Abstracting Probabilistic Models: Relations, Constraints and Beyond (Extended Abstract)

Citation for published version:
Belle, V 2020, Abstracting Probabilistic Models: Relations, Constraints and Beyond (Extended Abstract). in *Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning.* 17th International Conference on Principles of Knowledge Representation and Reasoning, Rhodes, Greece, 12/09/20.

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Proceedings of the 17th International Conference on Principles of Knowledge Representation and Reasoning

Abstracting Probabilistic Models: Relations, Constraints and Beyond (Extended Abstract)*

Vaishak Belle
University of Edinburgh & Alan Turing Institute
vaishak@ed.ac.uk

1 Introduction

Abstraction is a powerful idea widely used in science to explain phenomena at the required granularity. Think of explaining a heart disease in terms of its anatomical components versus its molecular composition. Think of understanding the political dynamics of elections by studying micro level phenomena (say, voter grievances in counties) versus macro level events (e.g., television advertisements, gerrymandering). In particular, in computer science, it is often understood as the process of mapping one representation onto a simpler representation by suppressing irrelevant information. The motivation is three-fold:

(a) When representing complex pieces of knowledge, abstraction can provide a way to structure that knowledge, hierarchically or otherwise, so as to yield descriptive clarity and modularity.

(b) Reasoning over large graphs, programs, and other structures is almost always computationally challenging, and so abstracting the problem domain to a smaller search space is attractive. Even in the case of tractable representations, such as arithmetic circuits (Darwiche and Marquis 2002), reasoning is polynomial in the circuit size, so clearly a smaller circuit is more effective.

(c) Lastly, and perhaps most significantly, abstraction features pervasively in commonsense reasoning, and there is much discussion in the fields of cognitive science and philosophy on the role of abstractions for explanations (Dedre and Christian 2017). Thus, abstractions will likely be critical for *explainable AI* (Gunning 2016), and indeed, much of that literature focuses on extracting high-level symbolic and/or programmatic representations from low-level data.

Formal perspectives on abstraction have matured considerably over the years (Giunchiglia and Walsh 1992; Banihashemi, De Giacomo, and Lespérance 2017). In particular, the work of (Banihashemi, De Giacomo, and Lespérance 2017) is noteworthy as it identifies how notions of soundness and completeness relate to the model-theoretic properties of a high-level abstraction and the corresponding low-level theory. However, the formal analysis of abstraction has largely focused on categorical (deterministic and non-probabilistic) domains; that is, both the high-level and the low-level representations are assumed to be categorical assertions. In that regard, existing frameworks are not immediately applicable to the fields of probabilistic modeling and statistical machine learning. Indeed, we do not yet have a full understanding of which aspects of one probabilistic model, representing some low-level phenomena, can be omitted when building a less granular (possibly non-probabilistic) model standing for a high-level understanding of the domain.

In this paper, we provide a semantical framework for analyzing such abstractions from first principles. We develop the framework in a general way, allowing for expressive languages, including logic-based ones that admit relational, deterministic and hierarchical constructs with stochastic primitives (Heckerman, Meek, and Koller 2004; Getoor and Taskar 2007). Representative examples of such languages include probabilistic databases and statistical knowledge bases, which have received considerable attention both in the academic and industry circles (Suciu et al. 2011; Richardson and Domingos 2006; Wu et al. 2012; Dong et al. 2014; Niu et al. 2012; Carlson et al. 2010).

We first motivate a definition of consistency between a high-level (probabilistic or logical) model and its low-level (probabilistic) counterpart, but also treat the case when the high-level model is missing critical information present in the low-level model. We go on to prove properties of abstractions, both at the level of the parameter as well as the structure of the models. Put differently, we first motivate a definition of abstraction purely at the level of the model theory, which then provides the basis for analyzing the properties of "unweighted abstractions." (That is, probabilities are simply ignored in that construction.) We use that analysis to investigate how "weighted abstractions" can be defined. We then study how to incorporate low-level evidence and reason about it in the high-level representation. This latter point is particularly relevant to applications where observations are almost at the sub-symbolic/low-level, whereas the reasoning needs to be performed at the abstract/high-level. Finally, we consider some observations about how abstractions can be derived automatically, which itself rests on proving numerous cases where verifying that we have a reasonable abstrac-

*Full version available as: *Abstracting Probabilistic Models: Relations, Constraints and Beyond. Knowledge-Based Systems, 2020* (doi.org/10.1016/j.knosys.2020.105976). The author was supported by a Royal Society University Research Fellowship.

tion can be performed effectively.

2 Example

To motivate the framework using an example, consider a probabilistic relational model (PRM) on entity-relationships for a university database $\mathcal{U}$ (adapted from (Heckerman, Meek, and Koller 2004)). The model instantiates constraints for a (parameterised) Bayesian network:

$$\text{Difficulty} \longrightarrow \text{Grades} \longleftarrow \text{IQ}$$

Such a PRM might include weighted assertions such as: (a) **0.7** $\mathit{diff}(x, E)$; (b) **0.1** $\mathit{diff}(x, M)$; (c) **0.2** $\mathit{diff}(x, H)$; (d) **0.25** $iq(x, L) \wedge \mathit{diff}(y, E) \wedge takes(x, y) \supset grades(x, y, u)$ for $u \in \{7, 8, 9, 10\}$, and so on. Here, the constants E, M, H, L stand for *easy, medium, hard, low* respectively; (a) says that for any given course, the probability that its difficulty level is easy is 0.7, and (d) says that for any low IQ student taking an easy course, the probability that his grade is (say) 7 is 0.25. Let us refer to this as the low-level theory $\mathcal{U}_l$.

A simple yet powerful type of abstraction to apply here is to abstract away the domain. For instance, we can lump the constants $\{M, H\}$ as N, standing for *not easy*, and lump the mentioned grade values together as $\{0, \ldots, 6\}, \{7, 8\}, \{9, 10\}$ and denote them as B, O, G, standing for *bad, ok* and *good* respectively. Then, we would obtain the following model, referred to as the high-level theory $\mathcal{U}_h$: (a) **.7** $\mathit{diff}(x, E)$; (b) **.3** $\mathit{diff}(x, N)$; and (c) **.5** $iq(x, L) \wedge \mathit{diff}(y, E) \wedge takes(x, y) \supset grades(x, y, u)$ for $u \in \{O, G\}$.

The idea is that now the user reasons only with $\mathcal{U}_h$, and so the class of queries (e.g., conditional probabilities) of interest would only involve predicates and constants from $\mathcal{U}_h$. In that regard, using our framework, we can formally show that the two models agree on a large class of probabilistic queries. More generally, we can show that abstraction can be understood both from the viewpoint of the parameters (i.e., weights and/or probabilities) and structure (i.e., the logical sentences).

3 Framework

A formal theory of abstraction can be approached in three stages: first, how should abstraction be defined between a high-level representation Δ_h and a low-level one Δ_l? Second, given Δ_h and Δ_l, how do we prove that Δ_h is an abstraction of Δ_l? Third, given Δ_l and a target high-level vocabulary, how do we find Δ_h?

Our framework provides constructions to address such questions. To begin with, consider that the logical symbols (predicates and constants) may differ arbitrarily between Δ_h and Δ_l, and so we need to establish a "mapping" that defines how the symbols from Δ_h map to formulas in Δ_l. We then say Δ_h is an abstraction of Δ_l iff every model of Δ_h is "isomorphic" to some model of Δ_l (relative to the mapping), and vice versa. Roughly, if a model of Δ_h evaluates an atom to true, then the corresponding model of Δ_l should evaluate the mapping of that atom (to a low-level formula) to true as well.

From this simple definition, we consider notions under which the high-level and low-level theories agree on queries, but also how low-level evidence (e.g., readings on a sensor) can be incorporated at the high-level. These observations are then shown to lead to ideas on how abstractions can be derived automatically.

4 Conclusions

Our framework and results rest on the simple notion of isomorphisms between models, and the alignment of atomic probabilistic events. It is worth noting that abstraction is a major topic in knowledge representation (Giunchiglia and Walsh 1992; Erol, Hendler, and Nau 1996; Saitta and Zucker 2013; Banihashemi, De Giacomo, and Lespérance 2017), but it is also of interest in diverse areas such as statistics, program synthesis, and automated planning. Given this increasing interest in abstraction, we hope our framework will be helpful in developing probabilistic abstractions for increased clarity, modularity and tractability, and perhaps interpretability of statistical learning models.

References

Banihashemi, B.; De Giacomo, G.; and Lespérance, Y. 2017. Abstraction in situation calculus action theories. In *AAAI*, 1048–1055.

Carlson, A.; Betteridge, J.; Kisiel, B.; Settles, B.; Hruschka Jr, E. R.; and Mitchell, T. M. 2010. Toward an architecture for never-ending language learning. In *AAAI*, 1306–1313.

Darwiche, A., and Marquis, P. 2002. A knowledge compilation map. *Journal of Artificial Intelligence Research* 17:229–264.

Dedre, G., and Christian, H. 2017. Analogy and abstraction. *Topics in Cognitive Science* 9(3):672–693.

Dong, X.; Gabrilovich, E.; Heitz, G.; Horn, W.; Lao, N.; Murphy, K.; Strohmann, T.; Sun, S.; and Zhang, W. 2014. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 601–610.

Erol, K.; Hendler, J.; and Nau, D. S. 1996. Complexity results for htn planning. *Annals of Mathematics and Artificial Intelligence* 18(1):69–93.

Getoor, L., and Taskar, B., eds. 2007. *An Introduction to Statistical Relational Learning*. MIT Press.

Giunchiglia, F., and Walsh, T. 1992. A theory of abstraction. *Artificial intelligence* 57(2-3):323–389.

Gunning, D. 2016. Explainable artificial intelligence (xai). Technical report, DARPA/I20.

Heckerman, D.; Meek, C.; and Koller, D. 2004. Probabilistic models for relational data. Technical report, Technical Report MSR-TR-2004-30, Microsoft Research.

Niu, F.; Zhang, C.; Ré, C.; and Shavlik, J. W. 2012. Deepdive: Web-scale knowledge-base construction using statistical learning and inference. *VLDS* 12:25–28.

Richardson, M., and Domingos, P. 2006. Markov logic networks. *Machine learning* 62(1):107–136.

Saitta, L., and Zucker, J.-D. 2013. *Abstraction in artificial intelligence and complex systems*, volume 456. Springer.

Suciu, D.; Olteanu, D.; Ré, C.; and Koch, C. 2011. Probabilistic databases. *Synthesis Lectures on Data Management* 3(2):1–180.

Wu, W.; Li, H.; Wang, H.; and Zhu, K. Q. 2012. Probase: A probabilistic taxonomy for text understanding. In *Proc. Int. Conf. on Management of Data*, 481–492.