1. Supplementary Material

1.1. Choice of predicates

As discussed in Section 5, we used a subset of predicates for our primary experiments because the full 50 predicates represent a large number of synonyms and supersets for each predicate. We identified these dependencies between predicates as a directed graph, and selected the leaf nodes (bottom row) as our chosen predicates in Figure 1.

1.2. Performance on all 50 predicates

Furthermore, we have included results on the full set of 50 predicates in Table 1. Note that we are unable to evaluate against our primary baseline, transfer learning, because we have utilized all potential source domain predicates in this experiment. We see that our method improves over the baseline approach using $n = 10$ labeled examples per relationship by 15.46 R@100 for PREDCLS. We see similar trends across the various ablations of our model and therefore, only report the our best model.
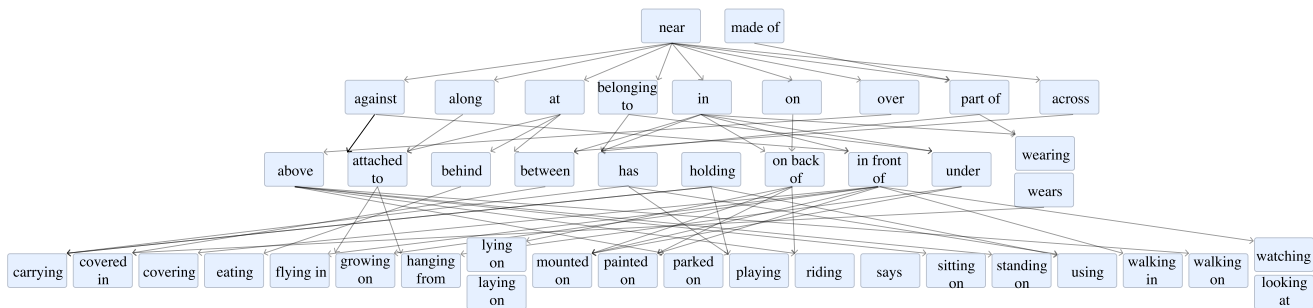
Figure 1. We define dependencies between predicates to determine which ones to include in the evaluation of our method. Directional arrows indicate supersets, and stacked nodes indicates synonyms. *Note:* `says` has no parents, so we treat this as a leaf node in our experimental setup.

Table 1. Results for top 50 predicates in Visual Genome.

Model	Scene Graph Detection			Scene Graph Classification			Predicate Classification		
	R@20	R@50	R@100	R@20	R@50	R@100	R@20	R@50	R@100
BASELINE [$n = 10$]	1.06	1.80	2.66	4.70	6.00	5.43	9.63	12.17	13.07
OURS (CATEG. + SPAT.)	**4.04**	**6.75**	**8.64**	**12.69**	**13.91**	**14.16**	**24.72**	**27.76**	**28.53**
ORACLE [$n_{\text{ORACLE}} = 44n$]	14.20	20.61	25.44	33.58	35.52	35.92	62.00	66.92	68.02