# Supplement to Chapter 8:
# Neuro Symbolic Reasoning with Differentiable Inductive Logic Programming

Supplment to Chapter 8 from Neuro Symbolic Reasoning and Learning - Current Advances and Future Directions

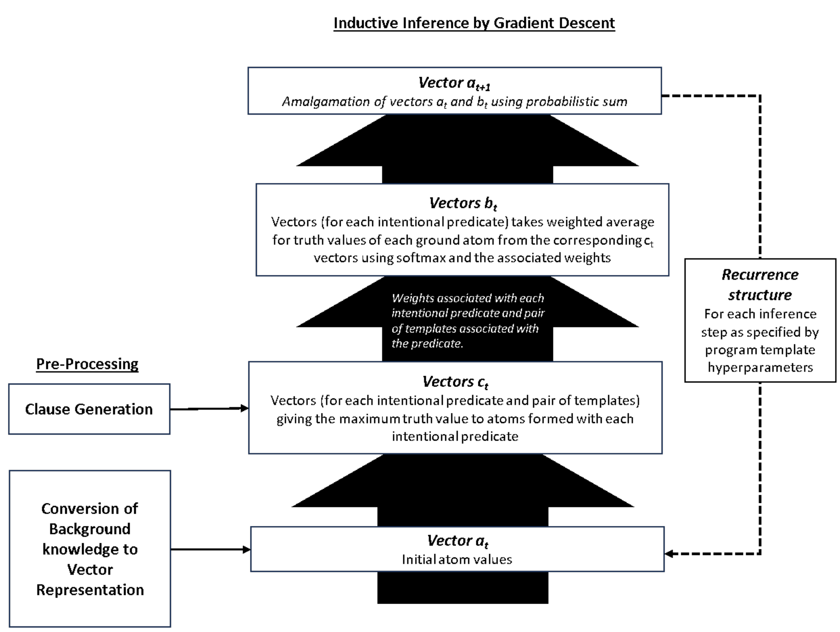


**Fig. 0.1** $\delta ILP$ architecture of [1].

## Complexity

In [1], the authors present an analysis of space (memory) complexity equal to the sum of the following two expressions:

$$2 \cdot n \cdot \mathcal{T} \cdot \sum_{i=1}^{|P_{in}|} \prod_{j}^{num_\tau} |cl(\tau_i^j)| \tag{0.1}$$

$$3 \cdot n \cdot |C| \cdot \sum_{i=1}^{|P_{in}|} \prod_{j}^{num_\tau} |cl(\tau_i^j)| \tag{0.2}$$

These correspond with memory requirements in term of the number of floats for primary vectors (Eq. 0.1) and intermediate vectors (Eq. 0.2) used in the neural architecture. Again, here $\mathcal{T}$ is the number of inference steps permitted and $n$ is defined as the number of ground atoms, which is bounded by the following:

$$n = |P| \cdot |C|^2 + 1 \tag{0.3}$$

Set $P_{in}$ is comprised of intensional predicates, which are predicates that only appear in rule heads (which includes invented predicates and any target predicates for which one wishes to make inferences about). The entire set of predicates ($P$) includes $P_{in}$ and $P_{ex}$, which are the extensional predicates (which only appear in rule bodies and never in the head). The number $num_\tau$ is the number of rules learned per intensional predicate, which in [1] is always set to 2.

The set $cl(\tau_i^j)$ contains the clauses (rules) generated by template $\tau_i^j$, and the size of this set is determined by several factors. First is the number of intensional predicates allowed in the body of each clause ($int$), and second is the number of variables that can be existentially quantified in the body ($v$). For $int, v$, we note that they are different for each intensional predicate, so we shall index these values in that case and use the subscript $max$ to refer to the max of such values. Third is the arity of the intensional predicates (which in [1] can be one of $\{0, 1, 2\}$ for each predicate, the maximum value of this we shall denote as $arity_i$ for the $i$ template and this is bounded by $arity_{max}$, which will be the maximum arity for any predicate), and fourth is the number of atoms per body (which is 2 in [1]; they argue that this is without loss of generality as more atoms per body can be permitted by creating more invented predicates; however it is noteworthy that doing so also increases computational complexity). We shall denote this last item with $body_{max}$.

Now we shall create a bound on $|cl(\tau_i^j)|$. We note that the total number of variables for a given template is $arity_i + int_i$. This means that for a given predicate in the body, there are at most $(arity_i + int_i)^{arity_{max}}$. The number of possible predicates for one of the atoms in the body is $|P_{ext}| + int_i$. We note that the number of variable combinations for a template and arity of $arity_i$ is $\frac{(arity_i+v)!}{v!} \leq (arity_i + v)^{arity_i}$. Hence, the number of possible variable arrangements and predicates for each atom is bounded by $(|P_{ext}| + int_i)(arity_i + v)^{arity_{max}}$. We can re-write the sum over all intensional predicates from Expressions 0.1 and 0.2 as follows:

$$\sum_{i=1}^{|P_{in}|} \prod_{j}^{num_\tau} |cl(\tau_i^j)| \leq \sum_{i=1}^{|P_{in}|} \prod_{j}^{num_\tau} \left((|P_{ext}| + int_i)(arity_i + v)^{arity_{max}}\right)^{body_{max}} \tag{0.4}$$

$$\leq |P_{in}| \cdot \left(|P|(arity_{max} + v)^{arity_{max}}\right)^{body_{max} \cdot num_\tau} \tag{0.5}$$

We note that based on expected use cases as expressed in [1], we have $num_\tau = 2$, $body_{max} = 2$, $arity_{max} = 2$, $v_{max} = 1$, and thus get the following.

$$6561 \cdot |P_{in}| \cdot |P|^4 \tag{0.6}$$

While this is an upper bound on space complexity, this is mainly due to smaller-arity sized predicates as well as pruning (e.g., two non-ground atoms create the same clause if their order is reversed). However, it is noteworthy that such exact pruning only reduces space complexity, but has minimal effect on time complexity (as such clauses are generated). We also note that the above result, especially taken in consideration with Expressions 0.1-0.3 is in line with a quintic run-time (in terms of predicates). We can re-write overall space complexity as being bounded by the following:

$$6561 \cdot (2 \cdot n \cdot \mathcal{T} + 3 \cdot n \cdot |C|) \cdot |P_{in}| \cdot |P|^4 \tag{0.7}$$

$$\leq 6561 \cdot (|P| \cdot |C|^2 + 1) \cdot (2 \cdot \mathcal{T} + 3 \cdot |C|) \cdot |P_{in}| \cdot |P|^4 \tag{0.8}$$

$$\approx 6561 \cdot (2 \cdot \mathcal{T} + 3 \cdot |C|) \cdot |P_{in}| \cdot |C|^2 \cdot |P|^5 \tag{0.9}$$

$$\leq K \cdot |P_{in}| \cdot |C|^3 \cdot |P|^5 \tag{0.10}$$

In line 0.10 (where $K$ is a large constant), it could also be the case that $\mathcal{T} > |C|$, and theoretically we may want this to be true (for correct reasoning); however, we note that in the current work on $\delta ILP$ , scalability precludes the full chaining of inference rules (and so $\mathcal{T}$ is typically set to a small natural number, e.g. $\mathcal{T} = 3$).

## References

1. Evans, R., Grefenstette, E.: Learning explanatory rules from noisy data. J. Artif. Int. Res. **61**(1), 1–64 (2018)