DeepPath		57.2		79.6	
DeepPath TransD		63.5		82.6	
DeepPath NoPre		53.9		79.6	
LSTMPath	AttnPath	58.1	58.5	81.2	81.4
LSTMPath MS / RR	AttnPath MS / RR	61.9	62.3	82.8	83.0
LSTMPath Force	AttnPath Force	65.4	**66.1**	85.5	**85.8**

Table 9: Ablation study: link prediction hits@1 of several compared methods.

Method		FB15K-237		NELL-995	
TransE		49.2		66.0	
TransH		21.3		70.0	
TransR		40.3		68.1	
TransD		46.2		71.9	
PRA		46.1		74.2	
MINERVA		21.7		66.3	
DeepPath		46.0		71.0	
DeepPath TransD		47.9		72.4	
DeepPath NoPre		37.5		71.1	
LSTMPath	AttnPath	46.9	47.3	72.9	73.3
LSTMPath MS / RR	AttnPath MS / RR	48.1	48.2	75.9	76.2
LSTMPath Force	AttnPath Force	51.4	**52.1**	79.7	**80.3**