

# Genetic Programming and Symbolic Regression

Trent McConaghy, PhD

ascribe@

**solido**  
DESIGN AUTOMATION

*Mysteries of the  
universe..*

**What does AI  
encompass?**



**Is Deep Learning  
cool or what?**

**WTF is genetic  
programming or  
symbolic regression?  
Why should I care?**



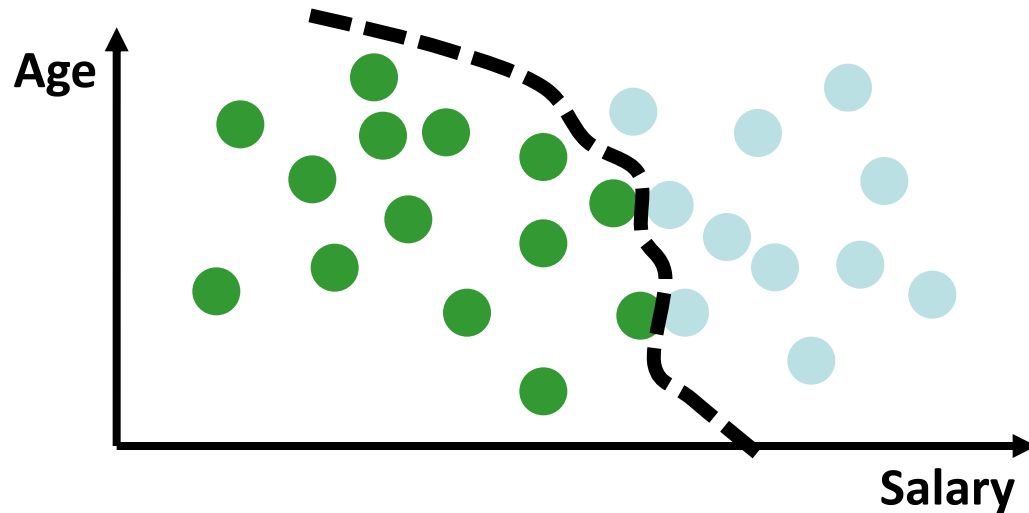
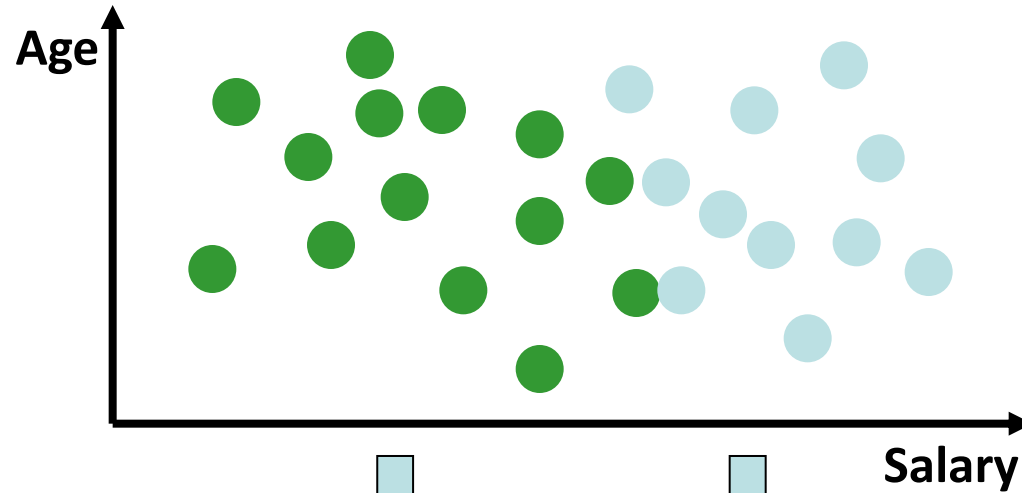
**How *does* Google find  
furry robots?**

**What *is* AI anyway?**

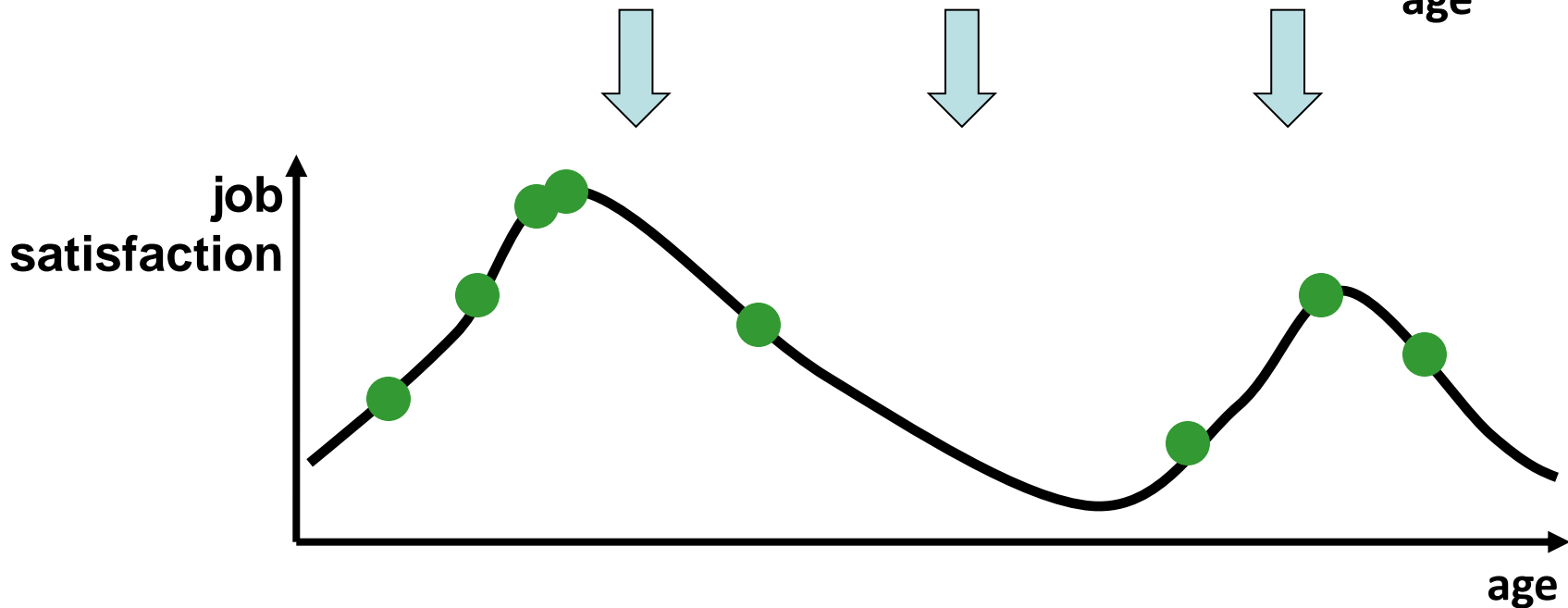
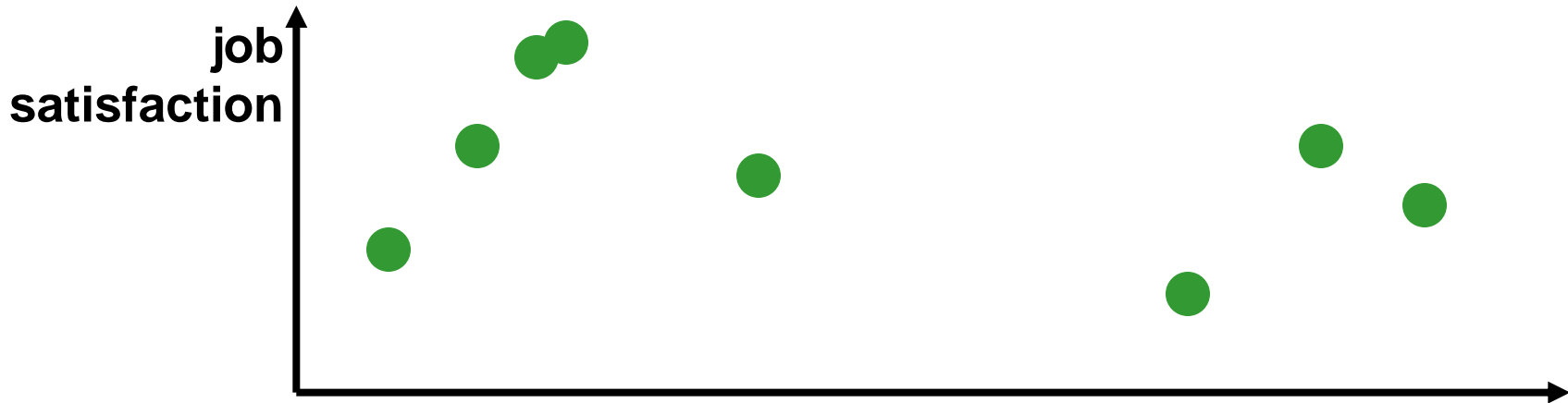
# Classification, in 2D

Credit profile:

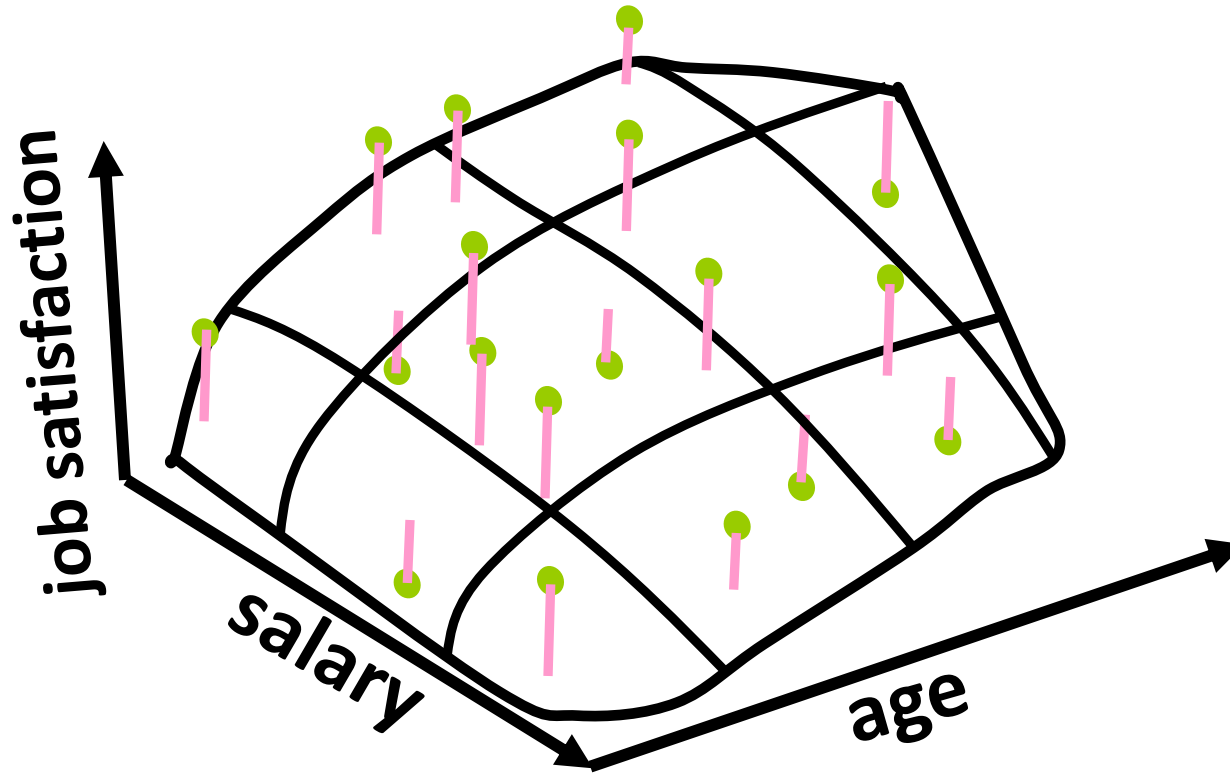
- Paid bills
- Didn't pay



# Regression, in 1D



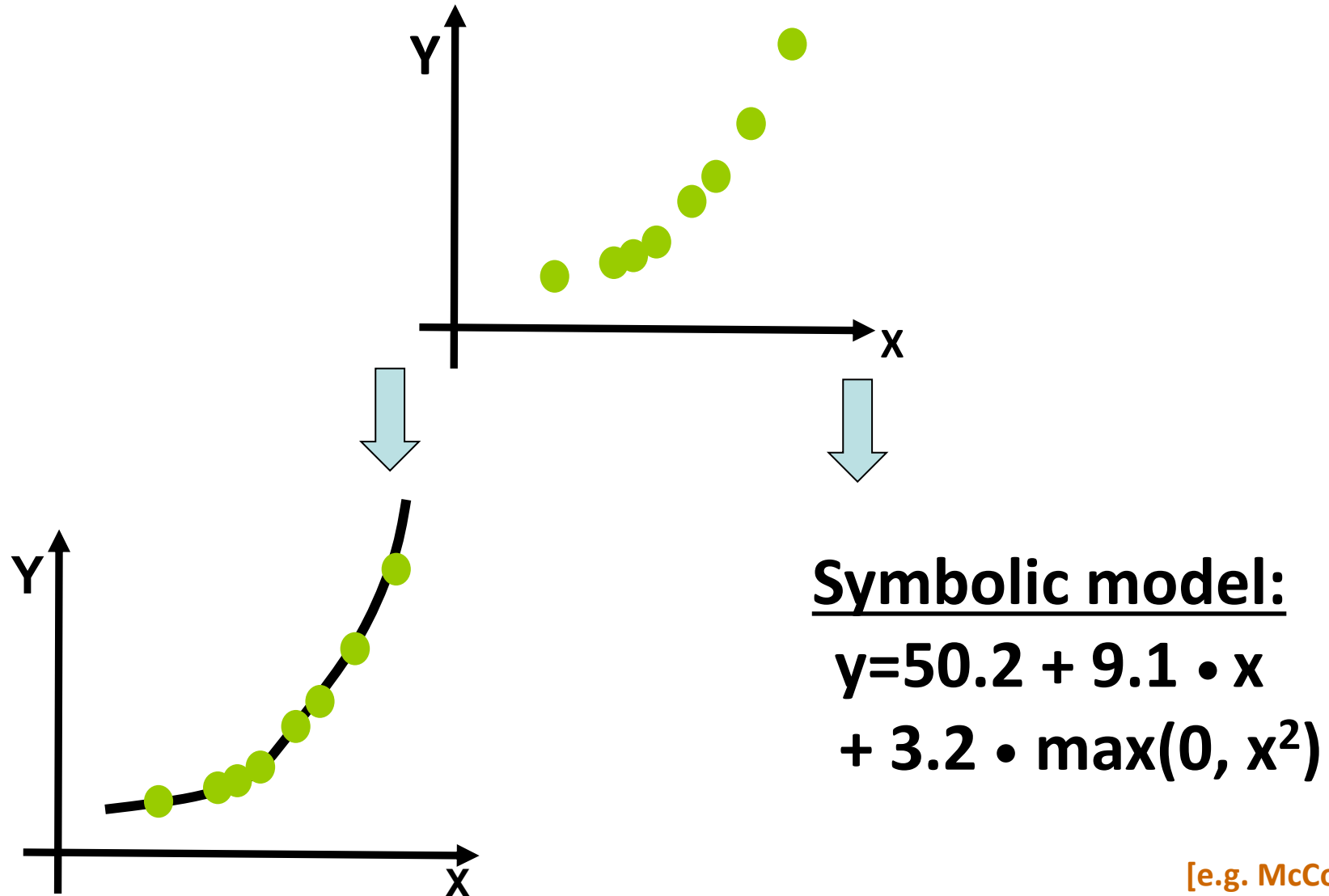
# Regression, in 2D



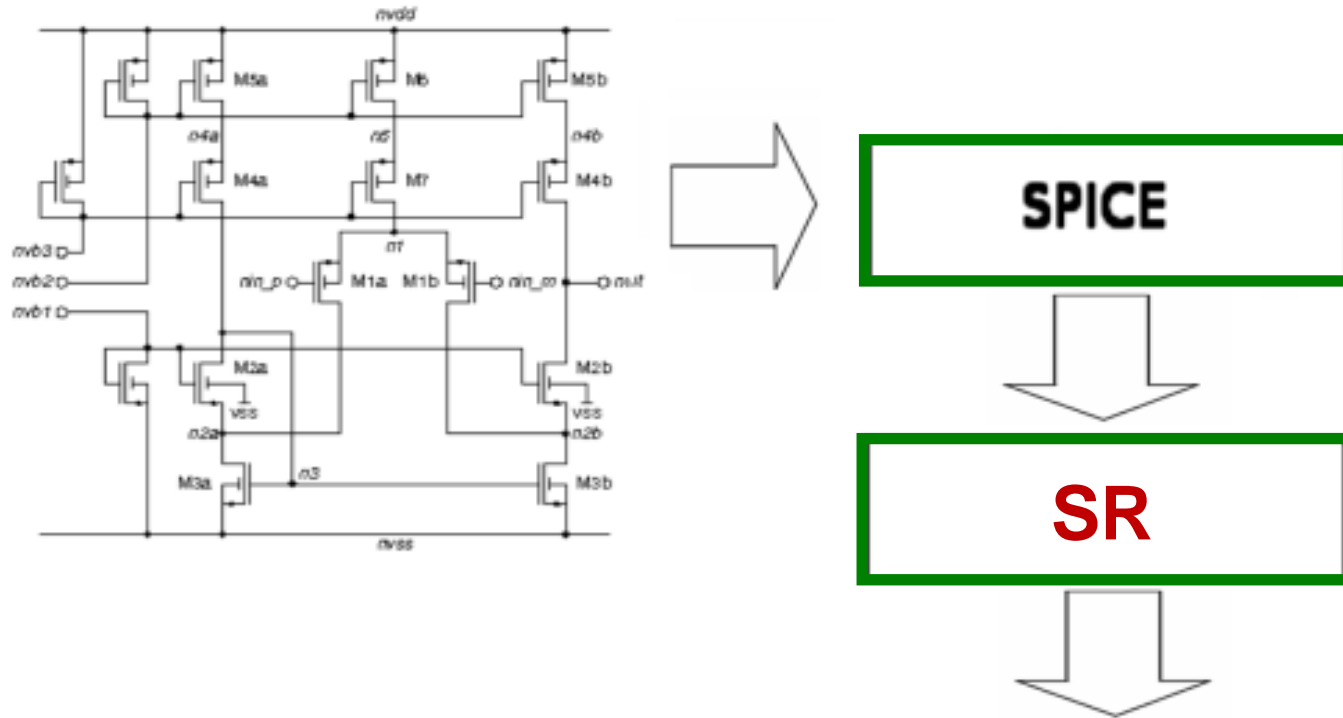
How: Polynomials, splines, neural networks, support vector machines, Gaussian process models, boosted trees, ... [\[many refs\]](#)

# Symbolic Regression (SR)

(Like regression, but output a symbolic model too)



# Example: SR on Circuits

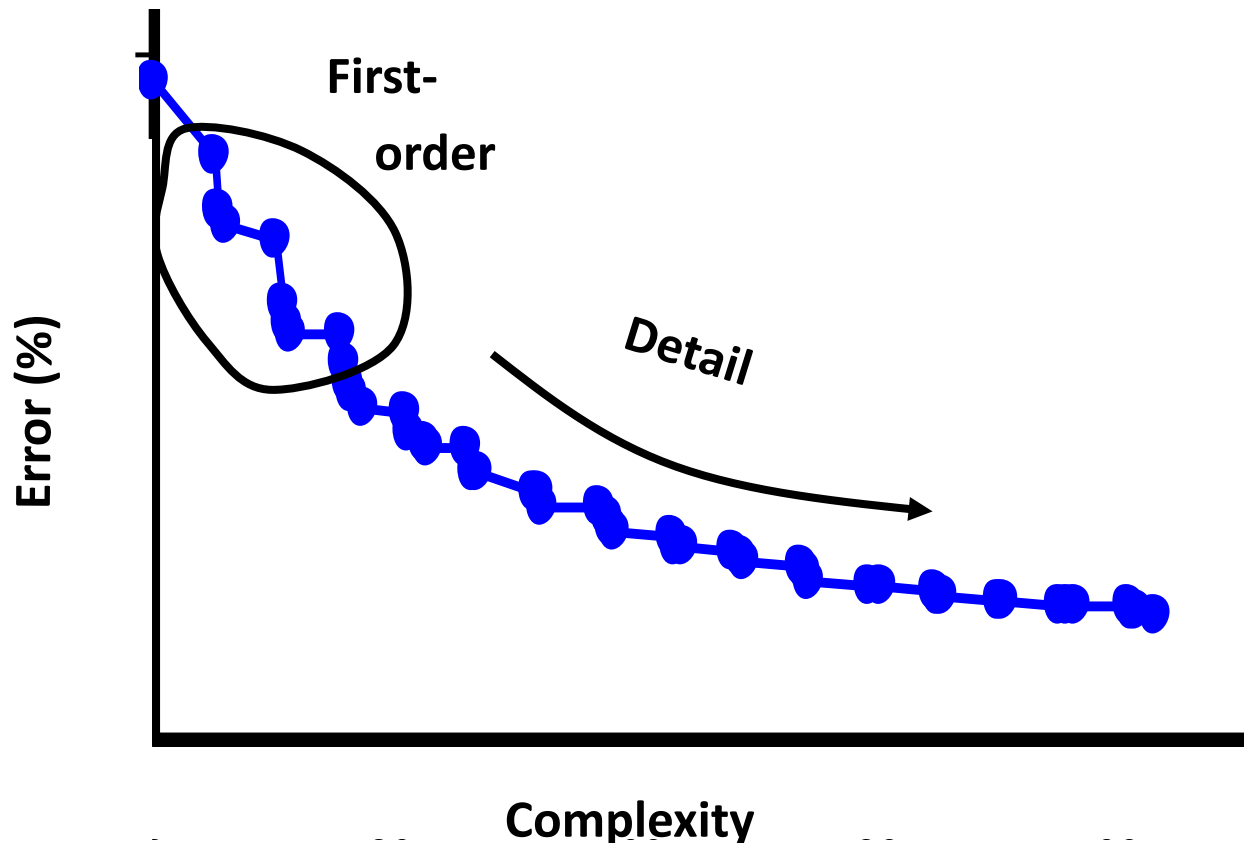


Perf.	Expression
$A_{LF}$	$-10.3 + 7.08e-5 / id1 + 1.87 * \ln(-1.95e+9 + 1.00e+10 / (vsg1*vsg3) + 1.42e+9 *(vds2*vds5) / (vsg1*vgs2*vsg5*id2))$
$f_u$	$10^{(5.68 - 0.03 * vsg1 / vds2 - 55.43 * id1 + 5.63e-6 / id1)}$
PM	$90.5 + 190.6 * id1 / vsg1 + 22.2 * id2 / vds2$
$V_{offset}$	$-2.00e-3$
$SR_p$	$2.36e+7 + 1.95e+4 * id2 / id1 - 104.69 / id2 + 2.15e+9 * id2 + 4.63e+8 * id1$
$SR_n$	$-5.72e+7 - 2.50e+11 * (id1*id2) / vgs2 + 5.53e+6 * vds2 / vgs2 + 109.72 / id1$



# SR Problem Definition, Redux

- Given  $(X, y)$
- Find whitebox *models*
- That minimize *error-complexity tradeoff*



# AI Has a Toolbox of Ways to Solve...

- Classification – Fraud detection, spam filtering ...
- Regression – Stock prediction, sensitivity analysis ...
- Whitebox regression – Scientific discovery ...
- Optimization – Airfoil design, circuit simulation ...
- Structural synthesis – Analog synthesis, robotics ...
- Pattern recognition – Face recognition, object recog ...
- System identification – Scientific discovery ...
- Ranking – Web search, ad serving, social discovery ...
- Control – Auto-driving autos, spacecraft trajectories ...
- ...

# AI Sub-fields

- machine learning
- neural networks
- evolutionary computation
- fuzzy logic
- data mining
- artificial general intelligence
- pattern recognition
- ..
- (nee) nonlinear programming
- (nee) databases
- ..

# AI Sub-fields of sub fields

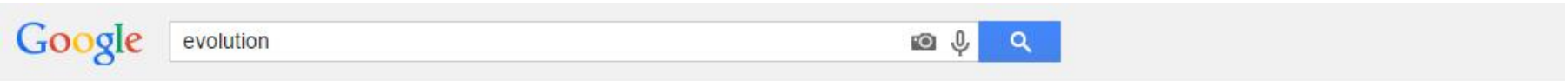
- machine learning + neural networks
  - recurrent neural networks
  - sparse linear regression
  - deep learning
  - ..
- evolutionary computation
  - evolutionary programming, evolution strategies
  - genetic algorithms
  - genetic programming
- ..

# **Genetic Programming (GP):**

A branch of a branch of AI

But a super-cool one..

# Evolution



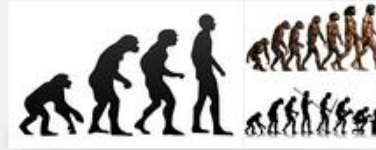
Web News **Images** Videos Maps More ▾ Search tools



Wwe  
**WTF?**



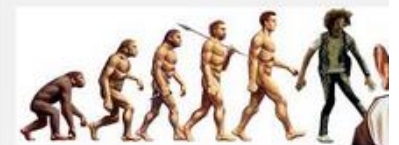
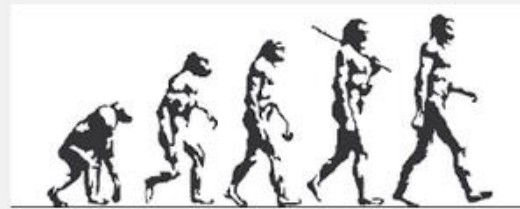
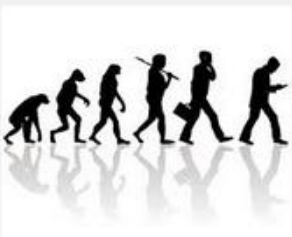
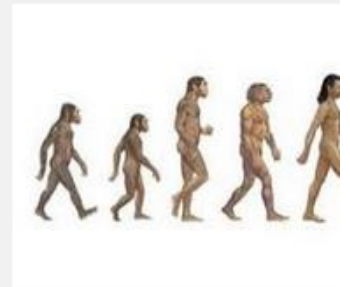
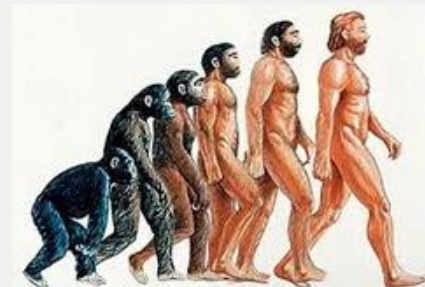
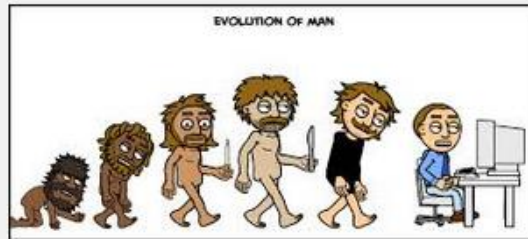
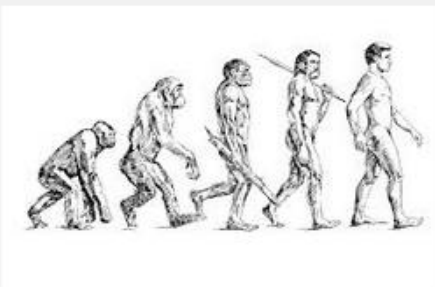
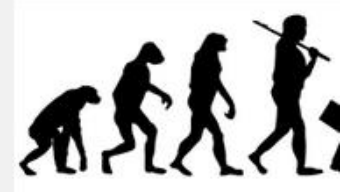
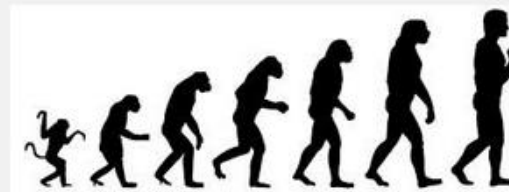
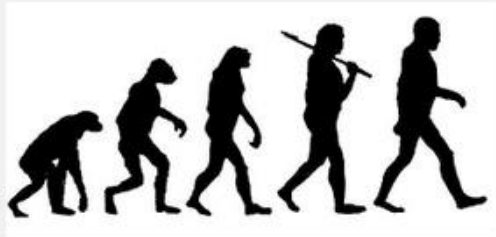
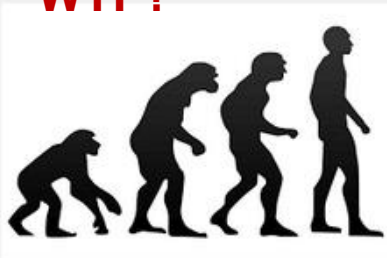
Animals



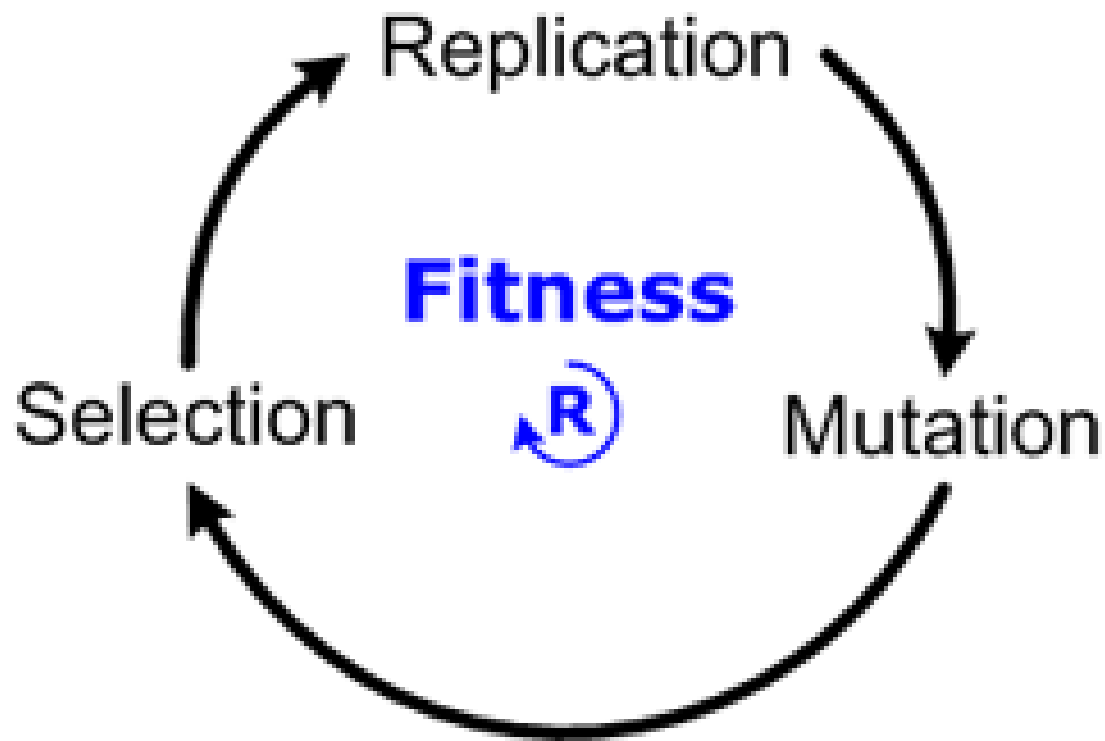
Human

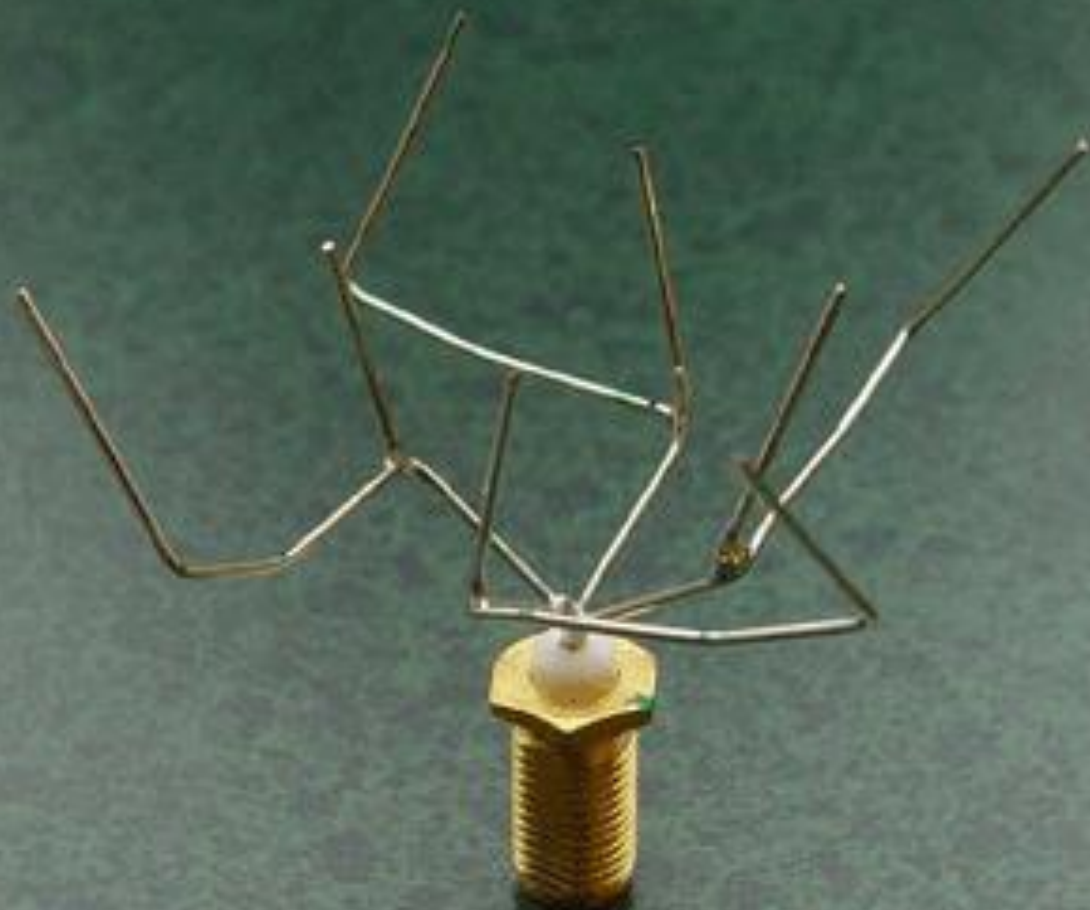


Man



# The Cycle of Evolution







# Genetic Programming

GENETIC PROGRAMMING  
by JOHN KOZL

HOW CAN WE GET A  
COMPUTER TO KNOW  
WHAT TO DO

AI  
50's  
Logic Based  
Representations  
SHIFT TO  
COFFS

DELIVER  
HIGH RETURN  
HUMAN COMPETITIVE  
INTELLIGENCE  
MACHINE

ROUTINE  
PATENTABLE  
LOGICAL DISCONTINUITY  
10:1 MOORE'S LAW JUMP

That  
will  
never  
work!

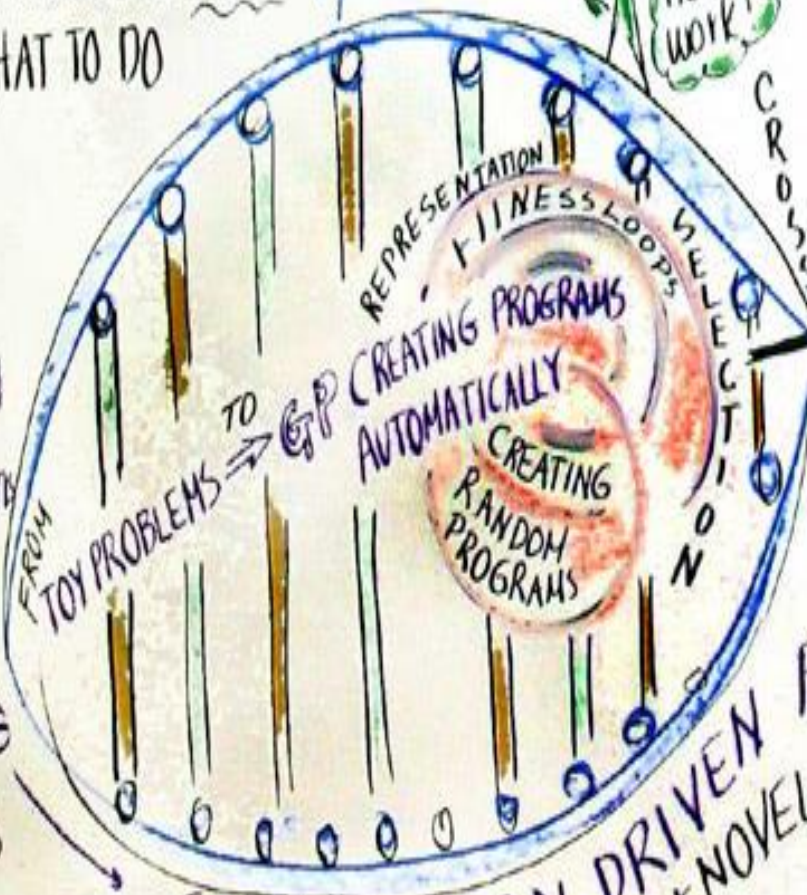
work  
along!

SCALABLE

TREE  
REPRESENTATION  
OF  
CIRCUITS

truss  
design  
G.P. AS A  
PATENT  
GENERATING  
MACHINE

LEADING EDGE  
15% of companies  
exploring +/or using  
-Brad Holtz

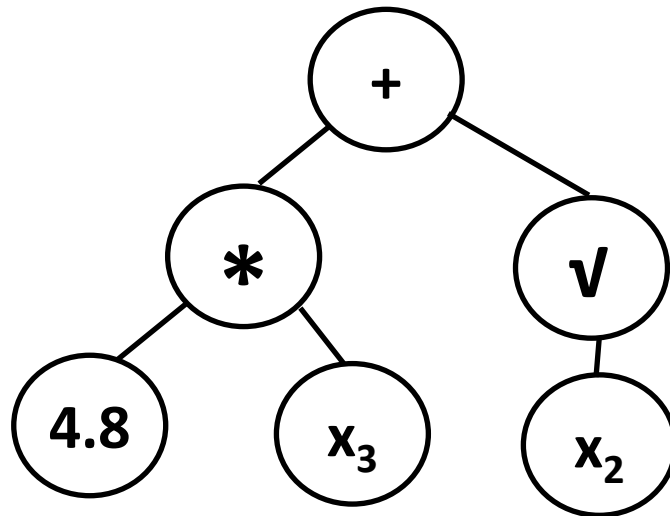


EVOLUTION DRIVEN BY HIGH LEVEL SPECS  
+ NOVELTY-DRIVEN  
AMENABLE TO  
BEOWULF

# GP for SR

“A function is a *tree*”

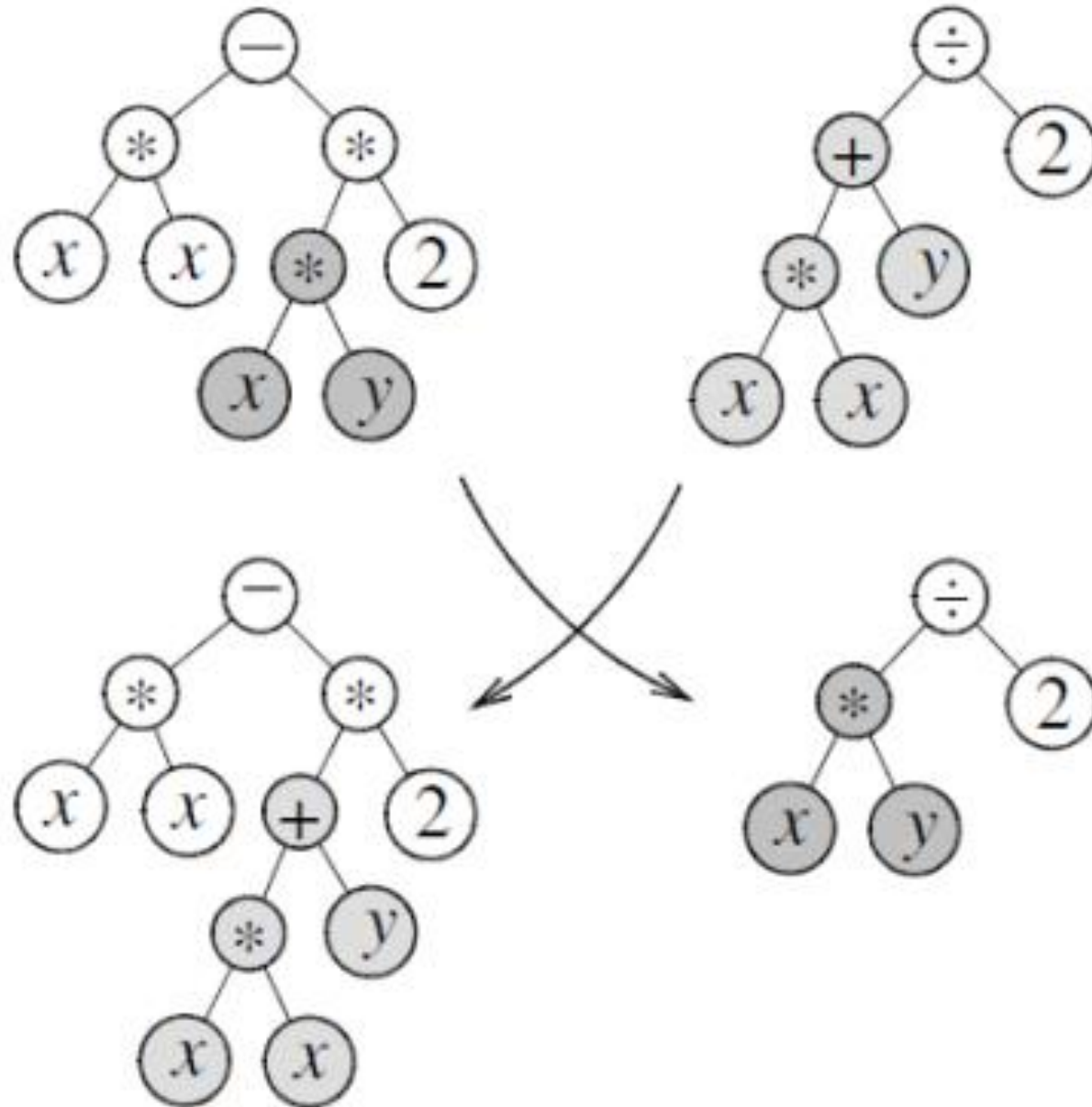
$$f(x) = 4.8 * x_3 + \sqrt{x_2}$$



Searches through the space of trees:

1. Initial random population; evaluate
2. Create children from parents via operators; evaluate
3. Select best; goto 2

# GP for SR: Crossover Operator





# SR with Vanilla GP.. And Problems

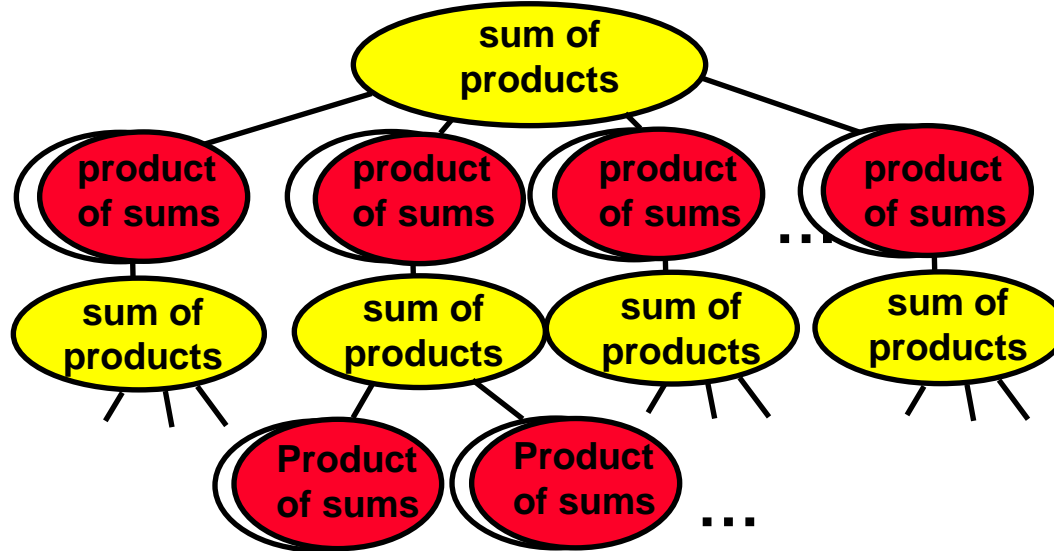
```
(+ (- (% (RLOG (COS X)) (* (RLOG 0.48800004)
                          (* (+ (- X X) (COS -0.8))
                             X)))
  (- (COS -0.8) (COS -0.8)))
(* (COS (- (COS (COS (+ (RLOG X)
                        (RLOG (COS X))))))
  (RLOG X)))
(* (COS (- (COS -0.8) (RLOG X)))
  (* (- (% (RLOG (COS X))
          (* (RLOG 0.48800004)
            (* (+ (- X X) (COS -0.8)) X)))
    (SIN X))
  (RLOG (COS (RLOG X)))))).
```

# SR with Vanilla GP.. And Problems (2)

```
(+ (* (* -0.5403 (+ 0.5741 -0.8861)) (% (*
0.296900000000000016 0.08089999999999997) (+ (% (% (-
-0.596200000000000001 0.39020000000000001) (- (+ (% (* (+ (*
0.2355000000000000004 0.150600000000000007) (* (*
-0.10289999999999999 -0.7332) 0.7723)) (*
0.2355000000000000004 0.150600000000000007)) (+ 0.6026 (+ (+
(% (- 0.372500000000000005 -0.34909999999999997) (- -0.776
-0.6013)) (- -0.52509999999999999 -0.009000000000000008))
(% (- 0.2969000000000000016 -0.34909999999999997) (- -0.776
-0.6013)))))) (* (+ -0.8861 (% -0.06019999999999992
0.05110000000000000145)) (% -0.06019999999999992
0.05110000000000000145)) (% -0.49659999999999993 0.4475)))
(+ (% (% (* (+ -0.19439999999999999 0.436600000000000001) (*
0.2355000000000000004 0.150600000000000007)) (+ 0.6026 (* (*
(+ (* -0.5403 -0.01719999999999993) (%
-0.06019999999999992 0.05110000000000000145)) (% (* (+
-0.19439999999999999 0.436600000000000001) (*
0.2355000000000000004 0.150600000000000007)) (% (%
0.421000000000000004 -0.4275) (- -0.481600000000000003
0.5708)))) 0.7723))) (- -0.8395 -0.1986)) (% (-
0.372500000000000005 -0.34909999999999997) (- -0.776
-0.6013)))) (% (% (+ 0.66980000000000002
0.871400000000000002) (% (- -0.829 -0.636) (-
0.763500000000000001 -0.15899999999999992))) (- (- (*
-0.5403 -0.01719999999999993) (- -0.8395 -0.1986)) (- (*
(* -0.5403 -0.01719999999999993) (- 0.6004 -0.4343)) (-
-0.951 (* (% 0.7803 0.9777) 0.319200000000000015))))))
(+ (* (* -0.5403 -0.01719999999999993)
-0.192400000000000002) (+ (+ -0.13339999999999996 0.7944)
0.6004))).
```

# CAFFEINE Approach

CAFFEINE = Canonical form functions in evolution



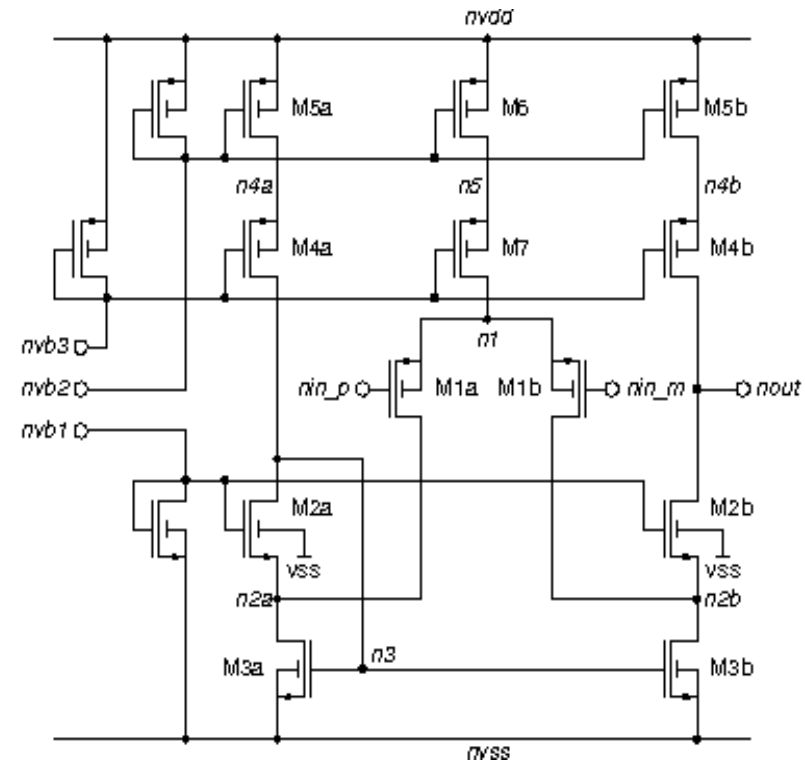
## Grammar to describe the canonical forms:

REPVC  $\Rightarrow$  VC | REPVC \* REPOP | REPOP  
 REPOP  $\Rightarrow$  REPOP \* REPOP | OP\_1ARG ( W +  
 REPADD ) | OP\_2ARG ( 2ARGS ) | ... 3OP, 4OP  
 2ARGS  $\Rightarrow$  W + REPADD, MAYBEW | MAYBEW,  
 W + REPADD  
 MAYBEW  $\Rightarrow$  W | W + REPADD  
 REPADD  $\Rightarrow$  W \* REPVC | REPADD + REPADD  
 OP\_2ARG  $\Rightarrow$  DIVIDE | POW | MAX | ...  
 OP\_1ARG  $\Rightarrow$  INV | LOG10 | ...

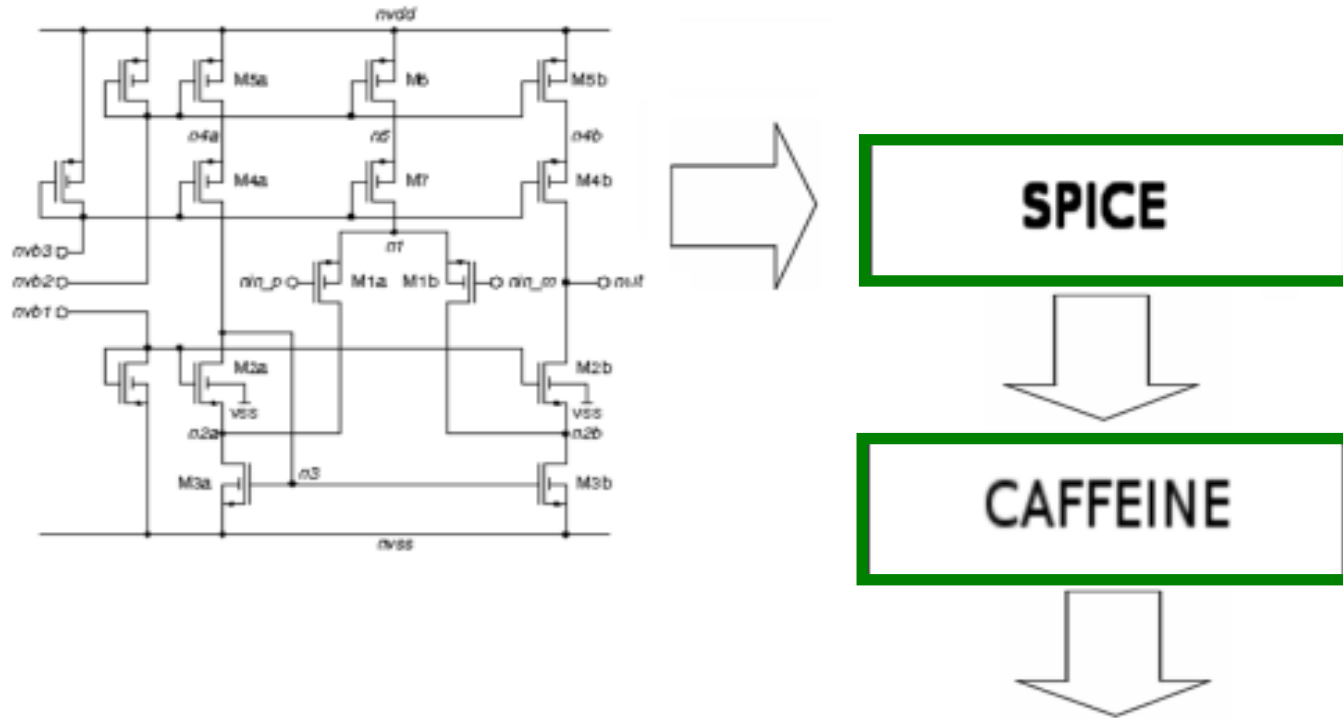
## Search the space with *grammatically-constrained* GP [Whig1995]

# Experimental Setup

- High Speed amplifier
- 13 design variables
  - $V_{ds}$ ,  $V_{gs}$ ,  $I_{ds}$  (operating-point driven formulation)
- orthogonal hypercube sampling
- 243 training samples
- 243 testing samples



# Example: GP for SR on Circuits



Perf.	Expression
$A_{LF}$	$-10.3 + 7.08e-5 / id1 + 1.87 * \ln(-1.95e+9 + 1.00e+10 / (vsg1*vsg3) + 1.42e+9 *(vds2*vds5) / (vsg1*vgs2*vsg5*id2))$
$f_u$	$10^{(5.68 - 0.03 * vsg1 / vds2 - 55.43 * id1 + 5.63e-6 / id1)}$
PM	$90.5 + 190.6 * id1 / vsg1 + 22.2 * id2 / vds2$
$V_{offset}$	$-2.00e-3$
$SR_p$	$2.36e+7 + 1.95e+4 * id2 / id1 - 104.69 / id2 + 2.15e+9 * id2 + 4.63e+8 * id1$
$SR_n$	$-5.72e+7 - 2.50e+11 * (id1*id2) / vgs2 + 5.53e+6 * vds2 / vgs2 + 109.72 / id1$

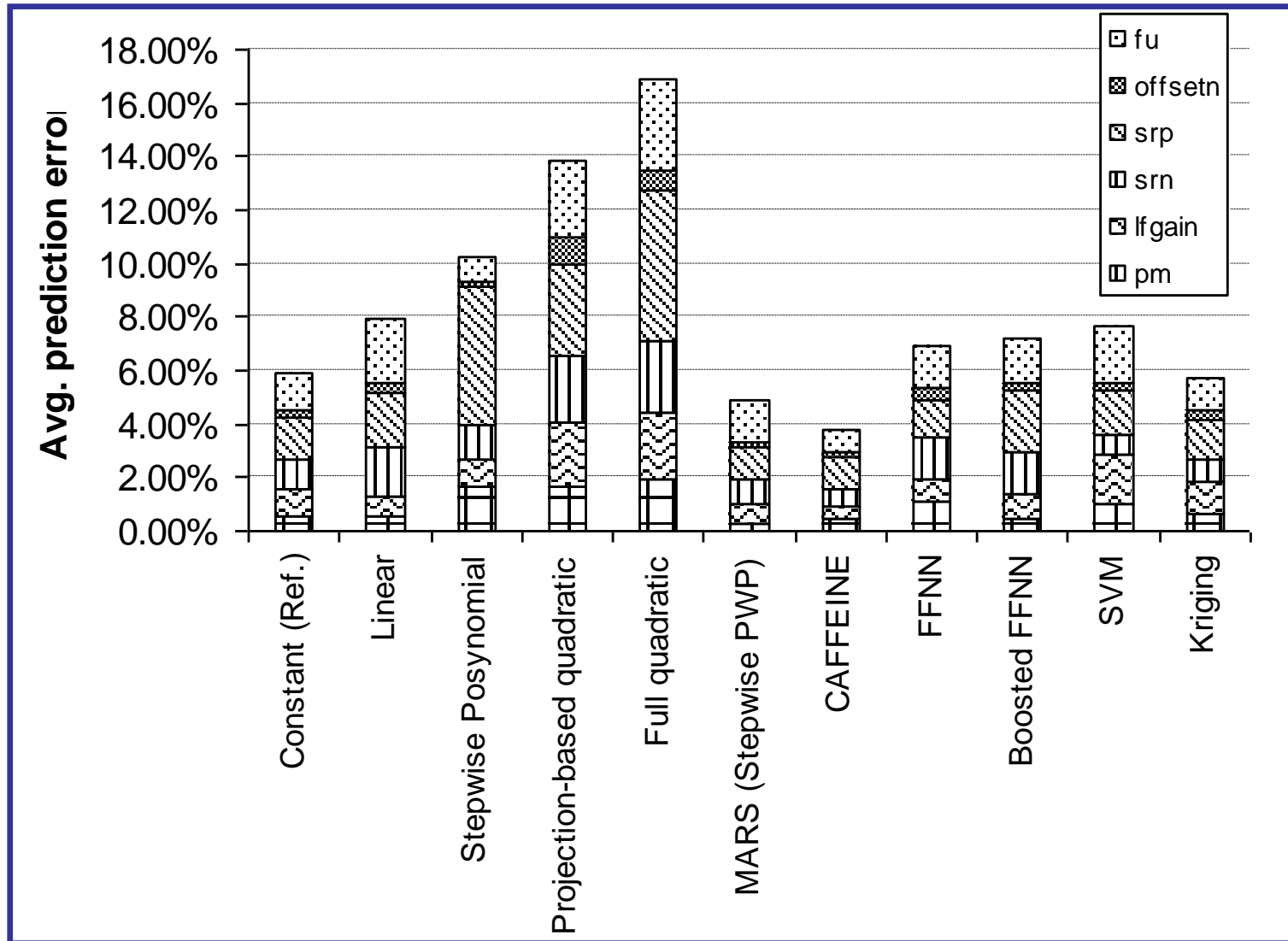


# CAFFEINE models with <10% error

Perf.	Target % error		Expression
	tr n	tst	
$A_{LF}$	10	10	$-10.3 + 7.08e-5 / id1$ $+ 1.87 * \ln( -1.95e+9 + 1.00e+10 / (vsg1*vsg3)$ $+ 1.42e+9 *(vds2*vds5) /$ $(vsg1*vgs2*vsg5*id2) )$
$f_u$	10	10	$10^{( 5.68 - 0.03 * vsg1 / vds2 - 55.43 * id1 + 5.63e-6 / id1 )}$
PM	10	10	$90.5 + 190.6 * id1 / vsg1 + 22.2 * id2 / vds2$
$V_{offset}$	10	10	$- 2.00e-3$
$SR_p$	10	10	$2.36e+7 + 1.95e+4 * id2 / id1 - 104.69 / id2 + 2.15e+9 * id2$ $+ 4.63e+8 * id1$
$SR_n$	10	10	$- 5.72e+7 - 2.50e+11 * (id1*id2) / vgs2 + 5.53e+6 * vds2 /$ $vgs2$

# CAFFEINE Prediction Performance

Predicts better than several state-of-the-art blackbox regression techniques on circuits benchmark suite (*and* gives whitebox models).



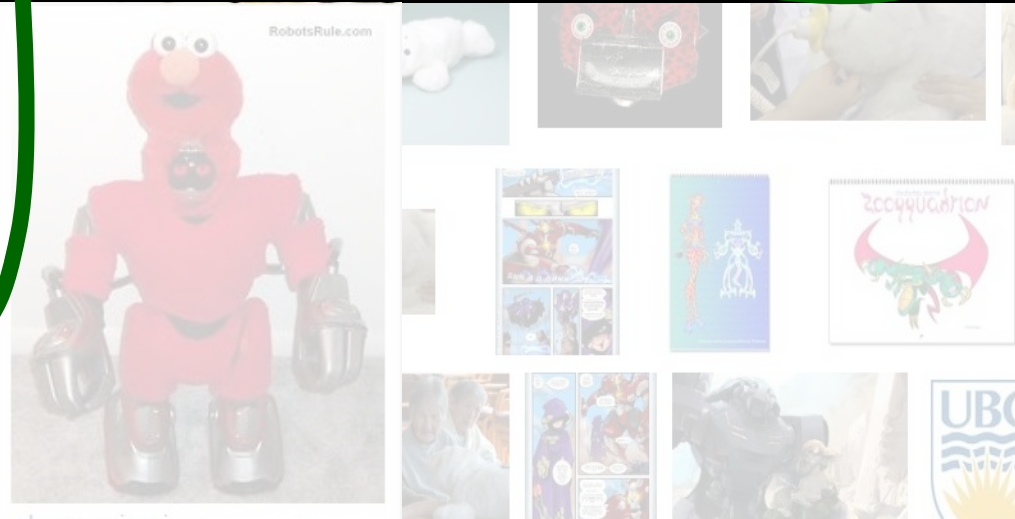
# Conclusion

**What does AI encompass?**



**WTF is genetic programming or symbolic regression?  
Why should I care?**

Is Deep Learning cool or what?



How *does* Google find furry robots?