Probab "p" is wrong

~~not~~ $p = w_t$

Formalize names

- OSC agent

- meta agent

  - Prob of Action

  $$P(OA) = \sum w_{At}$$

  - Punishment $\longrightarrow$ reduction of weight

  α : learning rate

See Pic

Learning Auto Nomen

Reward:

Penalty:

Reward — Penalty : use both of below

Linear Reward in action : $W_{act}\uparrow$ if Rew

$W_{act}\Delta$ in (no update) if no Rew

Penalty in action : Penal : $W_{act}\downarrow$

R : W_i no update

β normally: $\beta = 0.1 \alpha$

Not a fan of primes

Try Redundant DSCS