TOWARDS LOGICALLY CONSISTENT LANGUAGE MODELS VIA PROBABILISTIC REASONING

Diego Calanzone[*,+] **Antonio Vergari**[×,🎲] **Stefano Teso**[+,🎲]

ABSTRACT

Large language models (LLMs) are a promising venue for natural language understanding and generation tasks. However, current LLMs are far from reliable: they are prone to generate non-factual information and, more crucially, to contradict themselves when prompted to reason about beliefs of the world. These problems are currently addressed with large scale fine-tuning or by delegating consistent reasoning to external tools. In this work, we strive for a middle ground and introduce a training objective based on principled probabilistic reasoning that teaches a LLM to be consistent with external knowledge in the form of a set of facts and rules. Fine-tuning with our loss on a limited set of facts enables our LLMs to be more logically consistent than previous baselines and allows them to extrapolate to unseen but semantically similar factual knowledge more systematically.

1 INTRODUCTION

Developing reliable large language models (LLMs) and safely deploying them is more and more crucial, particularly when they are used as external sources of knowledge (Petroni et al., 2019). To do so, one would need LLMs to be *factual* (Wadden et al., 2020), i.e., generating content that satisfies some knowledge base (KB), and *logically self-consistent* (Li et al., 2019), i.e., being able not to contradict themselves when prompted to perform complex reasoning. Clearly, training on large datasets for question answering (QA) (Tafjord & Clark, 2021) alone cannot meet these desiderata (Evans et al., 2021; Lin et al., 2022; Liu et al., 2023).

Factuality and consistency are intimately related. Enforcing factuality alone generally boils down to fine-tuning an LLM on a large KB of atomic facts (Kassner et al., 2021). When predicting the truth values of these facts, a number of works try to enforce the simplest form of consistency: that the probability of a true fact shall be one minus the probability of its negation (Burns et al., 2022). Liu et al. (2023) use more sophisticated heuristics, fine-tuning on a large QA dataset with a combination of three objectives: binary cross entropy to classify true facts; multiclass cross entropy to choose the true statement among other options; a supervised contrastive loss to pull apart true and false facts. All these approaches require large KBs and extrapolating to unseen facts remains an open challenge.

When it comes to self-consistency w.r.t. more complex reasoning scenarios, e.g., ensuring that LLMs can perform modus ponens without contradicting themselves (Tafjord et al., 2022), one line of research focuses on employing external reasoning tools such as MAX-SAT solvers (Battiti, 2009) at inference time (Mitchell et al., 2022; Jung et al., 2022; Kassner et al., 2023). However, these approaches depend on the constant availability of a reasoner (and sometimes also of a natural language inference model (Mitchell et al., 2022)) which can increase the cost of inference for every reasoning step. At the same time, training the LLM to reason is not possible or hindered by the hardness of backpropagating through the solvers (Pogancic et al., 2020).

In this work, we show how to improve factuality and self-consistency of LLMs without external components by leveraging recent advancements in neuro-symbolic learning (De Raedt et al., 2021). This is done by turning complex reasoning tasks into logical constraints that can be incorporated in a semantic loss (Xu et al., 2018). This in turn enforces the LLM to perform principled probabilistic reasoning at training time over the possible truth assignments. Among previous works of research, Zhang et al. (2023) similarly applied the semantic loss to instill integrity constraints in the embedding space of entities in encoder-only models. It is worth noting that representations of beliefs can

[*,+] University of Trento, [×] University of Edinburgh, [🎲] joint supervision.

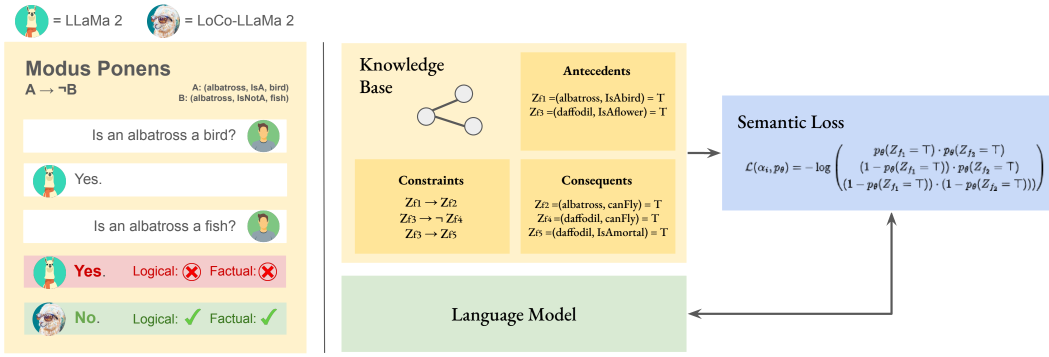

Figure 1: Our LoCo-LMs are trained by compiling implication constraints in a semantic loss and applying them to some atomic facts as antecedents. At test time, they can predict the missing consequents more consistently and factually than other baselines (Table 1).

be highly contextual on the context, as shown by Arakelyan et al. (2024) with NLI models, suggesting additional forms of consistency to be included. We empirically show how given incomplete factual knowledge, the LLM can learn truth beliefs for new facts while keeping logical consistency with prior knowledge. In our experiments, with a single offline training session, LLMs trained with our objective outperform models relying on external solvers, and are more factual and logically consistent in low-data regimes when compared to standard supervised fine-tuning.

2 Tackling Factual & Consistent Inference with LoCo-LMs

Setting. We view a pre-trained LLM as a collection of truth beliefs (facts) over which it can *factually reason*: when prompted with question like "Is a bison a mammal?", it can supply a binary prediction of the form "Yes"/"No" or "True"/"False". When given a limited set of textual statements $\mathcal{D}_f = \{f_1, \ldots, f_n\}$, such as "an albatross is a bird", and general logical implications (from here on, constraints) $\mathcal{D}_c = \{\alpha_1, \ldots, \alpha_m\}$, such as "IsAbird $\to$ canFly", we want an LLM queried about new beliefs $\{f_{n+1}, \ldots, f_k\}$ to output truth values that are *consistent with the known facts* $\mathcal{D}_f$ and the prior KB $\mathcal{D}_c$. We assume the facts are provided as *(subject, property)* pairs, and that the logical implications specify how new properties (consequents) follow from the initial ones (antecedents).

Logical constraints satisfaction. LLMs fit only on facts – via, e.g., a cross-entropy loss – struggle to infer consistent truth values for unseen consequents that satisfy the KB $\mathcal{D}_c$. To design logically-consistent LLMs (LoCo-LMs), we instead minimize a semantic loss (SL) (Xu et al., 2018) that penalizes a model proportionally to the probability it allocates to truth values inconsistent with one or more such logical constraints. To build intuition, consider two facts $f_1 =$ (daffodil, IsAflower), $f_2 =$ (daffodil, IsMortal) and a logical implication α_i stating that IsAflower entails IsMortal, or

$$\alpha_i := (z_{f_1} \to z_{f_2}) \iff (\neg z_{f_1} \vee z_{f_2}) \tag{1}$$

where z_f is the truth value assigned to a fact f. The SL in LoCo-LMs will penalize the LLM distribution p_θ if it allocates truth values inconsistent with α_i, specifically $(z_{f_1} = \top, z_{f_2} = \bot)$ which encodes the false belief that daffodils are flowers ($\top =$ true) but not mortal ($\bot =$ false). To this end, it relies on probabilities for these truth values extracted from the model. We obtain these probabilities directly by reading off the likelihood of utterances produced by the LLM, that is:

$$p_\theta(z_{f_1} = \top) = p_\theta(x_t = l \mid x_1, \ldots, x_{t-1} = \text{"Is a daffodil a flower?"}) \tag{2}$$

$$p_\theta(z_{f_2} = \top) = p_\theta(x_t = l \mid x_1, \ldots, x_{t-1} = \text{"Is a daffodil a mortal?"}) \tag{3}$$

Here, l is a textual truth value, e.g., "Yes" for $\top$ or "No" for $\bot$. More generally, the SL encourages consistency by maximizing the likelihood of assignments satisfying $\alpha_i \in \mathcal{D}_c$, and can be applied to more complex reasoning scenarios that involve not only implications but any logical constraint. From a constraint $\alpha_i \in \mathcal{D}_c$, a LoCo-LM is fine-tuned by minimizing a loss that is defined as:

$$\mathcal{L}(\alpha_i, p_\theta) = -\log \sum\nolimits_{\mathbf{z} \models \alpha_i} \prod\nolimits_{j:\mathbf{z} \models z_{f_j}} p_\theta(z_{f_j}) \prod\nolimits_{j:\mathbf{z} \models \neg z_{f_j}} 1 - p_\theta(z_{f_j}) \tag{4}$$

where $z \models \alpha_i$ means that a set of truth assignment to facts satisfies α_i and the argument of the logarithm computes the probability of α_i being true. Taking for instance the constraint in Equation (1), the assignments from the truth table satisfying the formula are $\{(\top,\top),(\bot,\top),(\bot,\bot)\}$, the semantic loss in a LoCo-LM would thus minimize the negative log likelihood of these assignments:

$$\mathcal{L}(\alpha_i, p_\theta) = -\log \begin{pmatrix} p_\theta(z_{f_1} = \top) \cdot p_\theta(z_{f_2} = \top) \ + \\ (1 - p_\theta(z_{f_1} = \top)) \cdot p_\theta(z_{f_2} = \top) \ + \\ (1 - p_\theta(z_{f_1} = \top)) \cdot (1 - p_\theta(z_{f_2} = \top)) \end{pmatrix} \tag{5}$$

Figure 1 illustrates the pipeline of LoCo-LMs.

Factuality. As described so far, the SL in LoCo-LMs fosters self-consistency, but can hinder factuality as it considers as equally valid assignments both the cases where antecendents can be true or false when consequents are true (i.e., $\{(\top,\top),(\bot,\top)\}$). Therefore, given a training set of ground facts $\mathcal{D}_f = \{f_1, \ldots, f_n\}$ assumed to be true, we embed this factual information in LoCo-LMs by adding it to the logical constraints via a logical conjunction. E.g., from the example constraint in formula 1, if we assume the fact $f_1 =$ (daffodil, IsAflower) is assumed to be true, we get:

$$z_{f_1} = \top, \quad \alpha'_i = (z_{f_1} \to z_{f_2}) \wedge z_{f_1} \tag{6}$$

The assignments satisfying this formula change to $\mu' = \{(\top,\top)\}$, $\mu' \models \alpha'_i$, and so does the optimization objective as in equation 5. A full derivation can be found in Appendix (C.1).

3 Experiments

We aim to answer these research questions: **RQ1)** can LoCo-LMs compete with approaches using external reasoners without doing so? **RQ2)** can they require less data than standard fine-tuning?

Data. We evaluate LoCo-LMs on the BeliefBank data set (Kassner et al., 2021). It consists of three pieces: a "calibration" set of $1,072$ annotated facts about 7 entities of the form *(subject, property, true/false)* used for training, a "silver" set of $12,636$ facts about 85 entities used for evaluation, and a set of 2224 valid logical implications $\mathcal{D}_c$. The SL requires defining a set of ground constraints. We derive these as follows. For each general constraint in $\mathcal{D}_c$, we lookup the subjects of all facts in the training set: if the antecedent or the consequent fact of the general constraint is known for that subject, we add the subject ground constraint to the dataset.

For RQ1, we generate two splits: *T1 facts*, appearing either as antecedents or consequents in the constraints; *T2 facts*, appearing exclusively as constraint consequents. The goal is to correctly guess the consequents by seeing only the antecedents and the constraints. In the calibration set, we count 796 antecedents and 276 consequents, spawning 14,005 grounded constraints. In the silver set, we count 9,504 antecedents and 3,132 consequents, spawning 169,913 grounded constraints. For RQ2, we split the dataset to test the effects of pure supervised fine-tuning: a portion of random facts from the calibration set is taken with the goal to predict the excluded antecedent or consequent facts.

Models. Following Mitchell et al. (2022), we work with a pre-trained Macaw-Large model (Tafjord & Clark, 2021) (770M parameters) capable of multi-angle question answering with fixed prompt templates. We adopt the same prompts as in the original paper: to query for binary beliefs, the model is queried: "`$answer$ ; $mcoptions$ = (A) Yes.  (B) No.  ; $question$ = Is [X] a [Y]?`", and the expected continuation would be: "`$answer$ = Yes./No.`". Since the training facts are presented in the form *(subject, property)*, we employ the functions used by Mitchell et al. (2022) to fill in natural language templates such as "`Is it true that [X] is [Y]?`" or "`Is it [X] a [Y]?`". Further training details are reported in Appendix A.

Competitors and Metrics. We consider two baselines: the ConCoRD logical layer (Mitchell et al., 2022) applied to Macaw-Large, using RoBERTa-ANLI (Liu et al., 2019) for relationship inference, and a pre-trained Macaw-Large model from Tafjord & Clark (2021) as zero-shot baseline. In Appendix B.2, we compare our models with the baseline in terms of runtime requirements. We evaluate our models for *factuality* and *logical self-consistency*. We measure the former with the F_1 score to account for the unbalance between false and true facts (Kassner et al., 2021). Factuality is measured on the two splits (antecedents and consequents) and the complete facts set for both the distributions (calibration, silver). As in (Li et al., 2019), we measure *logical self-consistency* as the fraction of non violated constraints $\alpha_i \in \mathcal{D}_c^{\text{test}}$:

$$1 - |\{\alpha_i \mid \neg(z_j \to z_k)\}| \;/\; |\{\alpha_i \mid z_j\}| \tag{7}$$

Table 1: **LoCo-LMs achieve better logical self-consistency and factuality** as measured via Equation (7) and F_1 scores when compared to classical supervised finetuning (SFT) and baselines using external reasoners such as ConCoRD (Mitchell et al., 2022) measured on test (silver set) facts. For RQ1 (Section 3), LoCo-LMs fine-tuned on T1 facts only outperform training-free baseline for all metrics. For RQ2, they boost performance in the presence of a small fraction of T1+T2 facts (5-10%). For larger dataset sizes, LoCo-LMs are competitive for consistency and factuality on consequents. A similar trend is visible on training data (Table 2).

	Method	Train Subset	Antecedents F_1	Consequents F_1	Total F_1	Logical consistency
RQ1	ConCoRD				0.91	0.91
	Macaw-Large		0.52	0.90	0.81	0.83
	SFT	T1	0.13	0.01	0.03	0.72
	LoCo-LM	T1	**0.79**	**0.98**	**0.96**	**0.99**
RQ2	SFT	T1+T2 (5%)	0.23	0.78	0.72	0.82
	LoCo-LM	T1+T2 (5%)	**0.67**	**0.83**	**0.81**	**0.92**
	SFT	T1+T2 (10%)	**0.55**	**0.97**	**0.91**	0.90
	LoCo-LM	T1+T2 (10%)	0.45	**0.97**	0.89	**0.93**
	SFT	T1+T2 (75%)	**0.85**	**0.99**	**0.97**	**0.98**
	LoCo-LM	T1+T2 (75%)	0.79	**0.99**	0.95	**0.98**

i.e., the proportion of unsatisfied formulas where the antecedent is believed to be true. According to literature, we implicitly refer to “self-consistency”, that is logical consistency based on the model’s own beliefs. Logical consistency scores are computed by iterating through the subjects, the general constraints and by querying each belief appearing in the formulas.

Results. We report factuality and logical consistency scores for the test (Table 1) and training distribution (Table 2). The test distribution involves entities unseen at training time, thus we assume the language model implicitly transfers some semantic knowledge based on similarities in the conceptual space; we further investigate this by calculating pairwise similarities between entity embeddings in Appendix B.3.

We firstly observe a net improvement in both factuality and logical consistency with semantic loss fine-tuning, compared to pre-trained Macaw-Large and the ConCoRD variant (Table 1, RQ1). Standard supervised fine-tuning on antecedent facts is insufficient: due to a class imbalance between true facts ($\sim 10\%$) and false facts ($\sim 90\%$), the model tends to label any statement as “false”; moreover, no knowledge about consequent facts is introduced. This is accentuated in the training distribution (Table 2): with semantic loss, the model attains higher factuality and logical consistency in both the standard setup and with a small amount of cheating data. Finally, assuming sufficient domain knowledge overlap (Appendix B.3), high logical consistency scores on the test distribution suggest the imposed logical structures in the conceptual space can be generalized.

Assuming the language model can access to a portion of consequent facts (Table 1, RQ2), semantic loss fine-tuned LoCo-LMs still yields better logical consistency and factuality for unseen consequents in low-data regimes (e.g. 5-10% of the T1+T2 dataset) compared to canonical supervised fine-tuning. When they are allowed to see more data (e.g. 75% of the T1+T2 dataset), traditionally finetuned models can “cheat” and directly learn about the consequents (somehow equivalent to memorizing a single row of the truth table). In this scenario, LoCo-LMs achieve compareble logical self-consistency and factuality over consequents, but not on the antecedents.

4 Conclusion and Further Work

Our results show that LoCo-LMs improve upon ConCoRD in terms of factuality and self-consistency in complex reasoning tasks, especially when queried on unseen facts. This suggests that probabilistic reasoning objectives can impose structure in a language model’s conceptual space. In future work, we aim to further investigate the transfer of logical structures across entities with semantic relations, such as e.g. abstraction or homology. We also consider the implications of scaling language models, in terms of logical consistency and training efficiency. Finally, we plan to extend our analysis to more complex logical operators (Vergari et al., 2021) and to consider more advanced probabilistic reasoning techniques that sport improved consistency guarantees (Ahmed et al., 2022).

REFERENCES

Kareem Ahmed, Stefano Teso, Kai-Wei Chang, Guy Van den Broeck, and Antonio Vergari. Semantic Probabilistic Layers for Neuro-Symbolic Learning. In *NeurIPS*, 2022.

Erik Arakelyan, Zhaoqi Liu, and Isabelle Augenstein. Semantic sensitivities and inconsistent predictions: Measuring the fragility of nli models, 2024.

Roberto Battiti. *Maximum satisfiability problemMaximum Satisfiability Problem*, pp. 2035–2041. Springer US, Boston, MA, 2009. ISBN 978-0-387-74759-0. doi: 10.1007/978-0-387-74759-0_364.

Collin Burns, Haotian Ye, Dan Klein, and Jacob Steinhardt. Discovering latent knowledge in language models without supervision, 2022.

Luc De Raedt, Sebastijan Dumančić, Robin Manhaeve, and Giuseppe Marra. From statistical relational to neural-symbolic artificial intelligence. In *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, pp. 4943–4950, 2021.

Owain Evans, Owen Cotton-Barratt, Lukas Finnveden, Adam Bales, Avital Balwit, Peter Wills, Luca Righetti, and William Saunders. Truthful ai: Developing and governing ai that does not lie, 2021.

Jaehun Jung, Lianhui Qin, Sean Welleck, Faeze Brahman, Chandra Bhagavatula, Ronan Le Bras, and Yejin Choi. Maieutic prompting: Logically consistent reasoning with recursive explanations, 2022.

Nora Kassner, Oyvind Tafjord, Hinrich Schütze, and Peter Clark. Beliefbank: Adding memory to a pre-trained language model for a systematic notion of belief, 2021.

Nora Kassner, Oyvind Tafjord, Ashish Sabharwal, Kyle Richardson, Hinrich Schuetze, and Peter Clark. Language models with rationality, 2023.

Tao Li, Vivek Gupta, Maitrey Mehta, and Vivek Srikumar. A logic-driven framework for consistency of neural models, 2019.

Stephanie Lin, Jacob Hilton, and Owain Evans. Truthfulqa: Measuring how models mimic human falsehoods, 2022.

Jiacheng Liu, Wenya Wang, Dianzhuo Wang, Noah A. Smith, Yejin Choi, and Hannaneh Hajishirzi. Vera: A general-purpose plausibility estimation model for commonsense statements, 2023.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach, 2019.

Ilya Loshchilov and Frank Hutter. Sgdr: Stochastic gradient descent with warm restarts. In *International Conference on Learning Representations*, 2016.

Eric Mitchell, Joseph J. Noh, Siyan Li, William S. Armstrong, Ananth Agarwal, Patrick Liu, Chelsea Finn, and Christopher D. Manning. Enhancing self-consistency and performance of pre-trained language models through natural language inference, 2022.

Fabio Petroni, Tim Rocktäschel, Patrick Lewis, Anton Bakhtin, Yuxiang Wu, Alexander H. Miller, and Sebastian Riedel. Language models as knowledge bases?, 2019.

Marin Vlastelica Pogancic, Anselm Paulus, Vít Musil, Georg Martius, and Michal Rolínek. Differentiation of blackbox combinatorial solvers. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net, 2020.

Oyvind Tafjord and Peter Clark. General-purpose question-answering with macaw, 2021.

Oyvind Tafjord, Bhavana Dalvi Mishra, and Peter Clark. Entailer: Answering questions with faithful and truthful chains of reasoning, 2022.

Antonio Vergari, YooJung Choi, Anji Liu, Stefano Teso, and Guy Van den Broeck. A compositional atlas of tractable circuit operations for probabilistic inference. *Advances in Neural Information Processing Systems*, 34:13189–13201, 2021.

David Wadden, Shanchuan Lin, Kyle Lo, Lucy Lu Wang, Madeleine van Zuylen, Arman Cohan, and Hannaneh Hajishirzi. Fact or fiction: Verifying scientific claims, 2020.

Jingyi Xu, Zilu Zhang, Tal Friedman, Yitao Liang, and Guy Van den Broeck. A semantic loss function for deep learning with symbolic knowledge, 2018.

Hanlin Zhang, Jiani Huang, Ziyang Li, Mayur Naik, and Eric Xing. Improved logical reasoning of language models via differentiable symbolic programming, 2023.

A TRAINING DETAILS

We fine-tune our models for 5 epochs keeping the learning rate fixed to $\gamma = 3 \cdot 10^{-4}$ on 1–2 nVidia A30 GPUs. Each model took approximately 35 minutes to train. We use AdamW (Loshchilov & Hutter, 2016) as optimizer with a default weight decay $\lambda = 10^{-2}$.

B EVALUATION DETAILS

B.1 EVALUATION ON THE TRAINING DISTRIBUTION

We implement the same evaluation functions from the original ConCoRD codebase (`https://github.com/eric-mitchell/concord/`). For ConCoRD, we report only the scores computed with the original code and the provided cache inference facts.

Table 2: **LoCo-LMs achieve better logical self-consistency and factuality** as measured via Equation (7) and F_1 scores when compared to classical supervised finetuning (SFT) and baselines using external reasoners such as ConCoRD (Mitchell et al., 2022) measured on train (calibration set) facts. For RQ1 (Section 3), LoCo-LMs fine-tuned on T1 facts only outperform training-free baseline for all metrics. For RQ2, they boost performance in the presence of a small fraction of T1+T2 facts (5-10%). For larger dataset sizes, LoCo-LMs are competitive for consistency and factuality on consequents.

	Method	Train size	Antecedents F_1	Consequents F_1	Total F_1	Logical consistency
RQ1	ConCoRD				0.91	0.91
	Macaw-Large		0.47	0.84	0.78	0.82
	SFT	T1	0.46	0.08	0.14	0.79
	LoCo-LM	T1	**0.98**	**0.99**	**0.99**	**1.00**
RQ2	SFT	T1+T2 (5%)	0.31	0.73	0.69	0.90
	LoCo-LM	T1+T2 (5%)	**0.34**	**0.77**	**0.72**	**0.92**
	SFT	T1+T2 (10%)	0.48	0.88	0.85	0.87
	LoCo-LM	T1+T2 (10%)	**0.52**	**0.95**	**0.89**	**0.91**
	SFT	T1+T2 (75%)	**0.69**	**1.00**	**0.97**	0.97
	LoCo-LM	T1+T2 (75%)	0.65	**1.00**	**0.97**	**0.99**

B.2 TEMPORAL EFFICIENCY ANALYSIS

Table 3: **LoCo-LMs require much less time for inference at the price of a single training step.** We report time elapsed in seconds for (1) training (calibration set only) our model; (2) inference (calibration + silver set) with our model or with ConCoRD+RoBERTa-ANLI. For ConCoRD, we sum time elapsed for QA inference, NLI inference and running the solver. All runs are measured on a single NVIDIA A30.

Method	Training time (s)	Inference time (s)
ConCoRD	-	3669.33
LoCo-LM	2124.48	**2405.28**

B.3 SEMANTIC OVERLAP BETWEEN TRAIN AND TEST

We measure the semantic overlap between the training and test distribution by constructing a Representation Dissimilarity Matrix (RDM) of Macaw's embeddings (token average) between training and test entities. The main assumption is that semantically similar subjects may have similar properties, as a proxy for domain knowledge transfer.

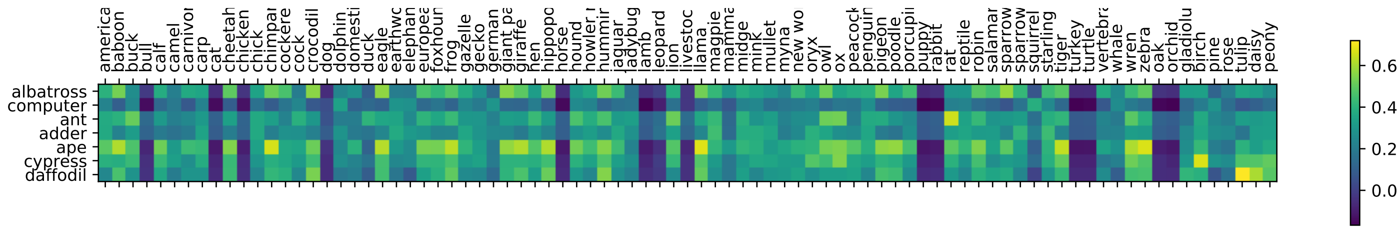

Figure 2: Pairwise cosine similarities between entities in the training distribution (calibration, rows) and the test distribution (silver, columns).

C DERIVATIONS

C.1 LOGICAL CONSTRAINTS WITH FACTUAL KNOWLEDGE

Here, we briefly discuss how we construct the semantic loss term for a given general constraint:

$$A \to B \iff \neg A \vee B$$

following the procedure outlined in Section 2. Given prior knowledge about the world, we embed it in the constraint via logical conjunction:

- Case $A = \top$:
  $$(A \to B) \wedge A \iff (\neg A \vee B) \wedge A$$
  We apply the DeMorgan's distributive laws:
  $$\begin{aligned} &(\neg A \vee B) \wedge A \\ \iff &(\neg A \wedge A) \vee (A \wedge B) \\ \iff &\perp \vee (A \wedge B) \\ \iff &A \wedge B \end{aligned}$$
  The semantic loss optimizes thus the satisfying cases:
  $$\mathcal{L}(\alpha, p_\theta) = -\log\left(p_\theta(A = \top) \cdot p_\theta(B = \top)\right)$$
  This similarly applies when $B = \perp$.
- Case $A = \perp$:
  $$(A \to B) \wedge A \iff (\neg A \vee B) \wedge \neg A$$
  We apply the DeMorgan's distributive laws:
  $$\begin{aligned} &(\neg A \vee B) \wedge \neg A \\ \iff &(\neg A \wedge \neg A) \vee (\neg A \wedge B) \\ \iff &\neg A \vee (\neg A \wedge B) \end{aligned}$$
  The semantic loss optimizes thus the satisfying cases:
  $$\mathcal{L}(\alpha, p_\theta) = -\log\begin{pmatrix} (1 - p_\theta(A = \top)) \cdot (1 - p_\theta(B = \top)) \;+ \\ (1 - p_\theta(A = \top)) \cdot p_\theta(B = \top) \end{pmatrix}$$
  This similarly applies to the case $B = \top$.