Probabilistic Circuits: Marginal Maximum a Posteriori Queries

Markus Pawellek
markus.pawellek@mailbox.org

Abstract

Marginal maximum a posteriori (MMAP) queries combine the aspects of both marginal (MAR) and maximum a posteriori (MAP) inference. Introducing marginal determinism, we are able to provide sufficient conditions for a probabilistic circuit (PC) to be tractable for MMAP queries. As a consequence, also the tractable computation of other query types, such as marginal entropy or mutual information, is made possible.

1 Introduction

The class of marginal maximum a posteriori (MMAP) queries is more advanced concerning the tractability for probabilistic circuits (PCs). Its computation complexity is highly dependent on the set of query variables and may typically be NP-hard. To reduce complexity, the structural property of marginal determinism is defined. Together with smoothness and decomposability, these are sufficient conditions to make MMAP queries tractable. Nevertheless, it is shown that these properties are indeed not necessary.

2 Preliminaries and Review

Marginal (MAR) queries in the following sense are of paramount importance when we want to reason about state of the world where not all random variables are fully observed.

$$p(E = e, Z \in I) = \int_I p(z, e) \, \mathrm{d}Z$$

For a PC, the ability to tractably compute MAR queries is equivalent to its smoothness and decomposability.

Maximum a posteriori (MAP) queries relate to the mode of the distribution in the following sense.

$$\arg\max_{q \in \mathrm{val}(Q)} p(q \mid e) = \arg\max_{q \in \mathrm{val}(Q)} p(q, e)$$

The tractable computation of MAP queries in a PC is characterized by the properties consistency and determinism.

Decomposability implies consistency. As a consequence, we would expect a PC needs to fulfill smoothness, decomposability, and marginal determinism.

3 Marginal Maximum A-Posteriori Queries

Definition: (MMAP Query Class)

$$\arg\max_{q \in \mathrm{val}(Q)} p(Q = q | E = e, Z \in I)$$

$$\arg\max_{q \in \mathrm{val}(Q)} \int_I p(q, e, z) \, \mathrm{d}z$$

4 Marginal Determinism

Marginal Determinism is a simple generalization of determinism.

Definition: (Marginal Determinism)

Let $Q \subset X$. A sum node is marginal deterministic with respect to Q if for any partial state $q \in \mathrm{val}(Q)$, the output of at most one of its input units is nonzero. A PC is marginal deterministic with respect to Q if all of its sum nodes containing variables in Q are marginal deterministic.

> **THEOREM:** (MMAP Conditions)
>
> *Let $Q \subset X$ and $\mathcal{G}$ be smooth, decomposable, and marginal deterministic with respect to Q PC. Then for any parameterization the sum-maximizer circuit of $\mathcal{G}$ tractably computes $MMAP$ queries over Q.*

PROOF:
Lorem ipsum dolor sit amet, consectetur adipisicing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum. □

5 Algorithm

With respect to the previous algorithms for MAR and MAP queries, only the computation for sum nodes changes. Let n be a sum unit in a feed-forward computation. Then its result r_n can be computed by the following.

$$a := \sum_{c \in \mathrm{in}(n)} \vartheta_{nc} r_c$$

$$b := \max_{c \in \mathrm{in}(n)} \vartheta_{nc} r_c$$

$$r_n = \begin{cases} a & : \varphi(n) \cap Q = \emptyset \\ b & : \text{else} \end{cases}$$

6 Application to Marginal Entropy

> **THEOREM:** (Marginal Entropy Computation)

7 Expressive Efficiency

In contrast to other structural properties, such as decomposability or determinism, marginal determinism is defined with respect to its query set Q. For $Q = \emptyset$, it corresponds to smooth and decomposable PCs (tractable for MAR). For $Q = X$, it corresponds to smooth, decomposable, and deterministic PCs (tractable for both MAR and MAP). The latter is more expressive efficient. Other subsets may not be as expressive efficient with respect to either a subset or a superset. Assuming all properties for all possible subsets of Q, full-support distributions would have to be fully factorized. This explains the restriction to the query set Q.

8 Conclusions

References

Choi, YooJung, Antonio Vergari, and Guy Van den Broeck (October 2020). "Probabilistic Circuits: A Unifying Framework for Tractable Probabilistic Models". In: URL: http://starai.cs.ucla.edu/papers/ProbCirc20.pdf.