1 Objective
Structured data, made up of interdependent components, is abundant. The goal is to exploit these interdependencies to improve the performance of prediction machines.

2 Learning & Prediction
Supervised learning can be understood as fitting a function to some observed input/output pairs.

If the learned function generalises well from the observed examples, we expect our predictions on new, unseen data to be accurate.

For discrete output spaces the learning task is called classification. A typical classification example is handwritten character recognition.

3 Exploiting Interdependencies
Classical machine learning systems treat interdependent components, like letters in a word, separately. Only recently, methods have been developed to incorporate these interdependencies, hoping for better predictions.

These structured prediction models have numerous applications, e.g. in web page classification, machine translation, 3D protein folding prediction, ...

4 Discriminative Models for Semi-Structured Data
Most successful amongst the structured prediction models are discriminative approaches like conditional random fields [1] and maximum margin Markov networks [2].

We research the use of these models for semi-structured data like XML documents, exploiting both structure and content of a file.

5 Experiments, Preliminary Results & Future Work
After having implemented a first version of a maximum margin Markov network, we carried out information extraction experiments for semi-structured data. Although the results need to be viewed as being preliminary, they confirm the validity of our approach.

A comparison of existing algorithms and their extension to a semi-supervised structured setting is left for future work.

6 Acknowledgements
The author gratefully acknowledges support through a research studentship from Microsoft Research Ltd.

7 References