Table 1. Model size and training time. The runtime is in hours (h).

	Carcinogenesis		Mutagenesis		Semantic Bible		Vicodi	
	|Params.|	Time	|Params.|	Time	|Params.|	Time	|Params.|	Time
NCES2$_{m=32}$	2,747,376	1.61h	1,974,682	0.84h	1,233,552	0.39h	3,610,516	3.79h
NCES2$_{m=64}$	2,755,568	1.86h	1,982,874	0.94h	1,241,744	0.45h	3,618,708	4.38h
NCES2$_{m=128}$	2,771,952	2.40h	1,999,258	1.20h	1,258,128	0.55h	3,635,092	5.61h

We report additional results that could not fit into the paper. Figure 1 shows the loss and *Soft Accuracy* curves of NCES2 when the embedding model ConEx is used. We can observe that the loss rapidly decreases in early epochs and starts to stagnate after approximately 100 epochs. As in the case of hard accuracy curves (see the main paper), the *Soft Accuracy* curves show a rapid increase on especially the largest datasets Carcinogenesis and Vicodi, and reach an accuracy of 95% on all datasets. This suggests that NCES2 successfully learns the mapping between instances data and class expressions in $\mathcal{ALCHIQ}(\mathcal{D})$.

Table 1 gives the number of trainable parameters and the training time for each NCES2 instance. As expected, we can observe that training times are higher on large datasets (up to 5 hours on Vicodi for NCES2$_{m=128}$) than on small datasets. Note that training is only required once per dataset, and many learning problems can be solved afterwards. This is not the case with search-based approaches as they always restart from the beginning for each new learning problem [1]. In this regard, our approach is more suitable for knowledge bases where many learning problems are to be solved.

We conducted further experiments where we used a different embedding model to evaluate the effectiveness of NCES2. We chose DistMult [2] for this purpose. The training curves of NCES2 when using DistMult are shown in Figure 2, and experimental results on class expression learning on the test sets are reported in Table 2. From Figure 2, we can observe as in the case of ConEx that the loss decreases and stagnates during training. We also observe an excellent performance regarding the accuracy curves as we reach 95% on all datasets. Moreover, the behaviour of the curves are similar to those with the ConEx embedding model: accuracy curves increase faster on the largest datasets Carcinogenesis and Vicodi. This is in line with our early hypothesis that NCES2 learns faster on large datasets (see the main paper). On the other side, Table 2 shows that NCES2 still outperforms search-based approaches in runtime on all datasets, and in F-measure on the largest datasets Carcinogenesis and Vicodi. This observation is the same as in the case of the ConEx model. We therefore expect NCES2 instances to work well with other state-of-the-art embedding models.

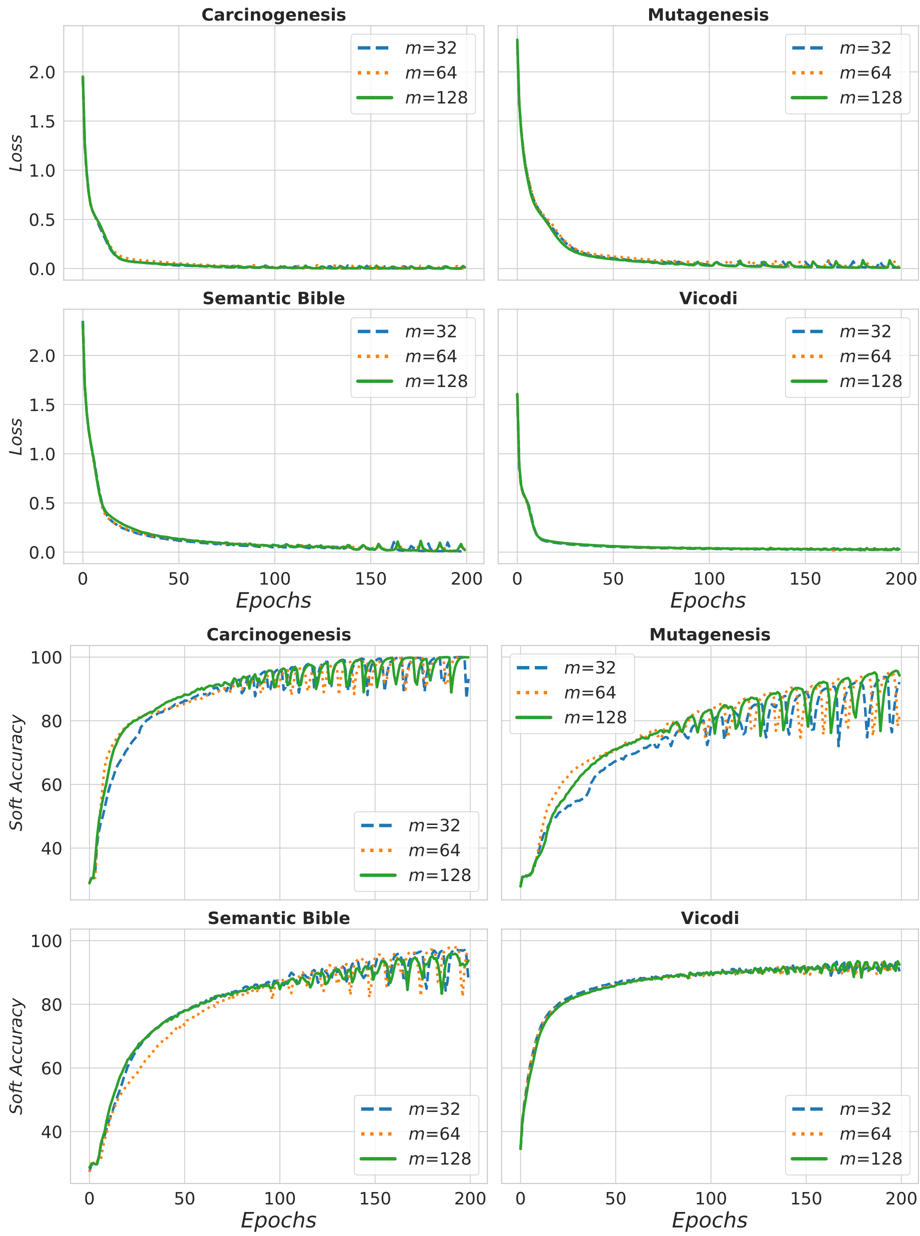

Fig. 1. Loss and *Soft Accuracy* curves using the ConEx embedding model.

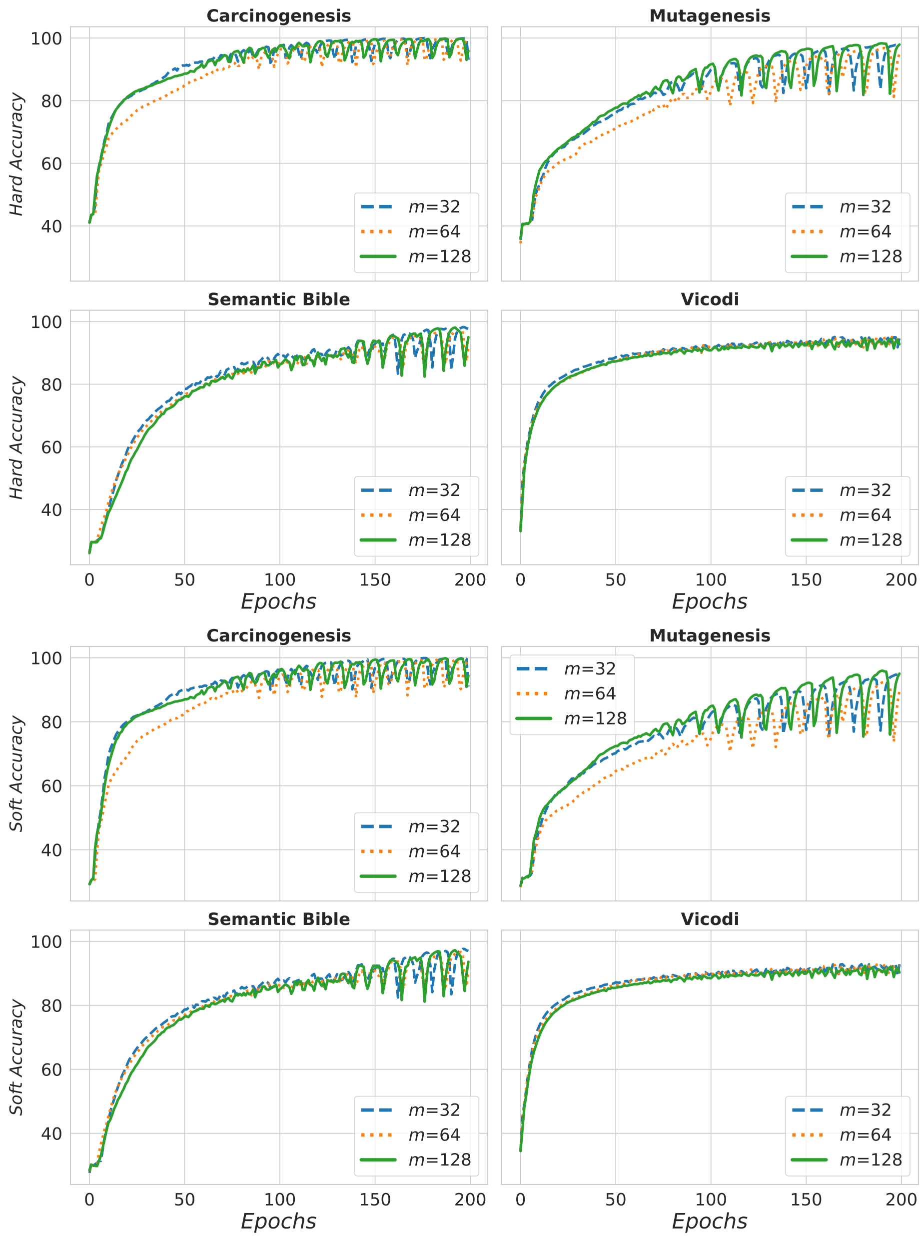

Fig. 2. Training accuracy curves using the DistMult embedding model.

Table 2. Evaluation of NCES2 instances per dataset using the Distmult embedding model.

	F$_1$ (%)			
	Carcinogenesis	**Mutagenesis**	**Semantic Bible**	**Vicodi**
CELOE	29.24±39.22	74.46±37.59	**88.60**±19.50	22.63±35.21
EvoLearner	89.34±15.80	**95.37**± 8.02	88.38±12.50	76.99±26.37
ELTL	14.46±28.48	36.33±34.98	35.21±31.74	8.58±22.94
ECII	18.91±31.46	34.33±31.53	32.79±32.18	29.20±30.81
NCES$_{m=32}$	87.18±30.21	77.11±35.53	77.73±30.93	89.93±23.25
NCES$_{m=64}$	86.02±30.57	73.22±38.28	74.05±33.13	81.62±32.24
NCES$_{m=128}$	86.23±31.20	73.66±38.53	66.90±38.23	80.87±31.91
NCES$_{m=\{32,64\}}$	91.17±26.63	85.14±28.80	77.96±31.50	89.23±25.64
NCES$_{m=\{32,128\}}$	**92.22**±25.41	84.68±31.94	78.48±33.35	89.80±25.03
NCES$_{m=\{64,128\}}$	90.73±26.71	86.29±30.65	76.64±33.42	**91.50**±22.68
NCES$_{m=\{32,64,128\}}$	91.75±24.80	86.98±29.54	79.07±31.71	90.03±26.33

	Accuracy (%)			
	Carcinogenesis	**Mutagenesis**	**Semantic Bible**	**Vicodi**
CELOE	62.96±22.56	87.33±17.80	95.85± 8.71	78.59±15.83
EvoLearner	**99.68**± 0.81	**99.52**± 2.17	**97.43**± 4.74	**97.79**± 6.74
ELTL	19.37±32.31	40.58±35.33	39.40±29.92	41.67±44.29
ECII	27.17±38.40	32.52±33.31	29.06±33.37	71.05±39.43
NCES$_{m=32}$	94.39±20.58	89.34±23.69	89.04±21.08	95.08±19.97
NCES$_{m=64}$	93.56±21.85	90.55±22.83	85.97±26.08	94.13±21.38
NCES$_{m=128}$	94.01±19.44	87.52±27.19	85.88±25.30	95.61±17.12
NCES$_{m=\{32,64\}}$	93.71±21.77	92.31±20.61	87.91±24.01	95.50±18.95
NCES$_{m=\{32,128\}}$	94.82±20.61	91.90±24.15	88.73±24.58	95.91±17.37
NCES$_{m=\{64,128\}}$	94.45±21.09	91.95±23.97	86.51±26.65	96.77±14.52
NCES$_{m=\{32,64,128\}}$	93.88±21.39	92.30±23.74	89.01±23.59	95.13±19.26

	Runtime (%)			
	Carcinogenesis	**Mutagenesis**	**Semantic Bible**	**Vicodi**
CELOE	268.90±116.04	165.27±145.11	172.04±140.27	334.99±43.87
EvoLearner	62.21±26.11	70.77±47.53	18.44±5.53	236.92±80.90
ELTL	26.15±2.11	15.83±16.56	4.73±0.98	335.90±205.39
ECII	25.62±6.11	20.40±4.00	6.73±1.67	37.12±25.12
NCES$_{m=32}$	**0.02**±0.00	**0.03**±0.00	**0.01**±0.00	**0.04**±0.00
NCES$_{m=64}$	0.02±0.00	0.03±0.00	0.02±0.00	0.04±0.00
NCES$_{m=128}$	0.03±0.00	0.03±0.00	0.02±0.00	0.04±0.00
NCES$_{m=\{32,64\}}$	0.05±0.00	0.05±0.00	0.04±0.00	0.06±0.00
NCES$_{m=\{32,128\}}$	0.06±0.00	0.06±0.00	0.04±0.00	0.07±0.00
NCES$_{m=\{64,128\}}$	0.06±0.00	0.06±0.00	0.04±0.00	0.07±0.00
NCES$_{m=\{32,64,128\}}$	0.09±0.00	0.10±0.00	0.06±0.00	0.11±0.00

Bibliography

[1] Kouagou, N.J., Heindorf, S., Demir, C., Ngomo, A.N.: Neural class expression synthesis. CoRR **abs/2111.08486** (2021)
[2] Yang, B., Yih, S.W.t., He, X., Gao, J., Deng, L.: Embedding entities and relations for learning and inference in knowledge bases. In: Proceedings of the International Conference on Learning Representations (ICLR) 2015 (May 2015)