

The University of Chicago

Department of Statistics

Seminar

---

Robert McCulloch

Graduate School of Business  
The University of Chicago

**BART: Bayesian Additive Regression Trees**

**Monday, February 23, 2004 at 4:00 PM**  
**133 Eckhart Hall, 5734 S. University Avenue**

## ABSTRACT

H. A. Chipman, E.I. George, R.E. McCulloch

We develop a Bayesian approach to fitting a model  $y = f(x) + \epsilon$  where  $f$  is represented by the sum of trees. We build upon the previous work in [1] to develop an MCMC algorithm for drawing from the posterior. Even though the model is complex and flexible, a relatively simple prior specification allows us to control the fit.

The model and associated algorithm have some interesting similarities to Boosting and backfitting. If the priors are set so as to heavily regularize individual trees, we see Boosting-like behaviour with a large number of weak learners, each contributing a small amount to the overall model. If instead we relax the regularization, then a smaller number of additive trees will contribute to the model. The iterated draws of each tree conditional on others is similar to a Bayesian version of backfitting [2].

The Bayesian framework and MCMC algorithm yield a posterior distribution, which can be used to assess uncertainty and predict using model averaging rather than model selection.

## References

- [1] Chipman, H., George, E. and McCulloch, R. (2002) "Bayesian Treed Models", *Machine Learning*, 48, 299-320.
- [2] Hastie, T. , and Tibshirani, R. (2000), "Bayesian Backfitting (with comments and a rejoinder by the authors)", *Statistical Science*, 15 (3) , 196-223