On the traditional classifier system...

- Holland, J. H. (1986). In *Machine Learning, An Artificial Intelligence Approach. Volume II*.
- Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*.
- Lashon Booker
- Stephanie Forrest
- John Holmes
- Tim Kovacs
- Rick Riolo
- Robert Smith
- Tom Westerdale
- (Stewart Wilson)
- Many others

What about real inputs?

- Temperature, concentration, age, ...
- Binary variable encodings, "enumeration encoding", Fuzzy (Bonarini)
- Use of reals internally: get accurate ranges and thresholds, see if it can be done

What to use as a test problem?

- "Real 6-Multiplexer" (!)
- Boolean 6-multiplexer: Example 011010 $\rightarrow$ 0

  $F_6 = b_0' b_1' b_2 + b_0' b_1 b_3 + b_0 b_1' b_4 + b_0 b_1 b_5$
- Make "RF_6"

  $x = (x_0, ..., x_5), 0.0 \leq x_i < 1.0$

  Thresholds $0.0 \leq \theta_i < 1.0$

  <u>Interpret</u> x_i as 0 if $x_i < \theta_i$, else 1; apply F_6

How does XCSR *differ* from XCS?

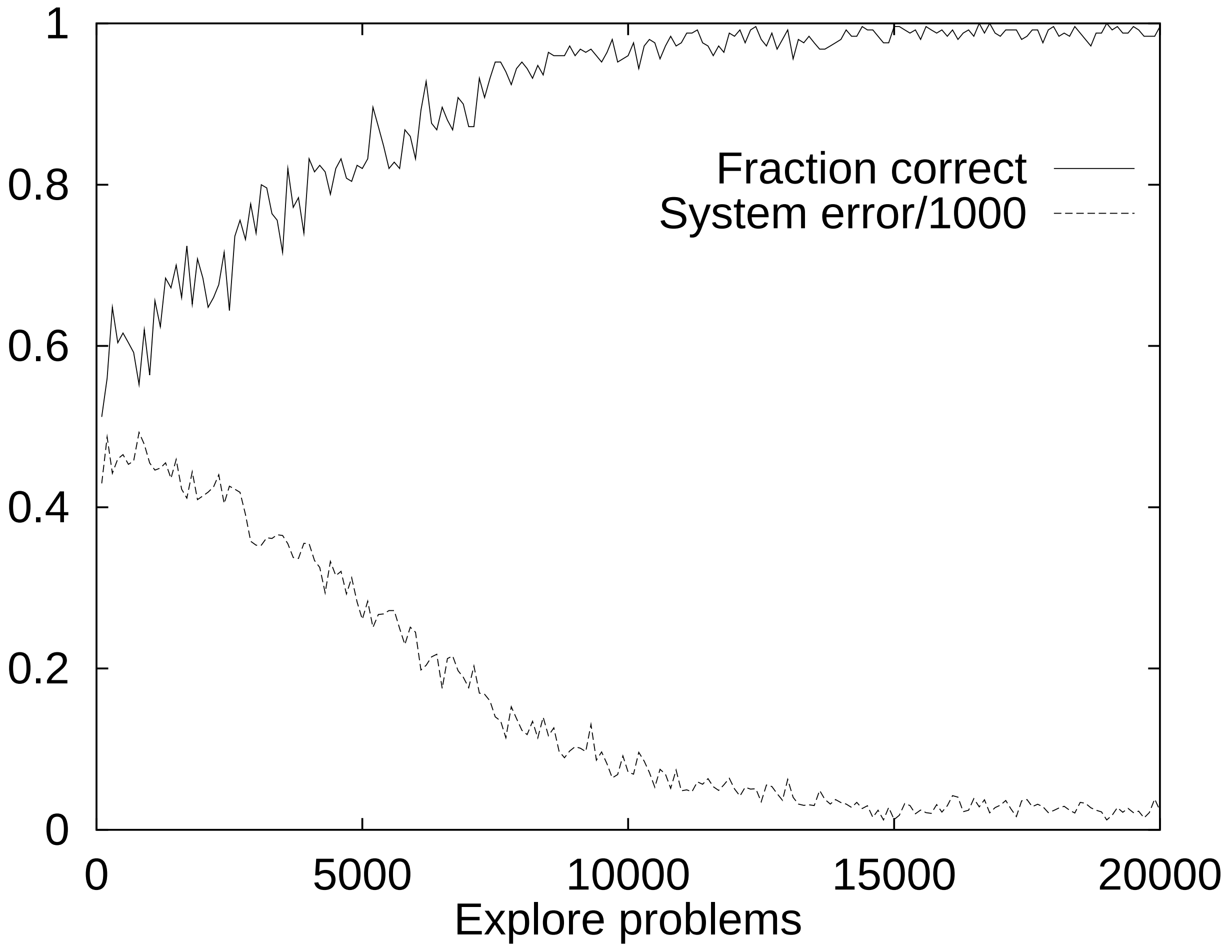
- Condition consists of "interval predicates" $int_i = (c_i, s_i)$, so condition $= (c_0, s_0, ..., c_5, s_5)$

- Classifier matches iff $c_i - s_i \leq x_i < c_i + s_i, \forall x_i$

- Mutation adds $\pm rand(m)$ to allele

- "Covering" creates classifier with condition in which $c_i = x_i$ and $s_i = rand(s_0)$

How about an experiment?



Experiment 1, all $\theta_i = 0.5$
Reward $R = 1000$ for correct, 0 for incorrect

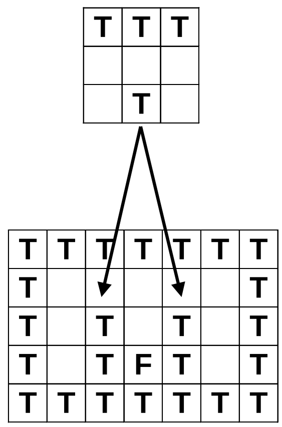
An action set [A] from Experiment 1

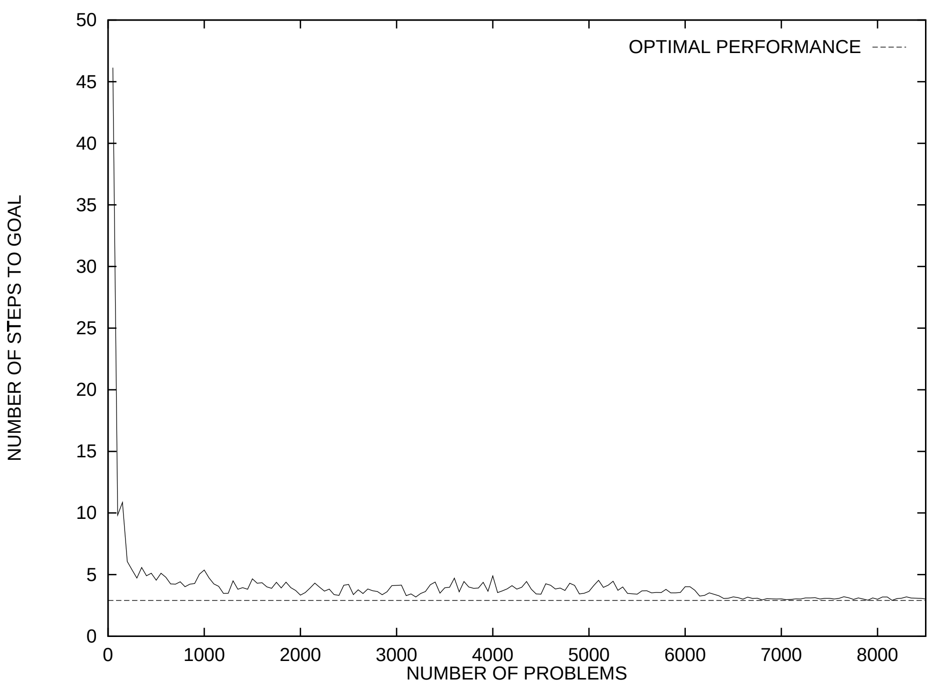
```
                                                                       ACT  PRED   ERR  FITN  NUM
 0. |OOOOOOOOOO|.....oOOOO|OOOOOOOOOO|.....oOOOO|oOOOOOOOOO|OOOOOOOo..| 1    0.  .000   14.   1
 1. |.....oOOOO|.....oOOOO|OOOOOOOOOO|OOOOOOOOOO|oOOOOOOOOO|OOOOo.....| 1    0.  .000   53.   2
 2. |.....oOOOO|.....oOOOO|OOOOOOOOOO|.....oOOOO|OOOOOOOOOO|OOOOo.....| 1    0.  .000   40.   1
 3. |.....oOOOO|.....oOOOO|OOOOOOOOOO|OOOOOOOOOO|oOOOOOOOOO|OOOOo.....| 1    0.  .000   50.   1
 4. |.....oOOOO|.....oOOOO|OOOOOOOOOO|OOOOOOOOOO|OOOOOOOOOO|OOOOo.....| 1    0.  .000   50.   1
 5. |.....oOOOO|.....oOOOO|OOOOOOOOOO|OOOOOOOOOO|OOOOOOOOOO|OOOOo.....| 1    0.  .000  140.   3
 6. |.....oOOOo|OOOOOo....|OOOOOOOOOO|OOOOOOOOOO|OOOOo.....|OOOOOOOOOO| 1   34.  .081    5.   2
 7. |....oOOOOO|.....oOOOO|OOOOOOOOOO|OOOOOOo...|OOOOOOOOOO|OOOOo.....| 1    0.  .000   56.   2
 8. |.....oOOOO|.....oOOOO|OOOOOOOOOO|....oOOOOO|OOOOOOOOOO|OOOOo.....| 1    0.  .000   41.   1
 9. |....oOOOOO|.....oOOOO|OOOOOOOOOO|.....oOOOO|OOOOOOOOOO|OOOOo.....| 1    0.  .000   58.   1
10. |.....oOOOO|.....oOOOO|OOOOOOOOOO|.....oOOOO|OOOOOOOOOO|OOOOo.....| 1    0.  .000   46.   1
11. |.....oOOOO|.....oOOOO|OOOOOOOOOO|....oOOOOO|OOOOOOOOOO|OOOOo.....| 1    0.  .000   85.   2
12. |.....oOOOO|.....oOOOO|OOOOOOOOOO|.....oOOOO|OOOOOOOOOO|OOOOo.....| 1    0.  .000   43.   1
```

```
                                                                 ACT  PRED   ERR  FITN  NUM
 0. |.572,.985|.924,.393|.322,0.99|.948,.417|.818,.812|.331,.404| 1    0.  .000   14.   1
 1. |.786,.264|0.89,.364|.602,0.99|0.23,.884|.796,.769|.228,.268| 1    0.  .000   53.   2
 2. |.794,.264|0.89,.364|.262,0.99|.868,.344|.665,.769|.228,.268| 1    0.  .000   40.   1
 3. |.794,.264|.807,.262|.602,0.99|0.23,.884|.796,.769|.228,.268| 1    0.  .000   50.   1
 4. |.794,0.28|.807,.262|.684,0.99|0.23,.884|.717,.769|.228,.268| 1    0.  .000   50.   1
 5. |.794,.264|.807,.262|.602,0.99|0.23,.884|.717,.769|.228,.268| 1    0.  .000  140.   3
 6. |.743,.232|.172,.404|.813,.903|0.41,.841|.092,.366|.506,.658| 1   34.  .081    5.   2
 7. |.775,.332|.807,.262|.476,.687|.275,.344|.716,.874|.205,.233| 1    0.  .000   56.   2
 8. |.786,.264|.807,.262|.288,0.99|.818,.322|.717,.783|.181,.269| 1    0.  .000   41.   1
 9. |.798,.357|0.89,.364|.247,0.99|.894,.344|.665,.732|.207,.233| 1    0.  .000   58.   1
10. |.798,.264|0.89,.364|.247,0.99|.894,.344|.665,.732|.207,.269| 1    0.  .000   46.   1
11. |.798,.264|.807,.262|.288,0.99|.818,.322|.717,.783|.207,.269| 1    0.  .000   85.   2
12. |.798,.264|.807,.262|.288, 1.0|.818,.274|.717,.783|.207,.269| 1    0.  .000   43.   1
```

Input $\vec{x} = (0.72, 0.55, 0.33, 0.57, 0.14, 0.27)$
Boolean interpretation: 110100

Woods101 (= McCallum's Maze)





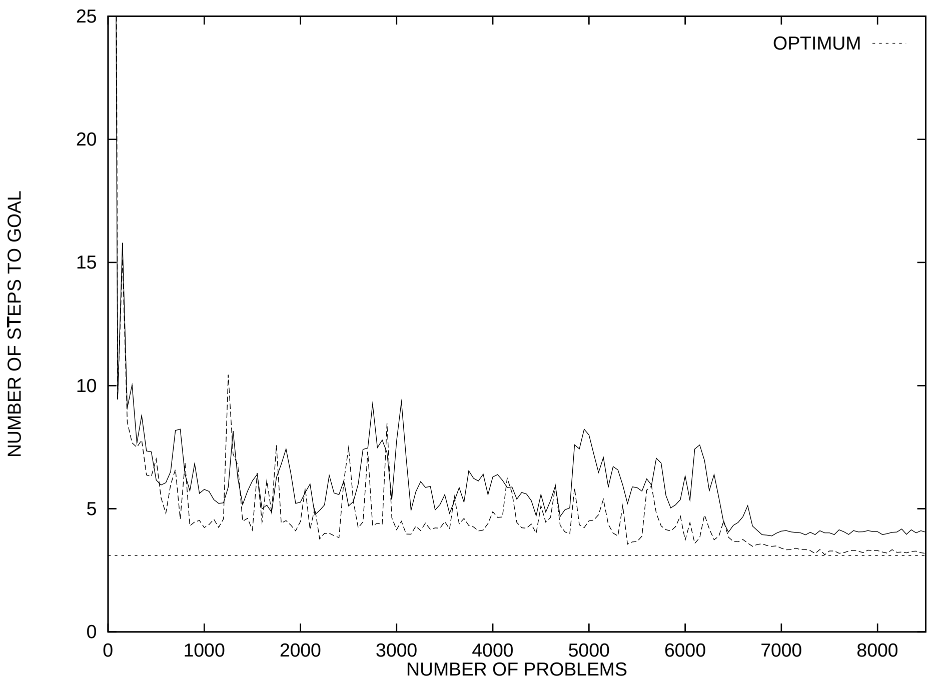
Woods101.5

T	T	T	T	T	T	T
T		T	F	T		T
T		T		T		T
T	T		T		T	T
T		T		T		T
T	T	T	T	T	T	T
T		T		T		T
T	T		T		T	T
T		T		T		T
T		T	F	T		T
T	T	T	T	T	T	T

(a)

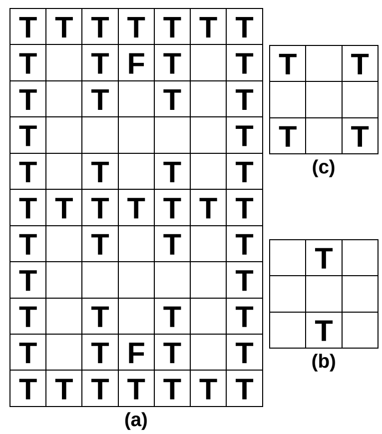
	T	
T		T
	T	

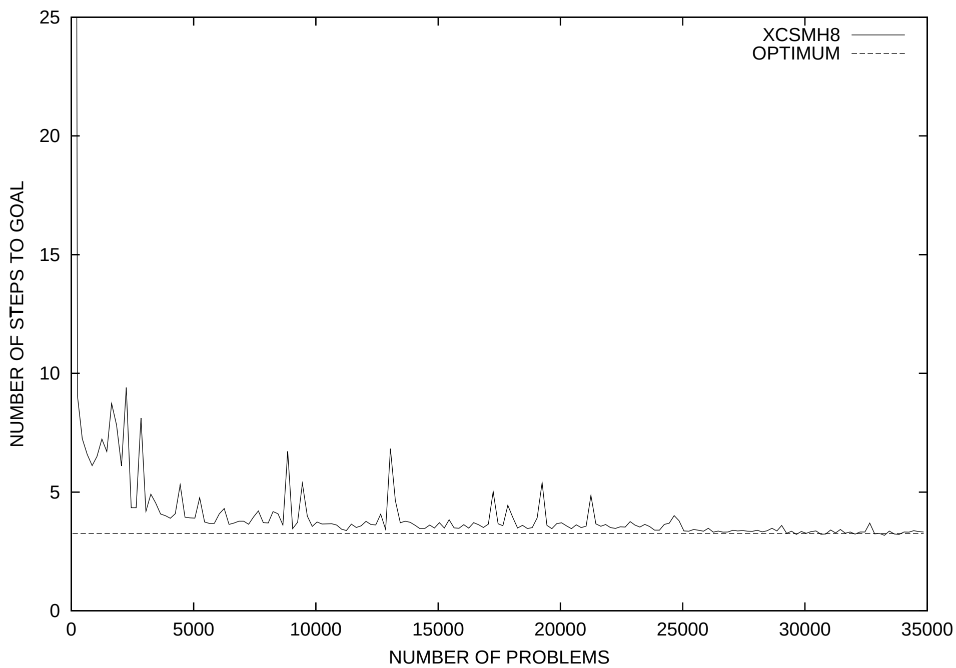
(b)



Optimum reached with register redundancy (4 bits vs. 2).

Woods102





Uses 8-bit register.

How is XCS different from other RL systems?

Rule-based, not PDP ("parallel distributed processing")

- Structure is created as needed
- Learning may often be faster because classifiers are inherently non-linear
- Learning complexity may be less than most PDP's
- Classifiers can keep and use statistics; difficult in a network
- Can "see the knowledge"
- Hierarchy and reasoning may be easier, since knowledge is in the form of rules