The Completion Time

Though we have employed two pruning strategies for acceleration, the detection of table changes accounts for the majority of the time for COMANTICS to complete a semantics inference task. Further, the detection of table changes highly depends on the size of the input and output tables. As a result, we design an additional quantitative experiment to assess how the completion time is affected by table sizes.

Data and Methods

To figure out how the size of input and output tables affects the efficiency of COMANTICS, we first divide the table-size space into five intervals, i.e., =0, (0, 100KB], (100KB, 1MB], (1MB, 5MB], and >5MB. As we discussed in the Characteristics Pruning strategy (see Section 4.1), most detection of table differences relies on the existence of input and output tables. Therefore, we set the first interval (i.e., =0) to see the efficiency of COMANTICS when there is no input or output table. Then we group the instances in our corpus (see Section 5.1) based on the size of their input and output tables under the five intervals. Note that if there are multiple input/output tables, we add up their table sizes. The distribution of the number of instances on the different table-size intervals is shown in Table S1. Finally, we record the average completion time in seconds that COMANTICS takes on these groups of instances, respectively.

Table S1. The number of instances under the different table-size intervals.

Input Size \ Output Size	=0	(0, 100KB]	(100KB,1MB]	(1MB, 5MB]	>5MB
=0	0	69	35	15	16
(0, 100KB]	0	452	2	2	0
(100KB,1MB]	0	37	92	3	0
(1MB, 5MB]	0	12	8	79	0
>5MB	0	9	5	9	76

Results and Analysis

Table S2. The average completion time in seconds that COMANTICS takes on different intersection groups of instances.

Input Size \ Output Size	=0	(0, 100KB]	(100KB,1MB]	(1MB, 5MB]	>5MB
=0		0.09	0.10	0.16	0.30
(0, 100KB]		2.76	3.72	7.86	
(100KB,1MB]		4.76	5.93	13.36	
(1MB, 5MB]		8.84	14.65	29.12	
>5MB		12.60	19.95	72.59	107.67

From table S1, we can find that all instances are with output tables. Besides, for nearly half of the instances (452/921, 49.08%), the sizes of their input and output tables are both fall into the (0, 100KB] interval.

From Table S2, we observe that the completion times in the first row (0.16s on average) are less than in others. Besides, we notice that as the size of input and output tables increases, the completion time COMANTICS takes increases drastically.