Unifying the Two Cases

[with Collins & Singer]

- two distinct cases:
  - weak learning assumption holds
    - $\mathcal{P} = \emptyset$
    - dynamics unclear
  - weak learning assumption does not hold
    - $\mathcal{P} \neq \emptyset$
    - can prove convergence of D_t's
- to unify: work instead with unnormalized versions of D_t's
  - standard AdaBoost: $D_{t+1}(i) = \dfrac{D_t(i)\exp(-\alpha_t y_i h_t(x_i))}{\text{normalization}}$
  - instead:

$$d_{t+1}(i) = d_t(i)\exp(-\alpha_t y_i h_t(x_i))$$

$$D_{t+1}(i) = \frac{d_{t+1}(i)}{\text{normalization}}$$

  - algorithm is unchanged