

# 論文の内容の要旨

論文題目: Efficient Inference and Training  
for Conditional Latent Variable Models  
(条件付隠れ変数モデルのための効率的な推論と学習)

SUN, Xu  
孫 栩  
ソン シュウ

## ABSTRACT

Real-world problems may contain latent dependencies (i.e., hidden sub-structures) that are difficult to capture with conventional structured classifiers, such as conditional random fields. In such cases, models that exploit latent variables are advantageous in learning, and conditional latent variable models have been applied successfully into real-world tasks in natural language processing and vision processing communities, including syntactic parsing and vision recognition. In the first part of this thesis, I perform experiments in a variety of tasks to confirm the advantages of conditional latent variable models, and investigate what kind of latent dependencies are learned by the model.

While the experiments confirmed the advantages of conditional latent variable models, the same experiments revealed two problems in applying such models for practical usages. First, establishing an efficient inference method on latent conditional models remains an open question. Second, because of the incorporation of latent variables, training a latent conditional model brings a heavy computational cost. To deal with those critical problems, I propose efficient methods for inference and training on latent conditional models.

In the inference stage, I propose the "latent-dynamic inference", which is able to produce the exact optimal label sequence on latent conditional models by systematically combining efficient search strategy (the A\* search) and dynamic programming (the Forward-Backward method). I also describe a straightforward solution on approximating the exact method, and show that the approximated version performs as well as the exact one, while the speed is much faster.

In the training stage, I propose a perceptron-style algorithm for fast online training of conditional latent variable models. Our method extends the perceptron algorithm for the learning task with latent dependencies. It relies on Viterbi decoding over latent variables, combined with simple additive updates. Compared to existing probabilistic training methods on conditional latent variable models, our training algorithm lowers the training cost significantly yet with comparable or even superior classification accuracy. I also analyze the convergence properties of the proposed training method, and show it is guaranteed to converge with a bound.

To evaluate the efficiency and accuracy of our proposals, I performed experiments on a variety of natural language processing tasks as well as an artificial data. Our experiments demonstrate that the proposed inference and training methods achieve good experimental results, and their computational costs are much lower than the traditional methods.

## 論文要旨

実世界の問題には、条件付確率場など従来の機械学習モデルでは表現しきれない隠れた依存関係が含まれることがある。このような問題に対しては、隠れ変数を用いたモデルが有効である。自然言語処理および画像処理の分野では、統語解析、画像認識など現実的な問題に条件付隠れ変数モデルが応用され、その有効性が確かめられている。本論文の最初の部分では、条件付き隠れ変数モデルをさらに様々な問題に応用し、その有効性を示すとともに、隠れた依存関係としてどのようなものが学習され、精度向上に寄与しているのかを検討する。

上記の実験は、条件付き隠れ変数モデルの有用性を示す一方で、このモデルを実用化するにあたって、2つの問題が存在すること明らかにした。ひとつは、条件付き隠れ変数モデルの下での推論の際の計算コストが非常に大きいという問題、もうひとつは、モデルの学習に必要な時間が非常に長いという問題である。本論文の後半では、これらの問題を解決するため、条件付隠れ変数モデルのための効率的な学習および推論の手法を提案する。

まず、条件付隠れ変数モデルを用いた系列ラベリングのための高速な推論手法を提案する。この手法では、効率的な探索と動的計画法を組み合わせることにより、最適な出力ラベル列(厳密解)を求める。また、この手法を一部変更し、高品質な近似解を厳密解法に比べはるかに高速に求める手法を提案する。

次に、パーセプトロンと類似したオンライン学習アルゴリズムを、条件付隠れ変数モデルに適用することを提案する。この手法では、隠れ変数列に対するビタビア アルゴリズムと加算形の重み更新ルールを組み合わせることにより、効率的な学習を実現する。最尤推定に基づく従来の手法と比べた場合、提案手法は学習コストを大幅に削減し、かつ、同等以上の分類精度を実現する。また、提案手法による重みパラメータの収束に関して理論的な解析を行い、パラメータ更新の回数に関する上限値が存在することを示す。

評価実験として、提案した学習および推論手法を種々の自然言語処理タスクおよび人工データに適用し、精度の高い結果が従来法よりもはるかに効率的に得られることを示す。