Recently, it has been shown:

1. AdaBoost does not necessarily maximize the margin (RSO 06)

2. Pätsch & Warmuth's bound is tight, in that
   AdaBoost (in a special case) can achieve a
   margin of exactly $\mathcal{I}(\phi)$. (RSO 06, 07)

It also turns out that Breiman was right:

1. arc-gd does maximize the margin (Breiman R+W)
   with a fast convergence rate (RSO 07)

2. AdaBoost still beats arc-gd experimentally,
   (Breiman, Regazzii & Shapire 06)

In fact, there have been a number of other algorithms
(with fast convergence rates) designed to maximize the margin,
none of which has been shown to beat AdaBoost experimentally:

- arc-gd (Breiman 99)
- AdaBoost $\gamma>1$ AdaBoost* (Pätsch & Warmuth 05)
- Optex, Coed Boosting (RSO 07)
- LP-AdaBoost (Coex & Sharmanski 06)

There are many unsolved problems - this is the current
state of theoretical research on generalization of AdaBoost!

Experimental work: There are lots of variations of
AdaBoost (almost all heuristically based, all
claim to beat AdaBoost.)

One more important fact about AdaBoost:

It solves the separate ranking problem at the
same time as it solves the problem of classification.

$\mathcal{I}(\phi)$