Continuous Action

Stewart W. Wilson
Prediction Dynamics, Concord, MA 01742
Department of General Engineering
University of Illinois at Urbana-Champaign IL 61801
wilson@prediction-dynamics.com

What are continuous actions?

- LCS actions are typically discrete:
  1/0, left/right/forward
- *Continuous* means system can choose action from a continuous range:
  turn 34.7 degrees
  reach 224.3 cm
- As always, action is chosen to maximize payoff.
- Aim—improved performance compared with discrete.

How can you do it?

$t(x), a \Rightarrow p$	Traditional classifier
$t(x) \Rightarrow p(x)$	Function approximation (XCSF)
$t(x), a \Rightarrow p_a(x)$	Add discrete actions (XCS-LP)
$t(x, a) \Rightarrow p(x, a)$	Continuous x and continuous a

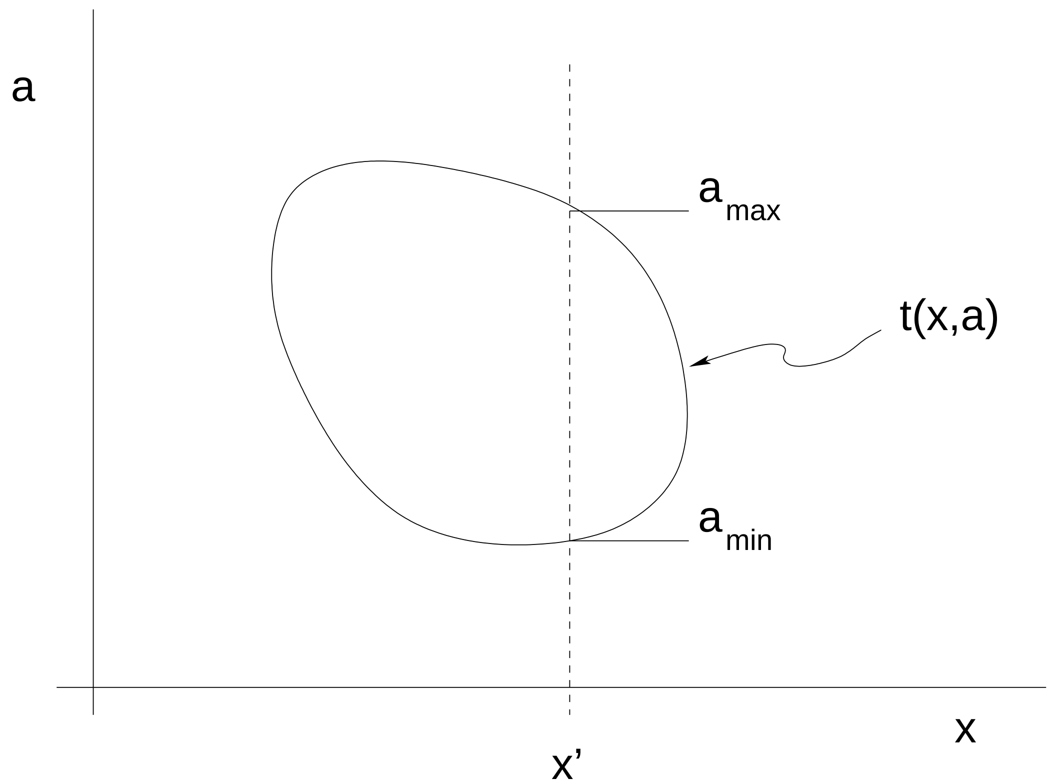
Best action, $a^*(x)$, is

$$a^*(x) = a : t(x, a)\,\text{true} \wedge p(x, a) = \max_{a' \in A} p(x, a')$$

How do you calculate a*(x)?

Assume landscape $P(x, a)$ has been sufficiently *sampled* (by exploration), so we have classifiers $t(x, a) \Rightarrow p(x, a)$.

Suppose $x = x'$, and consider a classifier's condition:



Let $p(x, a)$ be a *linear approximation*, I.e.,

$$p(x, a) = w_0 + w_1 x + w_2 a$$

Then $a^*(x) = a_{\text{max}}$ if $w_2 a_{\text{max}} \geq w_2 a_{\text{min}}$
a_{min} otherwise

Note: $a^*(x)$ depends *continuously* on x.

What functions for t(x,a)?

Butz introduced *general hyperellipsoidal* conditions for LCS function approximation.

They are true if

$(M(X-C))^2 < -2ln\theta$

The boundary forms a quadratic polynomial in $x_0, ..., x_n$.

For $t(x,a)$, we can use these functions, letting $a = x_0$.

How do you solve for a_{max}?

Once $x_1, ..., x_n$ are given (i.e., as input),

$(M(X - C))^2 = -2ln\theta$

becomes a quadratic equation in $x_0 = a$.

Its two roots are a_{max} and a_{min} !