Building an interface between probabilistic programming languages and lumen

Jonas Aaron Gütter
Friedrich Schiller Universität Jena
Matrikelnr 152127
Prof.Dr. Joachim Giesen
M. Sc. Phillip Lucas

27. August 2018

Zusammenfassung

(1) grober Fahrplan

- read related material:
  - understand what Probabilistic Programming Languages (PPLs) are
  - understand main idea of Lumen and what we want to do with it
  - understand and formulate "why & what"
- choose a PPL
  - work out requirements to chose PPL
  - work out preferred (but not necessarily required) features
  - chose a PPL based on these requirements and preferences
- get started with PPL
  - play around, learn how to use it, what it can do, etc also confirm identified 'pain point', i.e. understand and formulate what problem you are trying to solve, why this is relevant and outline how you plan to solve that pain point
- design a wrapper of chose PPL for Backend of Lumen
- give presentation about work so far, its justification, relevance, verification ideas, etc etc
- implement and test wrapper
- evaluate implementation in terms of goals set in beginning

(2) interessante PPLs:

- stan for python: https://pystan.readthedocs.io/en/latest/
- pymc3: https://docs.pymc.io/notebooks/getting_started.html#Case-study-2:-Coal-mining-disasters

- edward: http://edwardlib.org/getting-started

- pyro: http://pyro.ai/

Wegen eines Arbeitsplatzes und eines PCs erkundigen wir uns. Als Anmelde und Starttermin für Deine MA halten wir Mitte September (Mo, 17. September) im Auge.

Glossar

probabilistic model is a part of a probabilistic reasoning system. 3

1 Probabilistic Inference

There are three rules of probabilistic inference: The chain rule, the total probability rule, and the Bayes' rule. The following explanations are taken from [**?**].

1.1 Chain rule

The chain rule is used to calculate a joint probability distribution of several variables from local conditional probability distributions of these variables:

$$P(X_1, X_2, ...X_n) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2)...P(X_n|X_1, X_2, ...X_{n-1})) \tag{1}$$

1.2 Total probability rule

The total probability rule calculates the probability distribution over a subset of variables by summing out all the other variables, that is by summing the probability distributions for each combination of values of these variables:

$$P(\boldsymbol{X}|\boldsymbol{Z}) = \sum_{\boldsymbol{y}} P(\boldsymbol{X}, \boldsymbol{Y} = \boldsymbol{y}|\boldsymbol{Z}) \tag{2}$$

1.3 Bayes' rule

2 What is Probabilistic Programming

Modelle spezifizieren/beschreiben
[1]
effizienter in der Beschreibung von Modellen als herkömmliche Programmiersprachen [2]
unifying general purpose programming with probabilistic modeling [3]
probabilistic model

3 Existing PPLs

3.1 Stan for python

3.2 Pymc3

python library

fit Bayesian models, including Markov Chai Monte Carlo (MCMC) and variatinal inference (VI)

3.3 Edward

3.4 Pyro

4 Lumen

5 Literatur

Literatur

[1] Wikipedia contributors, "Probabilistic programming language — Wikipedia, the free encyclopedia," 2018. [Online; accessed 23-August-2018].

[2] L. Hardesty, "Probabilistic programming does in 50 lines of code what used to take thousands," Apr. 2015. [Online; accessed 23-August-2018].

[3] "probabilistic-programming.org." [Online; accessed 23-August-2018].