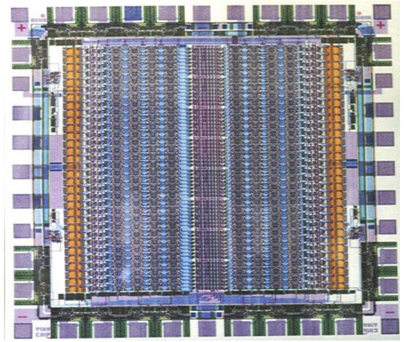
Machine Intelligence

J. T. Wunderlich, Ph.D.

Last updated in 2017

AGENDA

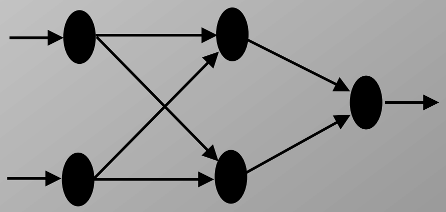
- Neural Networks vs Symbolic AI *(basics)*
  - *Wunderlich educational examples, 1990-present*
- Updated Wunderlich Mental Ability Matrix
  - *Philosophy, Psychology*
  - *Wunderlich publications 2003 -present*
- Design of two Neurocomputers
  - *Wunderlich 1991 and 1992*



- Case Study: IBM Watson
  - *Wunderlich related IBM Research in mid 1990's*
- Machine Intelligence in Robots
  - *Recent Wunderlich publications*
  - *Watch the beginning of this Oxford University Video now:*
    - *https://www.youtube.com/watch?v=rXVoRyIGGhU*

Two major Machine Intelligence fields of study

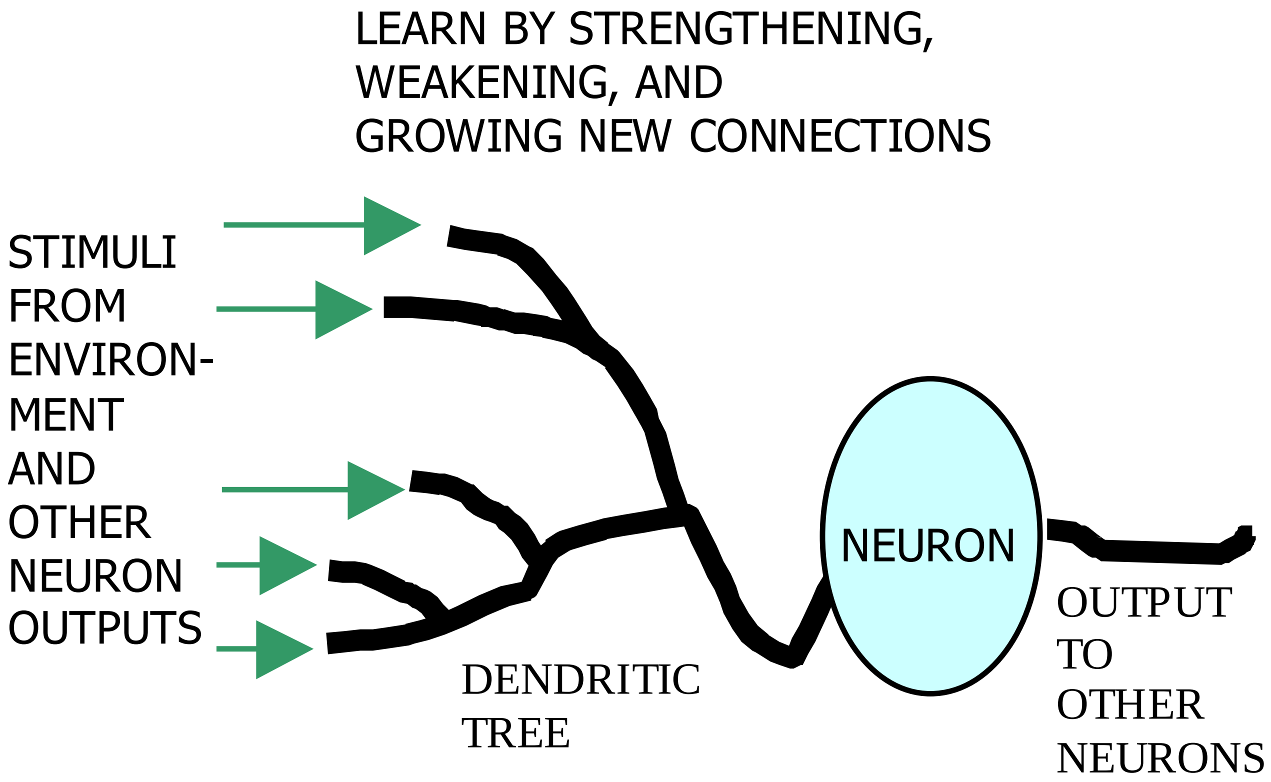
- Artificial **Neural Networks**
  - Connectionist architecture (hardware or software) *similar* to biological brain
  - Trained (learns) by changing strength of inter-neuron connections
  - Satisfy multiple constraints



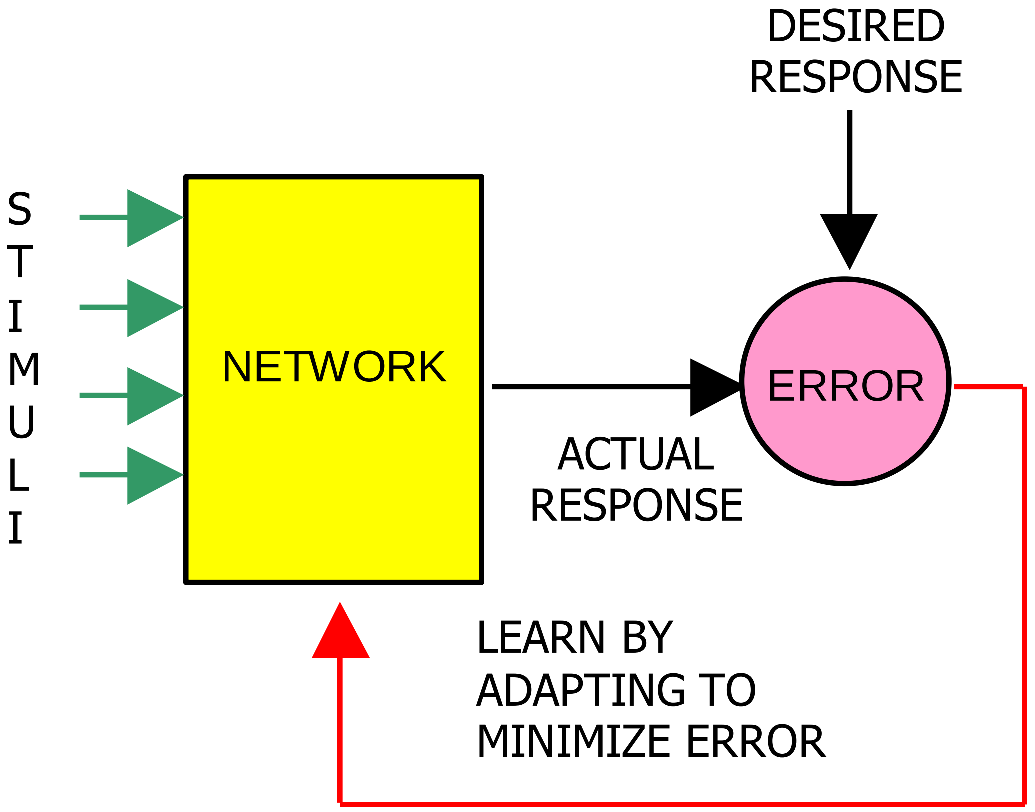
- **Symbolic AI** programs
  - Heuristics, inference, hypothesis-testing, and forms of knowledge representation
  - "Expert Systems"
  - Predicate Calculus, PROLOG, LISP
  - **Confidence Factors ("Values"), Probability Theory**

Neural Network

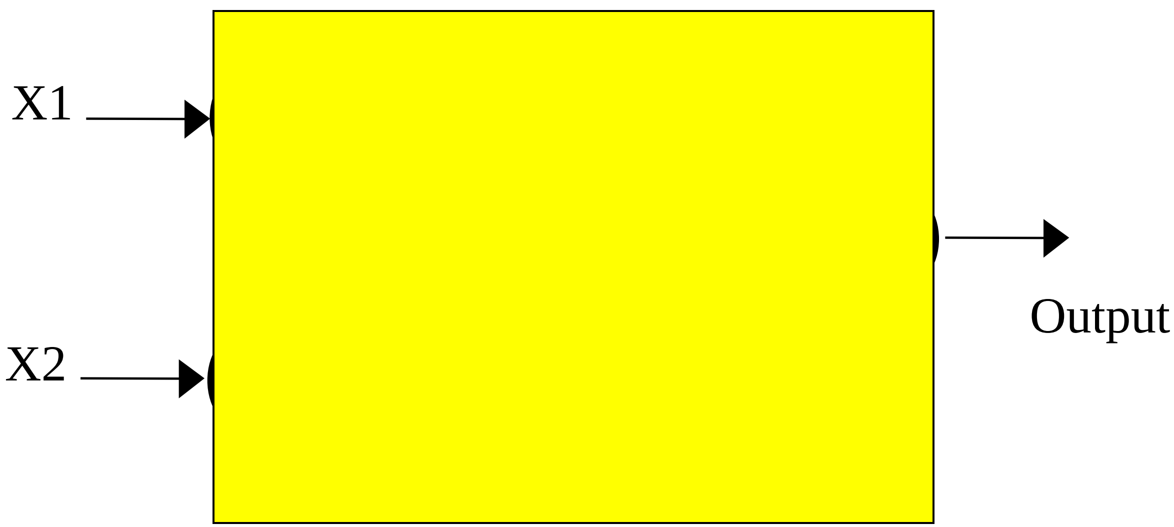
Biological neuron for "bottom-up" designs



Behavioral model for "top-down" designs

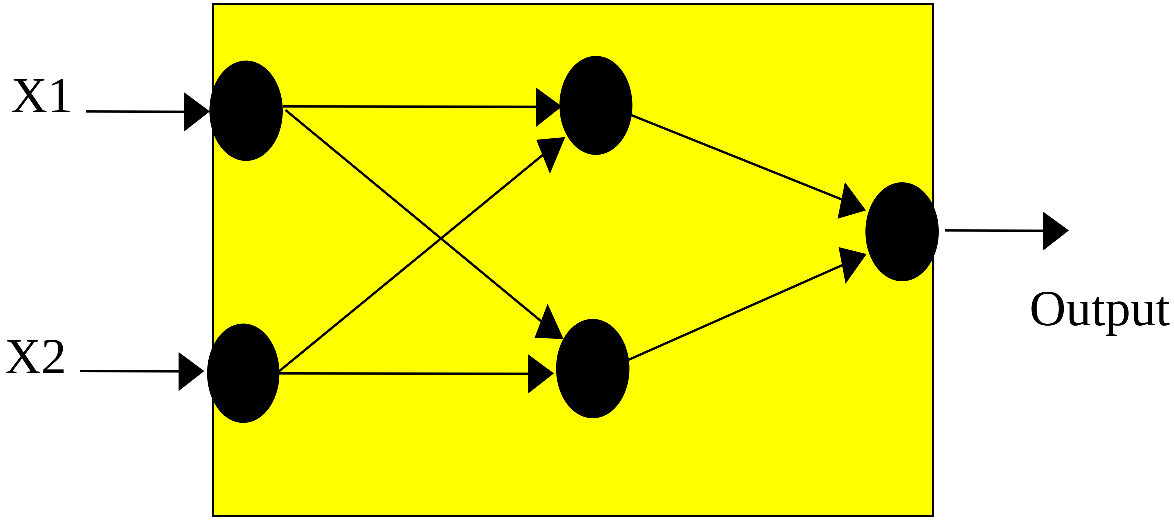


Example Neural Network – *Behavioral Model (“back propagation”)*



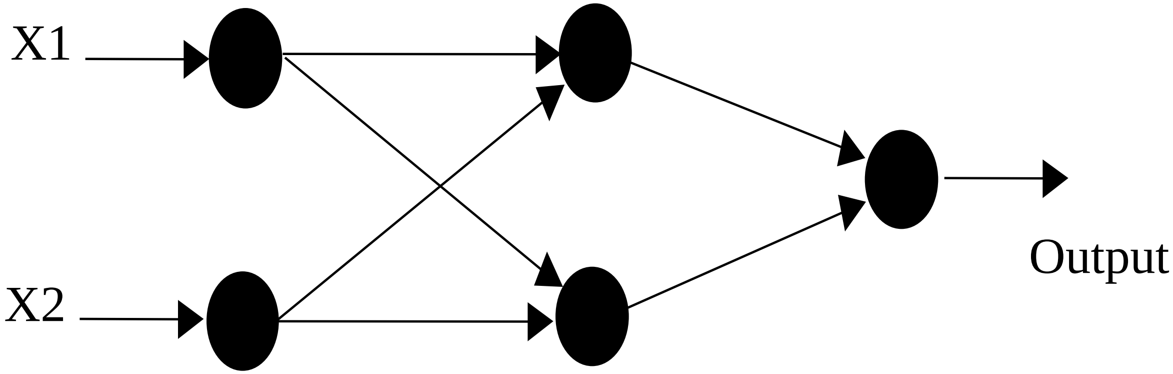
X1	X2	Output
0	0	**0**
0	1	**1**
1	0	**1**
1	1	**0**

Example Neural Network – *Behavioral Model ("back propagation")*



X1	X2	Output
0	0	**0**
0	1	**1**
1	0	**1**
1	1	**0**

Example Neural Network – *Behavioral Model (“back propagation”)*



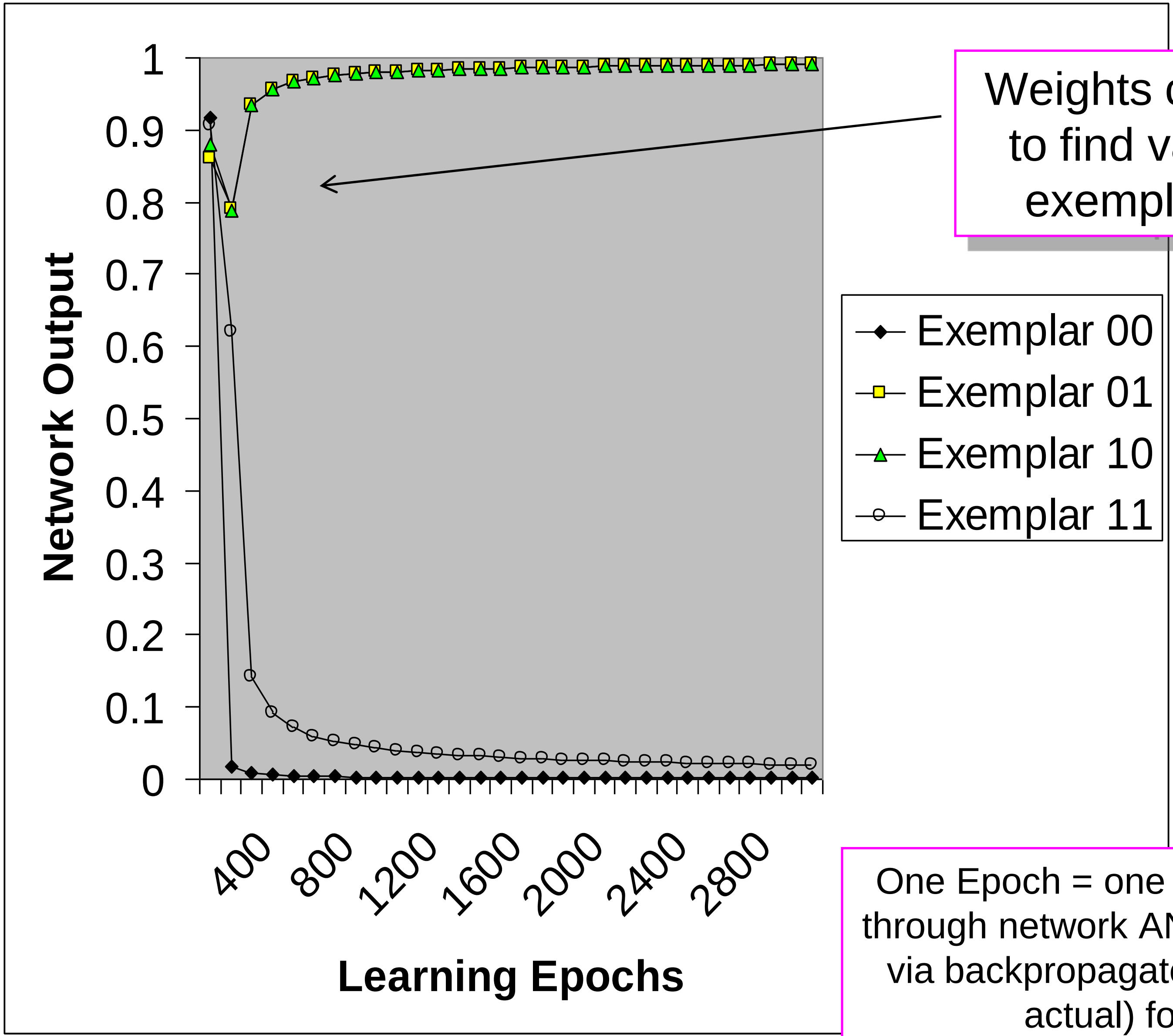
X1	X2	Output
0	0	**0**
0	1	**1**
1	0	**1**
1	1	**0**

LEARN by changing “weights” of connections between neurons:

$$\Delta W_{jk} = \eta * \left[(d_k - O_k) * O_k * (1 - O_k) \right] * O_j$$

$$\Delta W_{ij} = \eta * \left(O_j * (1 - O_j)\right) * \sum_k \left[(d_k - O_k) * O_k * (1 - O_k) * W_{jk} \right] * O_i$$

Example Neural Network – *Behavioral Model (“back propagation”)*

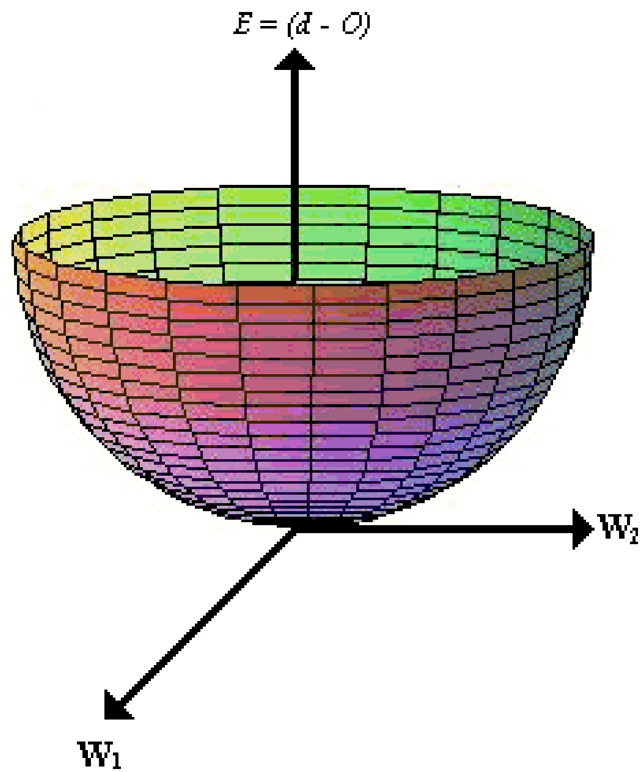


Weights changing *(struggling)* to find values that satisfy all exemplars simultaneously

One Epoch = one forward feed of all exmplars through network AND one change of all weights via backpropagated error (= desired output – actual) for each of exemplar

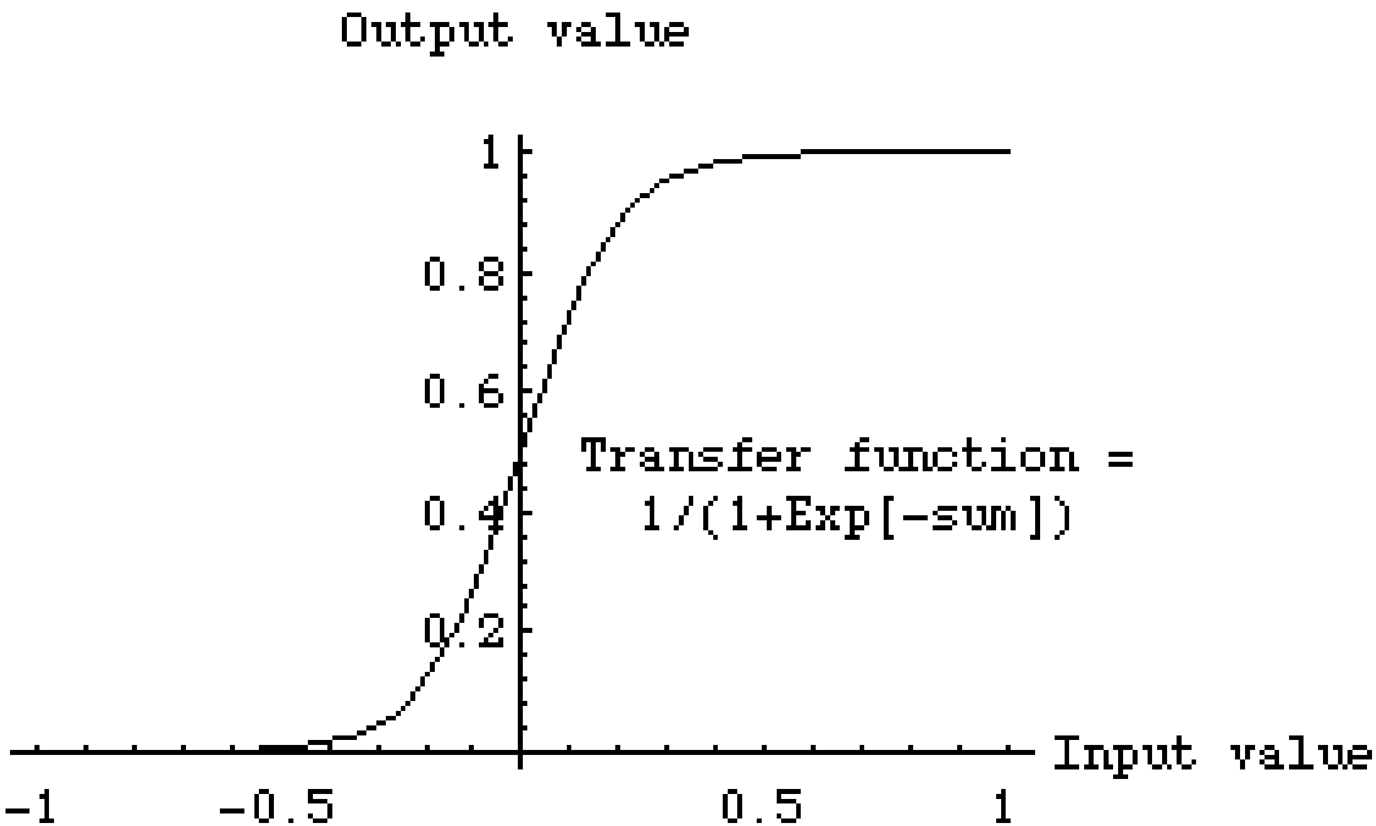
Neural Network LEARNING for this type Neural Net is gradient decent minimization on an Error Surface

where *"d"*= Desired Output" and *"O"* = present output



VIDEO by others: https://www.youtube.com/watch?v=WZDMNM36PsM

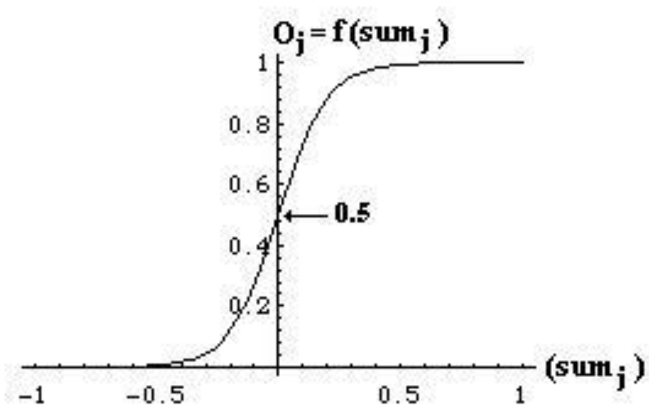
… where the neuron transfer function needs to be continuously differentiable

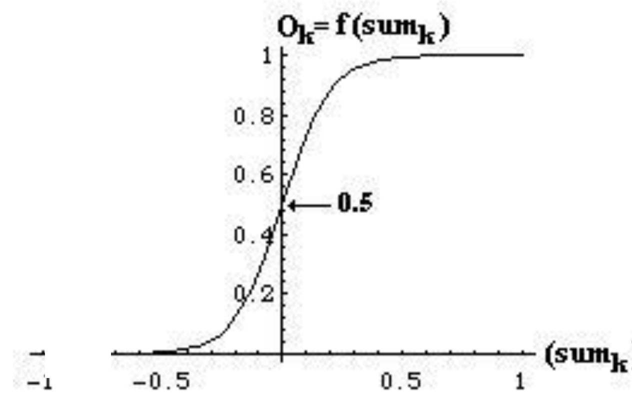


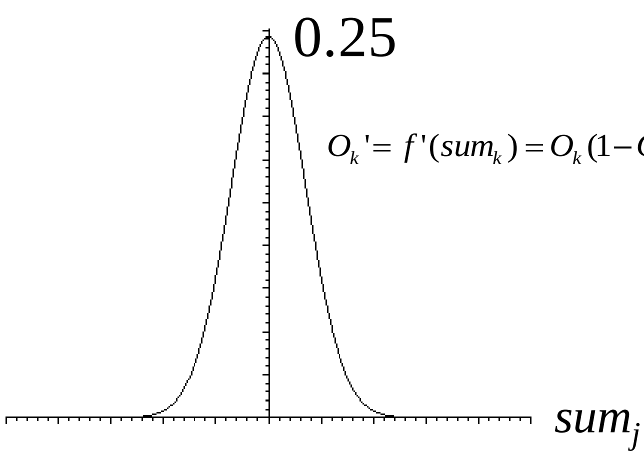
. . . and learning is most noticeable *(and mathematically magnified)* when neuron input summation is very small

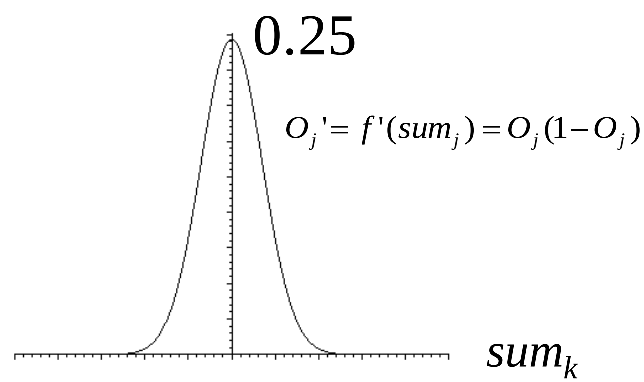
*(where **j** is the middle "hidden" layer and **k** is the output layer)*

i.e., It's in a ***"DECIDING position"***









Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's)

J. Wunderlich, 1991

AI Program to help pick a toy for a child

NOTE: Unlike Probability Theory, Confidence Factors do not need to add up to 100

User input:

1. Child's Age?
   - Program automatically assigns a CNF Confidence factor = 100
2. Do you want to spend more or less than $25?
   - Input CNF Confidence factor (e.g. CNF=65 for less than $25)
3. Child's gender?
   - Program automatically assigns a CNF Confidence factor = 100
4. Preference for type of toy? (Action, Cuddly, or Creative)
   - Input CNF Confidence factor s for each

Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's)
J. Wunderlich, 1991

ASSUMPTIONS:

1) Gender is not relevant for children younger than the age of one.
2) Children younger than the age of one cannot have a known preference for toys.
3) Children between the ages of one and three cannot have a known preference for toys.

CONTROL OF SEARCH SPACE:

As a result of the above assumptions, the rules have been ordered so that:

A) child_age is the first variable in the premise of every rule.
B) gender and child_preference questions are not asked when child_age = under_1
C) The child_preference question is not asked when child_age = one_to_three

TESTING OF EXPERT SYSTEM:

The expert system was tested under the three scenarios listed in the table below.
The resulting output is shown in the last column with the calculated confidence factors.
These confidence factors were calculated by VP-EXPERT using the following standard laws of Certainty:

A) CNF(P1 and P2) = MIN(CNF(P1), CNF(P2))
B) CNF(P1 or P2) = MAX(CNF(P1), CNF(P2))

And when two or more rules support the same result R:

C) CNF(R1) + CNF(R2) - (CNF(R1) * CNF(R2)) when CNF(R1) and CNF(R2) are positive
D) CNF(R1) + CNF(R2) + (CNF(R1) * CNF(R2)) when CNF(R1) and CNF(R2) are negative
E) (CNF(R1) + CNF(R2))/(1-MIN(|CNF(R1)|,|CNF(R2)|) otherwise

Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's)

J. Wunderlich, 1991

The example below shows the CNF calculations for suggestedd_toy = dress_up_doll

For the premise of RULE 12:

[(child_age=four_to_six) AND (price=under_25) AND (gender=female) AND (child_preference=cuddly_toy)]
CNF=1 AND CNF=0.65 CNF=1 AND CNF=0.55

Using law #A above; CFN(premise) = MIN(1,0.65,1,0.55) = 0.55

Using law #C above; [CFN(conclusion) = CFN(premise) * CFN(RULE 12)] = [0.55 * 0.9] = 0.49 see NOTE 7

Laws #C,#D, or #E are not used here because only RULE 12 fires to support the goal (suggested_toy=dress_up_doll)

TEST TRACE #	INPUT: child_age (see NOTE 1)	INPUT: price (see NOTE 2)	INPUT: gender (seeNOTE 3)	INPUT: chid_preference (see NOTE 4)	RULE FIRED AND ITS CNF (see NOTE 7)	OUTPUT: suggested_toy
1	under_1	under_25 (CNF=65)	N.A.	N.A.	CNF(R1)=95	teething_toy (CNF=61)
		over_25 (CNF=20)			CNF(R2)=90	mobile_for_crib (CNF=58)
					CNF(R3)=95	plastic_rattle (CNF=61)
					CNF(R4)=75	sterling_silver_rattle (CNF=15)
2	one_to_three	under_25 (CNF=65)	male	N.A.	CNF(R5)=90	roly_poly (CNF=58)
		over_25 (CNF=20)			CNF(R6)=90	tricycle (CNF=18)
					CNF(R7)=85	hammer_and_pegs_game (CNF=55)
3	four_to_six	under_25 (CNF=65)	female	action_toys (CNF=25)	CNF(R9)=95	lincoln_logs (CNF=61)
	CNF=100	over_25 (CNF=20)	CNF=100	cuddly_toys (CNF=55)	CNF(R11)=95	doll_house (CNF=18)
				creative_toys (CNF=75)	CNF(R12)=90	dress_up_doll (CNF=49)
					CNF(R14)=85	toy_tea_set (CNF=55)

Note 7: CFN's for RULES were assigned by J. Wunderlich Typically these will be assigned by the "Knowledge Engineer" after consultation with the "Domain Expert"

Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's) *J. Wunderlich, 1991*

CFN's for RULES were assigned by J. Wunderlich Typically these will be assigned by the "Knowledge Engineer" after consultation with the "Domain Expert"

```
RULE 1
IF child_age = under_one AND
price = under_25
THEN suggested_toy = teething_toy CNF 95;

RULE 2
IF child_age = under_one AND
price = under_25
THEN suggested_toy = mobile_for_crib CNF 90;

RULE 3
IF child_age = under_one AND
price = under_25
THEN suggested_toy = plastic_rattle CNF 95;

RULE 4
IF child_age = under_one AND
price = over_25
THEN suggested_toy = sterling_silver_rattle CNF 75;

RULE 5
IF child_age = one_to_three AND
price = under_25
THEN suggested_toy = roly_poly CNF 90;

RULE 6
IF child_age = one_to_three AND
price = over_25
THEN suggested_toy = tricycle CNF 90;

RULE 7
IF child_age = one_to_three AND
price = under_25 AND
gender = male
THEN suggested_toy = hammer_and_pegs_game CNF 85;
```

Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's) *J. Wunderlich, 1991*

CFN's for RULES were assigned by J. Wunderlich Typically these will be assigned by the "Knowledge Engineer" after consultation with the "Domain Expert"

```
RULE 9
IF child_age = four_to_six AND
price = under_25 AND
child_preference = creative_toys
THEN suggested_toy = lincoln_logs CNF 95;

RULE 10
IF child_age = four_to_six AND
price = over_25 AND
gender =  male AND
child_preference = action_toys
THEN suggested_toy = go_cart CNF 85;

RULE 11
IF child_age = four_to_six AND
price = over_25 AND
gender = female AND
child_preference = creative_toys
THEN suggested_toy = doll_house CNF 90;

RULE 12
IF child_age = four_to_six AND
price = under_25 AND
gender = female AND
child_preference = cuddly_toys
THEN suggested_toy = dress_up_doll CNF 90;

RULE 13
IF child_age = four_to_six AND
price = over_25 AND
gender = male AND
child_preference = action_toys
THEN suggested_toy = hot_wheels_set CNF 95;

RULE 14
IF child_age = four_to_six AND
price = under_25 AND
gender = female AND
child_preference = creative_toys
THEN suggested_toy = toy_tea_set CNF 85;

RULE 15
IF child_age = four_to_six AND
price = under_25 AND
gender = male AND
child_preference = creative_toys
THEN suggested_toy = army_men CNF 90;
```

Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's)

J. Wunderlich, 1991

TEST TRACE #	INPUT: child_age (see NOTE 1)	INPUT: price (see NOTE 2)	INPUT: gender (seeNOTE 3)	INPUT: chid_preference (see NOTE 4)	RULE FIRED AND ITS CNF (see NOTE 7)	OUTPUT: suggested_toy
1	under_1	under_25 (CNF=65)	N.A.	N.A.	CNF(R1)=95	teething_toy (CNF=61)
		over_25 (CNF=20)			CNF(R2)=90	mobile_for_crib (CNF=58)
					CNF(R3)=95	plastic_rattle (CNF=61)
					CNF(R4)=75	sterling_silver_rattle (CNF

Test-trace #1 →

```
suggested_toy
!   Testing 1
!   !   child_age
!   !   !   (= under_one CNF 100 )
!   !   price
!   !   !   (= under_25 CNF 65 )
!   !   !   (= over_25 CNF 20 )
!   (= teething_toy CNF 61 )
!   Testing 2
!   (= mobile_for_crib CNF 58 )
!   Testing 3
!   (= plastic_rattle CNF 61 )
!   Testing 4
!   (= sterling_silver_rattle CNF 15 )
!   Testing 5
!   Testing 6
!   Testing 7
!   Testing 8
!   Testing 9
!   Testing 10
!   Testing 11
!   Testing 12
!   Testing 13
!   Testing 14
!   Testing 15
```

Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's)

J. Wunderlich, 1991

TEST TRACE #	INPUT: child_age (see NOTE 1)	INPUT: price (see NOTE 2)	INPUT: gender (seeNOTE 3)	INPUT: chid_preference (see NOTE 4)	RULE FIRED AND ITS CNF (see NOTE 7)	OUTPUT: suggested_toy
2	one_to_three	under_25 (CNF=65)	male	N.A.	CNF(R5)=90	roly_poly (CNF=58)
		over_25 (CNF=20)			CNF(R6)=90	tricycle (CNF=18)
					CNF(R7)=85	hammer_and_pegs_game (CNF=5

Test-trace #2 →

```
gested_toy
 Testing 1
 !   child_age
 !   !   (= one_to_three CNF 100 )
 Testing 2
 Testing 3
 Testing 4
 Testing 5
 !   price
 !   !   (= under_25 CNF 65 )
 !   !   (= over_25 CNF 20 )
 (= roly_poly CNF 58 )
 Testing 6
!    (= tricycle CNF 18 )
!    Testing 7
!    !   gender
!    !   !   (= male CNF 100 )
!    (= hammer_and_pegs_game CNF 55 )
!    Testing 8
!    Testing 9
!    Testing 10
!    Testing 11
!    Testing 12
!    Testing 13
!    Testing 14
!    Testing 15
```

Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's)

J. Wunderlich, 1991

TEST TRACE	INPUT				RULE FIRED AND ITS CNF	OUTPUT
#	child_age (see NOTE 1)	price (see NOTE 2)	gender (seeNOTE 3)	chid_preference (see NOTE 4)	(see NOTE 7)	suggested_toy
3	four_to_six	under_25 (CNF=65)	female	action_toys (CNF=25)	CNF(R9)=95	lincoln_logs (CNF=61)
	CNF=100	over_25 (CNF=20)	CNF=100	cuddly_toys (CNF=55)	CNF(R11)=95	doll_house (CNF=18)
				creative_toys (CNF=75)	CNF(R12)=90	dress_up_doll (CNF=49)
					CNF(R14)=85	toy_tea set (CNF=55)

MODIFIED Test-trace #3

with age changed to one_to_three, and

gender changed to male

This changed OUTPUT to:

roly_poly (CNF 58)

tricycle (CNF 18)

hammer_and_pegs_game (CNF 55)

```
suggested_toy
!    Testing 1
!    !    child_age
!    !    !    (= one_to_three CNF 100 )
!    Testing 2
!    Testing 3
!    Testing 4
!    Testing 5
!    !    price
!    !    !    (= under_25 CNF 65 )
!    !    !    (= over_25 CNF 20 )
!    (= roly_poly CNF 58 )
!    Testing 6
!    (= tricycle CNF 18 )
!    Testing 7
!    !    gender
!    !    !    (= male CNF 100 )
!    (= hammer_and_pegs_game CNF 55 )
!    Testing 8
!    Testing 9
!    Testing 10
!    Testing 11
!    Testing 12
!    Testing 13
!    Testing 14
!    Testing 15
```

Perhaps unintended 1991 gender-bias in RULES and /or in CNF's of RULES should be adjusted

Example **Symbolic AI** rule-based Expert System with Confidence Factors (CNF's)

J. Wunderlich, 1991

Perhaps unintended 1991 gender-bias in RULES and/or in CNF's of RULES should be adjusted

```
RULE 7
IF child_age = one_to_three AND
price = under_25 AND
gender = male
THEN suggested_toy = hammer_and_pegs_game CNF 85;

RULE 10
IF child_age = four_to_six AND
price = over_25 AND
gender =  male AND
child_preference = action_toys
THEN suggested_toy = go_cart CNF 85;

RULE 11
IF child_age = four_to_six AND
price = over_25 AND
gender = female AND
child_preference = creative_toys
THEN suggested_toy = doll_house CNF 90;

RULE 12
IF child_age = four_to_six AND
price = under_25 AND
gender = female AND
child_preference = cuddly_toys
THEN suggested_toy = dress_up_doll CNF 90;

RULE 13
IF child_age = four_to_six AND
price = over_25 AND
gender = male AND
child_preference = action_toys
THEN suggested_toy = hot_wheels_set CNF 95;

RULE 14
IF child_age = four_to_six AND
price = under_25 AND
gender = female AND
child_preference = creative_toys
THEN suggested_toy = toy_tea_set CNF 85;

RULE 15
IF child_age = four_to_six AND
price = under_25 AND
gender = male AND
child_preference = creative_toys
THEN suggested_toy = army_men CNF 90;
```

Anna Elizabeth Wunderlich, born June15th, 2002

Survey

On a blank peace of paper answer the following questions and hand it in without your name on it:

1. Who is the **smartest person** ever?
2. What specific things **define** their **smartness**?
3. Name a few of the most important things that define **<u>you</u>** as **smart?**

How does Watson fit on Dr Wunderlich's **Mental Ability Matrix** ?

		Can human do?	Can bug do? (spider)	Can Conventional Computer Program do?	Can Symbolic AI Program do?	Can Artificial Neural Network do?	Comments
	BASIC ANIMAL ABILITIES:						
1	Acquire and retain knowledge	yes	yes	yes	yes	yes	
2	Solve problems	yes	yes	yes	yes	yes	
3	Learn and adapt	yes	yes	no	somewhat	yes	Evolution
4	Motor coordination	yes	yes	somewhat	somewhat	somewhat	Survival
5	Acquire energy	yes	yes	somewhat	somewhat	somewhat	Survival
6	Protect self	yes	yes	somewhat	somewhat	somewhat	Survival
7	Sensory processing	yes	yes	yes	yes	yes	
8	Real-time thought	yes	yes	yes	yes	yes	
9	React instinctively	yes	yes	no	not yet	not yet	
10	Anticipate	yes	yes	yes	yes	yes	
11	Predict	yes	yes	yes	yes	yes	
12	Communicate	yes	yes	yes	yes	yes	
13	Generalize	yes	yes	no	somewhat	yes	
14	Associate	yes	yes	somewhat	somewhat	yes	
15	Recognition patterns	yes	yes	somewhat	somewhat	yes	
16	Robust under partial failure	yes	yes	no	no	yes	
17	Autonomous thought	yes	yes	if programmed	somewhat	soon	How tightly to hold the leash?
18	Drive to reproduce	yes	yes	no	not yet	not yet	
19	Stability, repeatability, predictability	somewhat	somewhat	yes	yes	somewhat	Uncertainty
20	Multitask	to a point	yes	yes	no	yes	

How does Watson fit on Dr Wunderlich's **Mental Ability Matrix** ?

	COMPLEX ABILITIES:	Can human do?	Can bug do? (spider)	Can Conventional Computer Program do?	Can Symbolic AI Program do?	Can Artificial Neural Network do?	Comments
21	Abstraction	yes	unlikely	no	no	somewhat	
22	Intuition	yes	unlikely	no	not yet	soon	
23	Common sense	yes	yes	no	not yet	soon	
24	Manipulate tools	yes	no	yes	yes	yes	Evolution
25	Heuristics	yes	yes	somewhat	yes	no	
26	Inference	yes	yes	somewhat	yes	somewhat	
27	Hypothesis testing	yes	somewhat	somewhat	yes	no	
28	Self-discipline, impulse-control	yes	unlikely	no	somewhat	no	
29	Ethical behavior	yes	unlikely	no	somewhat	somewhat	If coded/trained
30	Selective awareness (filtering)	to a point	yes	yes	yes	yes	
31	Open to inspection	somewhat	somewhat	yes	yes	somewhat	
32	Emotions	yes	Somewhat	no	not yet	soon	
33	Imagination	yes	Somewhat	no	not yet	soon	
34	Creativity	yes	Somewhat	no	not yet	soon	
35	Passion	yes	unlikely	no	not yet	soon	
36	Playfulness	yes	unlikely	no	not yet	soon	Evolution
37	Empathy	yes	unlikely	no	not yet	soon	
38	Courage	yes	unlikely	no	not yet	soon	
39	Leadership	yes	unlikely	no	not yet	not yet	
40	Self awareness	yes	unlikely	no	not yet	not yet	
41	Awareness of mortality	yes	unlikely	immortal?	Immortal?	Immortal?	Replaceable parts
42	Group psychology, Social Networking, and Living in the Cloud(s)	yes	unlikely	somewhat	somewhat	somewhat	Networking, Crowd-sourcing, Socially-networked Design

Questions to help define what is "Intelligence"

- What can a **human** do?
- What can a lesser animal do? (a **spider** for example)
- What can a **conventional computer** do?
- What can **symbolic Artificial Intelligence (AI)** do?
- What can an artificial **Neural Network (NN)** do?

Questions to help define what is “Intelligence”

- All of the following slides need to be updated to answer “What can Watson do?”

- And more specifically

  “What can recent variations of Watson do?”

5 Questions, 42 Mental Abilities to help define what is “Intelligence”

- Answer:
  1. What can a **human** do?
  2. What can a **spider** do?
  3. What can **conventional computer** do?
  4. What can **Symbolic AI** do?
  5. What can **Neural Network** do?

- for:
  - 20 **“Basic Animal Abilities”**
  - 22 **“Complex Abilities”**

(1) Acquire & retain knowledge, (2) Solve problems

- Humans rely heavily (perhaps too much) on Standardized tests (SAT,GRE,MCAT,etc) for these
- Bug solves placing web.
  - Knowledge of environment and prey
- Conventional and ***intelligent* machines**
  - Solve problems and retain knowledge
  - Only differ in memory capacity, method of storage, and class of solvable problems

(3) Learn & Adapt

- Animals learn & adapt to new environments
- In real and evolutionary time
- Conventional computers don’t (need human)
- **Symbolic AI** can reason
  - Somewhat adaptable to new input
- **Neural Networks** better
  - Generalize when presented new inputs
  - Extremely fast in real-time when embedded

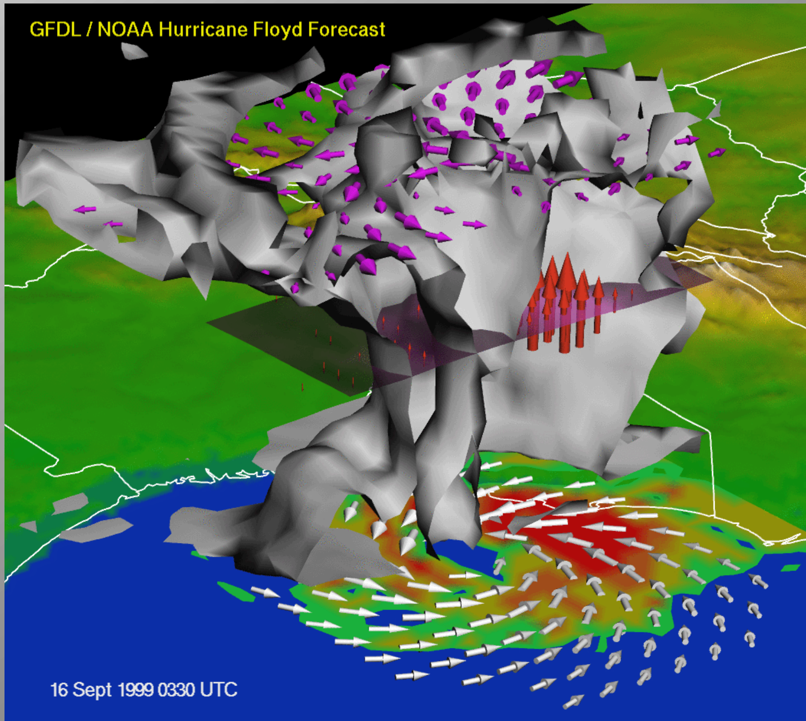
(4) Motor Coordination
(5) Acquire Energy
(6) Protect self

- “Mobility,” “Acquisition,” “Protection”
  - Essential for animal survival

- Partially solved by conventional and ***intelligent* machines**
  - Robotic Motor control
  - Power Supplies
  - Firewalls

(7) Sensory processing
(8) Real-time thought
(9) React instinctively
(10) Anticipate
(11) Predict

- Animals sense, think quickly/instinctually
  - Anticipate outcomes
  - Predict based on extrapolating info
- Conventional and ***intelligent* machines** do this
  (with the exception of instinct)
  - Neural Networks outperform symbolic A.I. when dealing with new stimulus
    - Can be much faster (especially if embedded)



(12) Communication

- Animals, conventional computers, ***intelligent machines*** all communicate
- Nothing close to human natural language processing
  - **Symbolic AI** attempted for decades
- **Neural Networks** very successful in speech recognition (IBM ViaVoice, DragonSpeak, etc.)
  - Including understanding “context”
  - And perhaps not loosing meaning in translations
  - *Krushchev’s "**We will bury you”*** was actually *"Whether you like it or not, history is on our side. We will dig you in“* or ***“We will outlive you”***

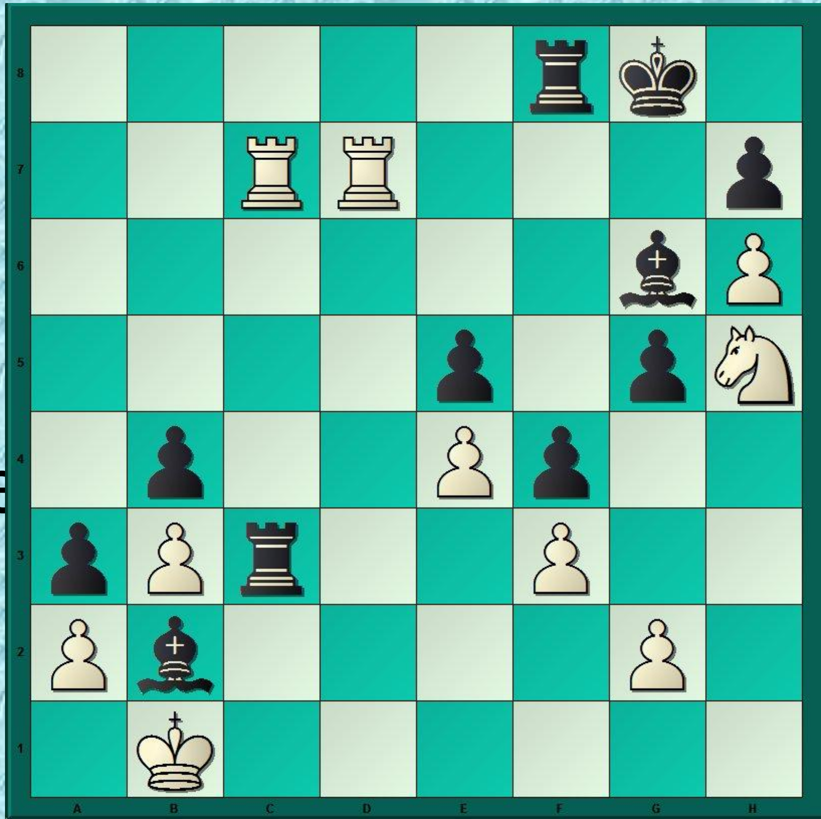
VIDEO:
http://www.youtube.com/watch?v=OwJHg9UBNPE

(13) Generalize

- *"To derive or induce a general principle or concept from particulars"*
- Animals do well
- Other than simple sorting, conventional computers and code typically don't
  - Give specific responses to specific inputs
- **Symbolic AI** only do somewhat
- **Neural Networks** good at
  - Generalize so outputs produced to "best fit" (classify) a set of inputs (<u>even when differ from what trained with</u>)

(14) Associate,
(15) Recognize patterns

- All animals do well
- No animal surpasses human’s ability to associate concepts and memories
- Conventional computers correlate data and recognize patterns
- **Symbolic AI** does better
  - But limited by fixed structure
  - “*State-space*” fixed regardless of efficiency of search
- **Neural Networks** even bette
  - Widely used for recognizing image and speech



(16) Robust under partial failure

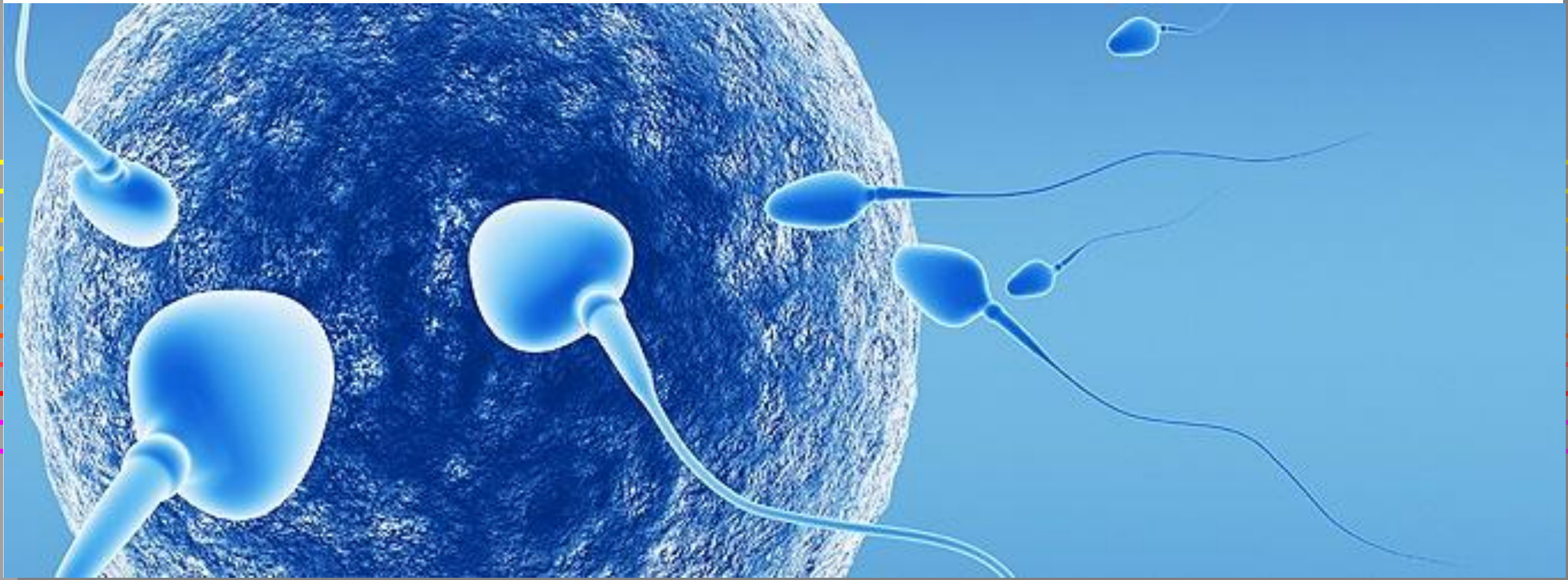
- Evolution insures animals partially function when some subsystems fail (including brain)
- Conventional computers can't
  - Bit error can cause "lock-up"
- **Symbolic AI** also likely to not function when the underlying computer hardware fails
- **Neural Networks** robust under partial failure
  - Will partially function if embedded NN hardware or NN software simulation looses some neurons or inter-neuron connections; however if NN is running on a computer that fails, it will fail

(17) Autonomous thought,

- Animals free to make decisions
- Conventional and **symbolic AI** programs respond in pre-programmed way, but, for example, **robotic autonomous global path-planners** are very much capable of incorporating symbolic AI
- **Neural Networks** can be also be part of these planners

(18) Drive to reproduce

- Animals free to make decisions
  - Struggle with programming dictated by genes
    - Including drive to reproduce
- Far in the future, if ever, for ***intelligent* machines**



(19) Stability, Repeatability, Predictability

- Some uncertainty with all animals
  - *Allows free will?*
- Conventional and **symbolic AI** don't have free will, but are generally repeatable, predictable, stable, and therefore traceable
- **Neural Networks** can produce unexpected results, especially with never-seen input, and even slightly different results with identical inputs
  - This is why FDA has denied permitting their use in pharmaceutical quality control

(20) Multitask

- Biological evolution yields brains with multiple subsystems
  - Regulation
    - Pulmonary, respiratory, temperature, and motor control
  - Pre-processing
    - Visual cortex, etc.
  - Higher reasoning
  - But we may be overloading !
- Conventional computers becoming better
- In **symbolic AI** when written for multi-processor
- Significant for **Neural network** learning

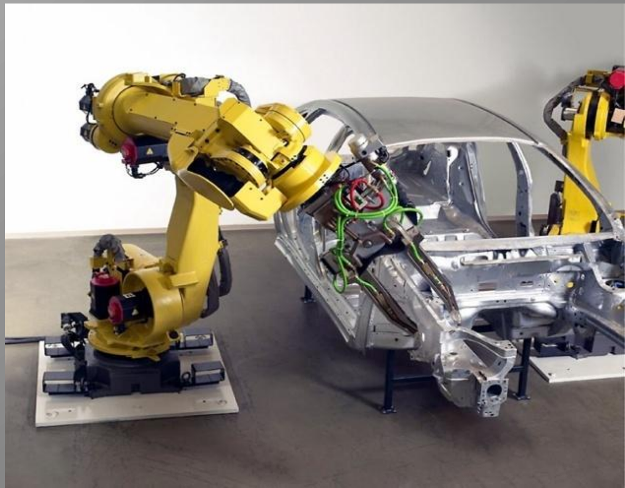
(21) Abstraction
(22) Intuition
(23) Common sense

- Abstract: "*.. only intrinsic form … no attempt at pictorial representation or narrative content"*
- Intuition: *"Knowing without conscious reasoning"*
- Common sense: *"Sound and prudent but often unsophisticated judgment "*

- All Unlikely for a bug
- Conventional and **symbolic AI** programs can't do
- **Neural Networks** generalize (abstraction?)

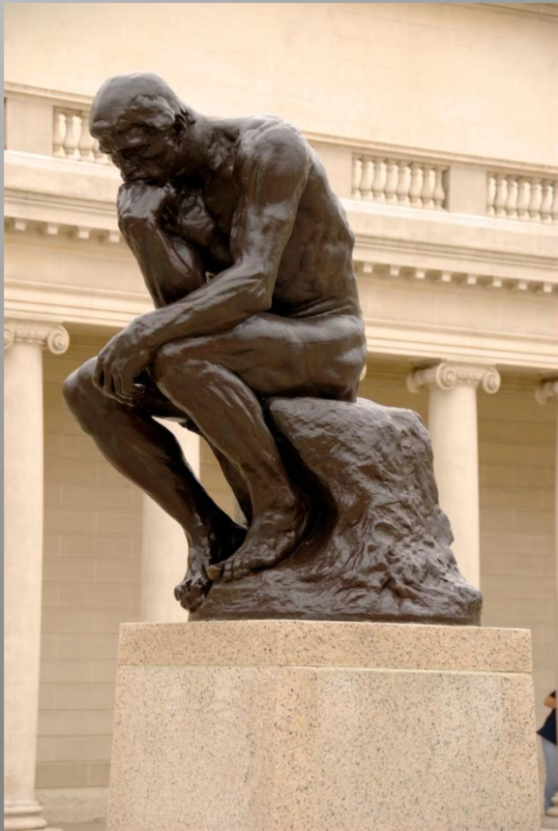
(24) Manipulate tools

- Evolved animals
  - Humans became bipedal so they could use front legs for manipulation of tools
- Spider can't envision extensions of its appendages
- Conventional and ***intelligent* machines** can
  - Signals to actuators to position & orient tools (robotic-arm)
  - Not only what arm holds, but arm itself is a tool to manipulate physical world



(25) Heuristics
(26) Inference
(27) Hypotheses testing

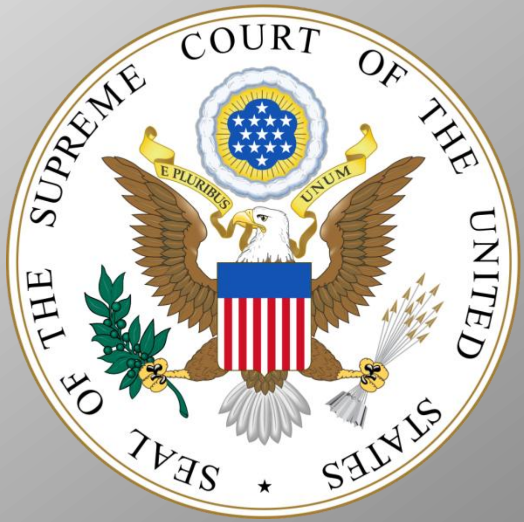
- Spiders, unlike humans, don’t typically:
  - Consider every way to react
  - Recognize when one scenario infers another
  - Solve problems by testing multiple hypotheses



- Conventional computers do (somewhat)
- **Symbolic AI** (especially “Expert Systems”) do exactly this !
- Most **Neural Networks** can’t
  - But somewhat infer results

(28) Self-discipline & Impulse control, (29) Ethical behavior

- Humans override "animal drives" and develop rules to maintain civilization
- Bugs don't



- Conventional computers don't
- **Symbolic AI** programs can incorporate rules
  - And therefore implied "values"
- **Neural Networks** can't
  - However could be trained to respond "ethically"

(30) Selective awareness (filtering)

- Animals focus on task
  - Ignoring distractions
- Find images semi-obscured by camouflage or clutter
- Conventional and **symbolic AI** programs pre-process
  - Signal and image processing
- **Neural Networks** perform well with “fuzzy” input

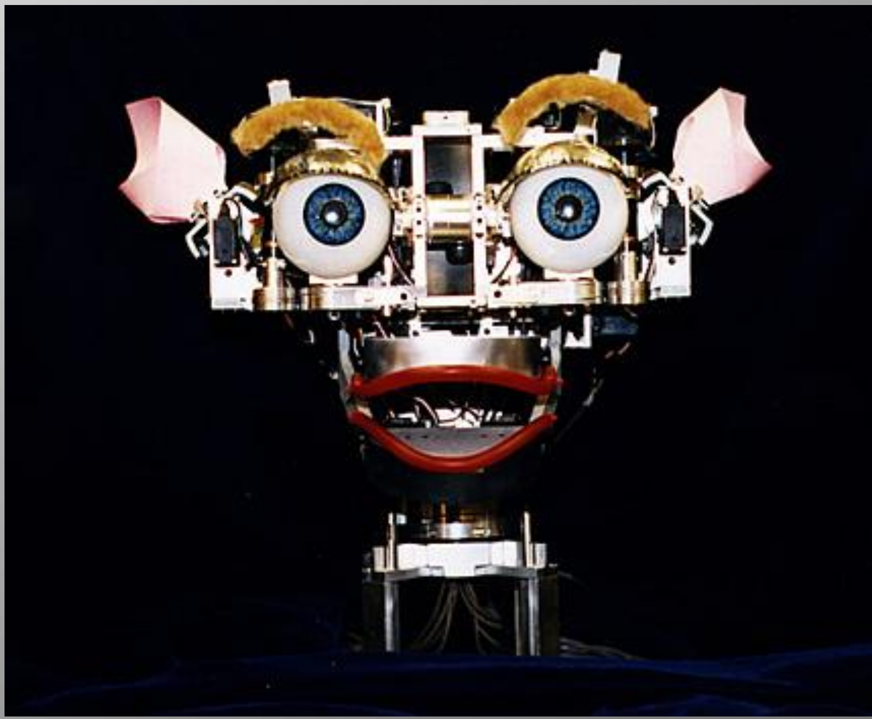
(31) Open to inspection

- Tracing mental thoughts less “exact” then tracing execution of conventional or **symbolic AI** program

- **Neural Networks** less open to inspection
  - Many compromises made in changing inter-neuron weights during learning

(32) Emotions
(33) Imagination
(34) Creativity
(35) Passion
(36) Playfulness

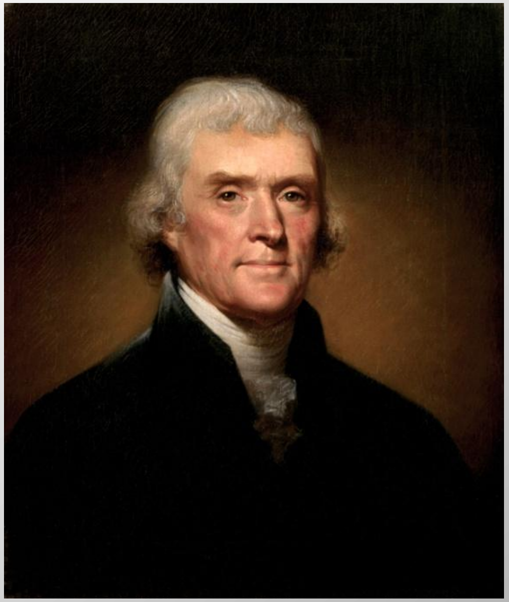
- To feel, imagine & create, have passion & ambition, and experiment through playful curiosity is still primarily human
- Play needed for evolution
- **Bugs do exhibit very basic emotions and creativity**
- No conventional or ***intelligent*** **machine** fully does this YET, but we're getting closer every day. These processes require multiple parts of our brains.



VIDEO: http://www.youtube.com/watch?v=Li4w9DPPNYI

VIDEO: http://www.youtube.com/watch?v=8KRZX5KL4fA

(37) Empathy
(38) Courage
(39) Leadership

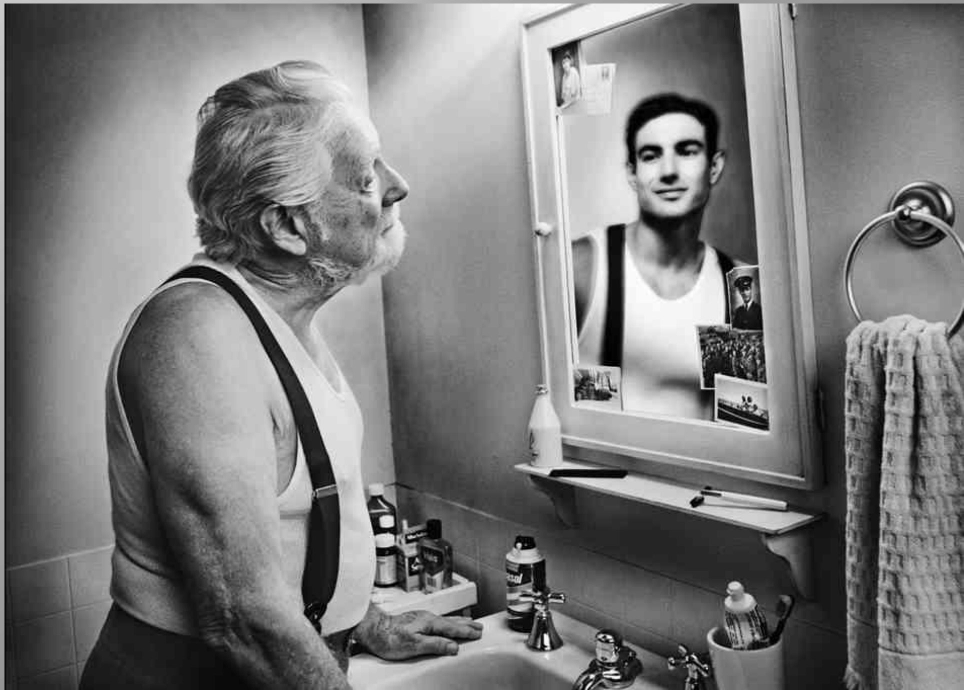


- Empathizing feelings of others, taking risks for others, and displaying leadership (vision, compassion, motivation) is primarily human

- Unlikely for a bug
- No conventional or ***intelligent*** **machines** do this yet, and it’s highly unlikely free people would want this

(40) Self-Awareness,
(41) Awareness of mortality

- Humans see themselves, their lives, their influence on others, their influence on the future, and their mortality
- Unlikely for bugs
- Conventional computers don't
- ***Intelligent* machines** becoming self-aware???
- But machine is "immortal" if ample supply of replacement parts



(42) Group Psychology, Social Networking, and getting lost in the Cloud(s)

- Humans play, work, raise children, and wage war as teams, and can collectively share beliefs
- Other animals work as a collective (not spiders)
- Conventional computers (including mobile devices), and ***intelligent*** **machines** can be networked
  - This has lead to group psychology!
  - Could social networking be a *"macro"* extension of the *"micro"* distributive brain-center processes needed for our most complex mental abilities?


浄土宗は、法然上人のみ教えをうけ、過去、いま、そして明日への「いのち」を慈しみ、新しい共生(ともいき)文化を創造します。
浄土宗21世紀劈頭宣言
法然共生
愚者の自覚を
家庭にみ仏の光を
社会に慈しみを
世界に共生を
浄土宗 総本山 知恩院


名城大学附属高等学校

More on Networking, Crowd-sourcing, and Socially-networked Design at:

PAPER: http://users.etown.edu/w/wunderjt/Green_Social_Designs_Japan_paper_19.pdf
TALK: http://users.etown.edu/w/wunderjt/Green_Social_Designs_Japan_TALK_19_PLUS.pdf

2012 Elizabethtown College student socially-networked architecture

2013 Key-note talk in Osaka, Japan

Wunderlich, J.T. and Wunderlich, J.J. (2013). Green architecture and environmental design using rapid-prototyping social-networking sandbox tools, followed by professional architectural software. *Asian Conference on Sustainability, Energy & the Environment* (ACSEE 2013), June 6-9, Osaka, Japan.

PAPER: http://users.etown.edu/w/wunderjt/Green_Social_Designs_Japan_paper_19.pdf

TALK: http://users.etown.edu/w/wunderjt/Green_Social_Designs_Japan_TALK_19_PLUS.pdf

A collaboration between the United Nations and Mojang, the developers of Minecraft

26 November 2012 Last updated at 06:28 ET

Minecraft to aid UN regeneration projects

Development plans for 300 places around the world are being modelled in Minecraft so residents can help decide how the locations will change.

Residents will be able to take a virtual stroll around the Minecraft models

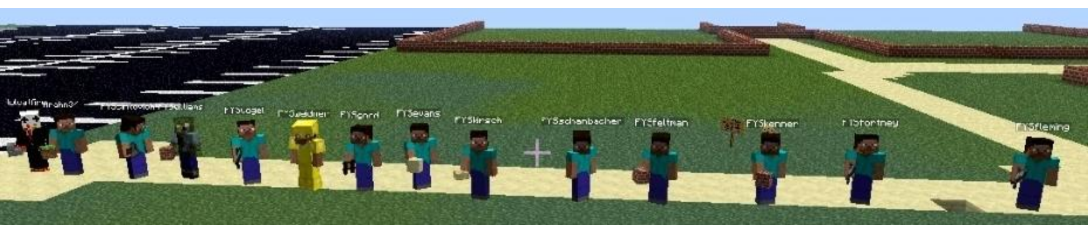
Called Block by Block, the programme is part of a collaboration between Minecraft-maker Mojang and UN Habitat.

Urban locations will be recreated on computer using Minecraft allowing residents to take a virtual tour.

They will also be able to change the model and help decide how regeneration cash should be spent.

(42) Group Psychology, Social Netwoking, and getting lost in the Cloud(s)

- Humans play, work, raise children, and wage war as teams, and can collectively share beliefs
- Many other animals work as a collective
- Conventional computers (including mobile devices), and ***intelligent* machines** can be networked
  - Could social networking be a *"macro"* extension of the *"micro"* distributive brain processes needed for our most complex mental abilities? – and possibly enhance civilization with a balance of autonomy and group harmony.



Minecraft to aid UN regeneration projects

Development plans for 300 places around the world are being modelled in Minecraft so residents can help decide how the locations will change.

Called Block by Block, the programme is part of a collaboration between Minecraft-maker Mojang and UN Habitat.

Residents will be able to take a virtual stroll around the Minecraft models

Urban locations will be recreated on computer using Minecraft allowing residents to take a virtual tour.

They will also be able to change the model and help decide how

Related Stories

Watch PBS Frontline Video "Digital Nation"
http://www.pbs.org/wgbh/pages/frontline/digitalnation/view/

and question Mental Abilities:

#20 ***"MULTITASKING“***

#30 ***“SELECTIVE AWARENESS (FILTERING)”***

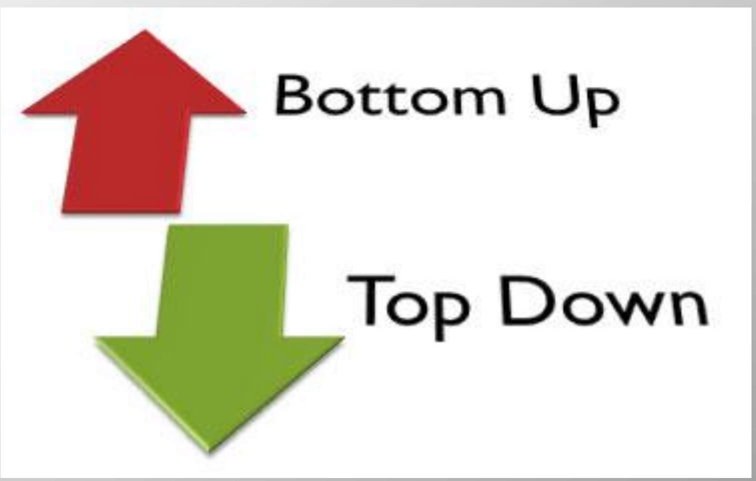
#42 ***"GROUP PSYCHOLOGY, SOCIAL NETWORKING, and LIVING IN THE CLOUD(s)"***

Then read:

Wunderlich, J.T. (2003). **Defining the limits of machine intelligence.** In *Proceedings of IEEE SoutheastCon, Ocho Rios, Jamaica,* [CD-ROM]. IEEE Press. Available on-line (updated): http://users.etown.edu/w/wunderjt/ITALY_2009/PUBLICATION_SUBMITPAPERdefininglimitsREVISED16b.pdf

Wunderlich, J.T. (2004). **Top-down vs. bottom-up neurocomputer design.** In *Intelligent Engineering Systems through Artificial Neural Networks, Proceedings of ANNIE 2004 Int'l Conference, St. Louis, MO.* H. Dagli (Ed.): Vol. 14. (pp. 855-866). ASME Press. ["Novel Smart Engineering System Design Award," 2nd runner-up best paper from over 300 submissions]; Available on-line: http://users.etown.edu/w/wunderjt/ITALY_2009/PUBLICATION_SUBMIT_FINAL_ANNIE2004_WUNDERLICH_61_TO_PRINT_fixed_after.pdf

Intelligent Machine Platforms and Devices



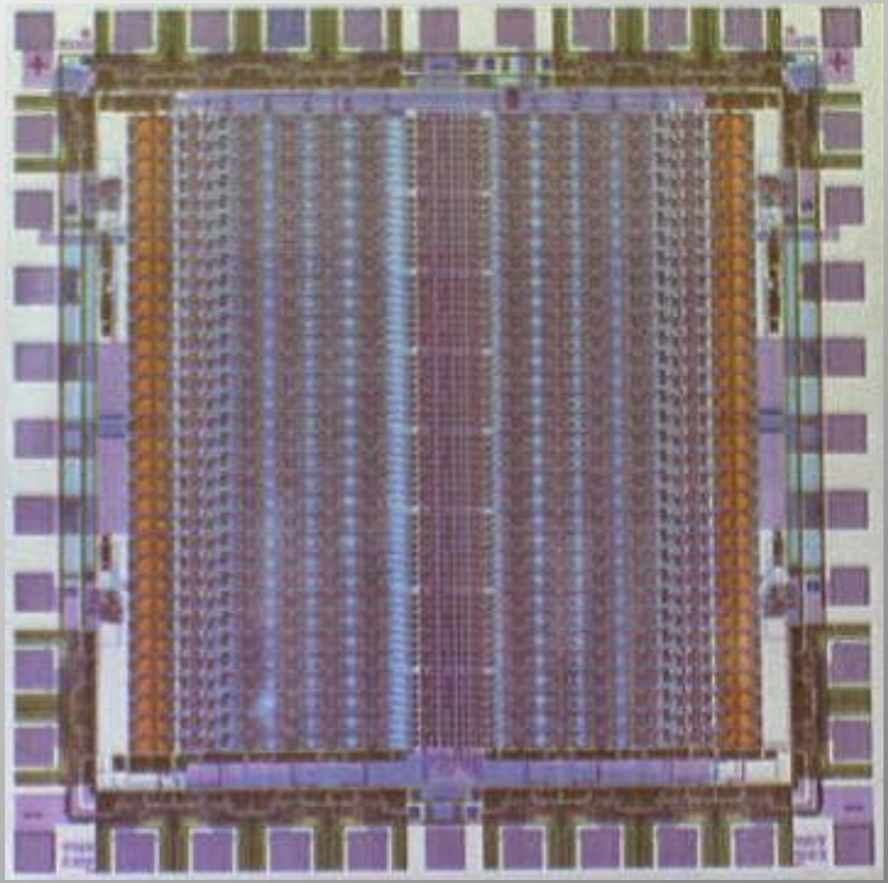
- Two approaches:

  1. Circuits similar to brain biology ("**Bottom-Up**" design)

  2. Produce results *similar* to human thought ("**Top-Down**" design)
     - Symbolic A.I. Heuristics, Inference, Hypothesis Testing
     - Neural Network Vector and Matrix Calculations

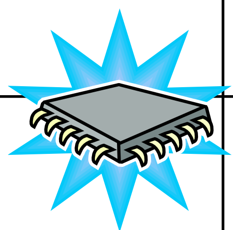
"Bottom-Up" circuits similar to brain biology (e.g., Artificial Dendritic Tree)

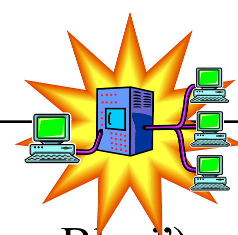
by Wunderlich and Elias, 1992



"Top-Down" Computing for Machine Intelligence

"LEVEL"	DEVICES	MACHINE INTELLIGENCE
Embedded	Microcontroller, Microprocessor ASIC	High-speed real-time-learning Neural Network. Not for "Symbolic AI"
PC	Microprocessor	ok for Neural Network simulations and "symbolic AI programming"
PC Server or Workstation	Multiple microprocessors Silicon Graphics , SUN , IBM RS6000	Good for Neural Network simulations and "symbolic AI programming"
Mini-Computer	IBM AS400, Amdahl, HP, Hitachi	Good for Neural Network simulations and "symbolic AI programming"
Super-Computer	SMP: IBM S/390 MPP: IBM SP2 (e.g.,"Deep Blue") Vector: CRAY	Great for Neural Network simulations and "symbolic AI programming"





Read more at:

- Wunderlich, J.T. (2003). **Defining the limits of machine intelligence**. In *Proceedings of IEEE SoutheastCon, Ocho Rios, Jamaica*, [CD-ROM]. IEEE Press.

- Wunderlich, J.T. (2004). **Top-down vs. bottom-up neurocomputer design**. In *Intelligent Engineering Systems through Artificial Neural Networks, Proceedings of ANNIE 2004 International Conference, St. Louis, MO.* H. Dagli (Ed.): Vol. 14. (pp. 855-866). ASME Press.

- Wunderlich, J.T. (2008). **Two single-chip neurocomputer designs; one bottom-up, one top-down**. *(invited journal paper in peer-review)*

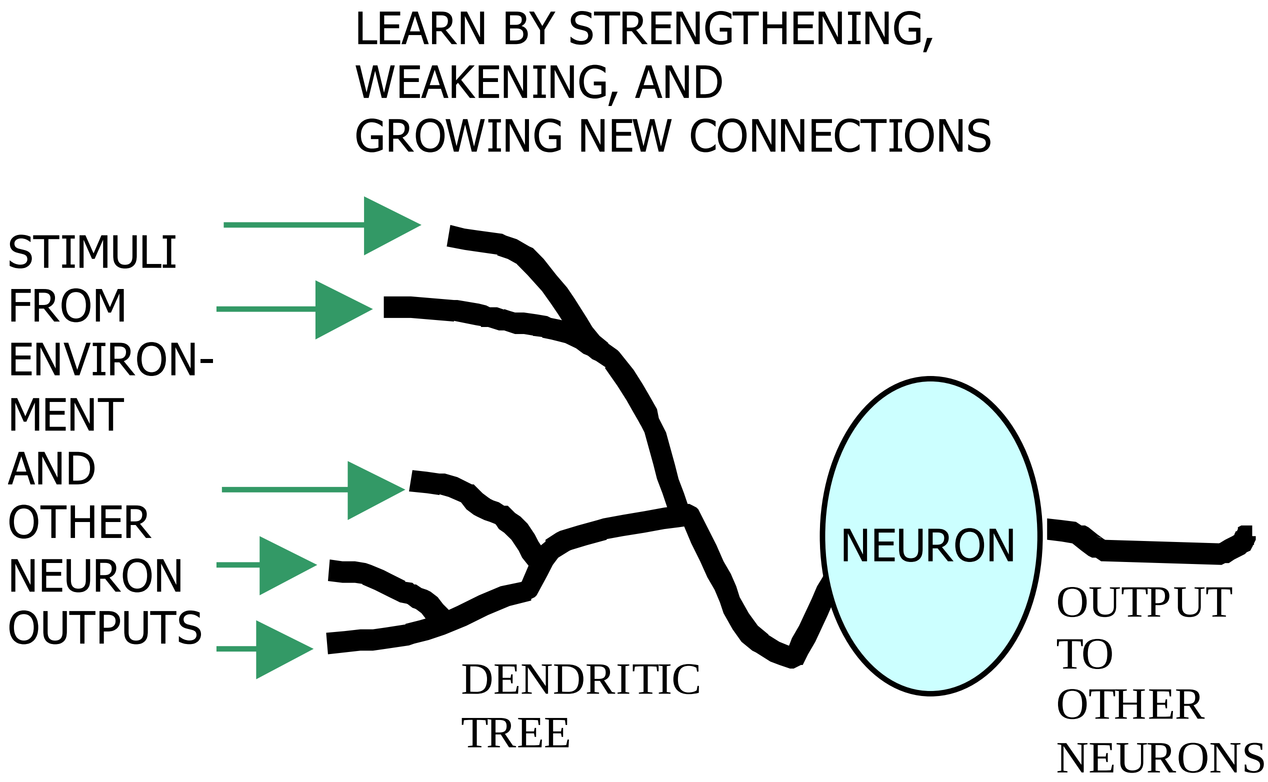
TOP-DOWN VS. BOTTOM-UP NEUROCOMPUTER DESIGN

JOSEPH T. WUNDERLICH PhD

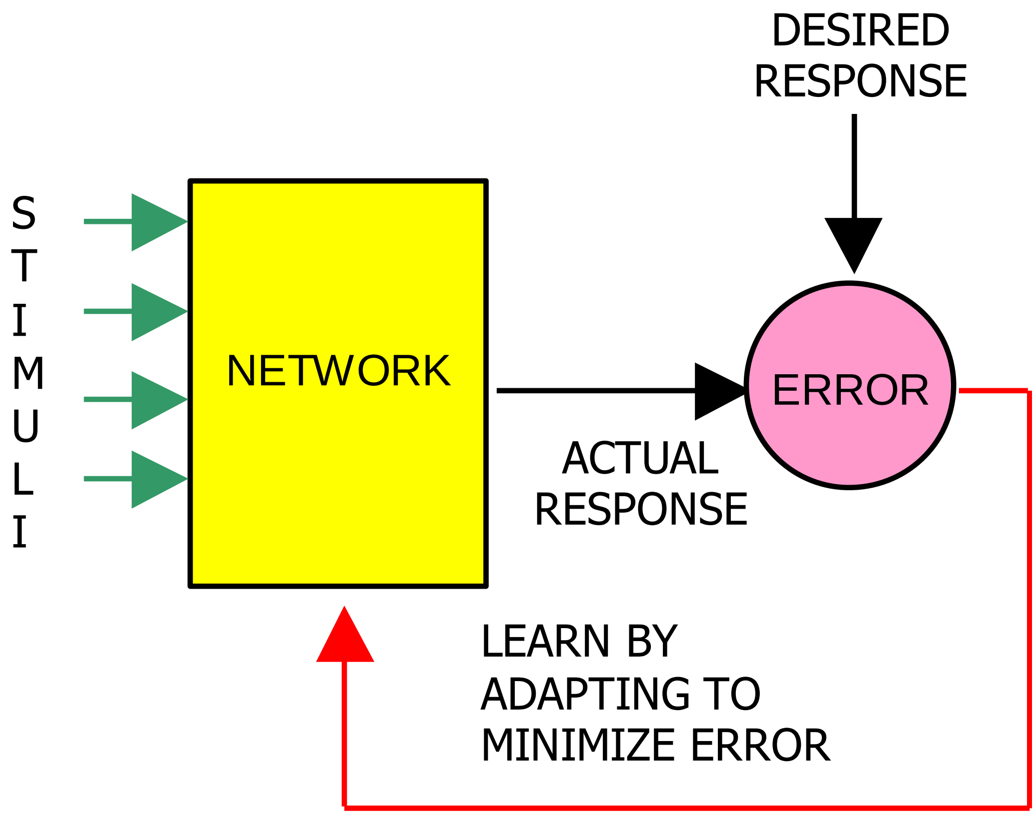
Agenda

- Top-down vs. bottom-up
  - Concepts
  - Mathematics
  - Learning
  - Two examples
    - Bottom-up artificial dendritic tree VLSI chip
    - Digital neurocomputer chip for top-down Neural Networks

Biological neuron for "bottom-up" neurocomputer design

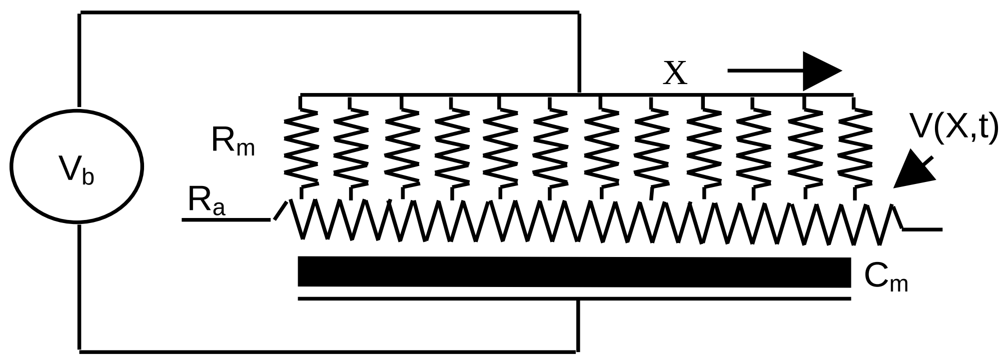


Behavioral model for "top-down" neurocomputer design



Analog circuit representation of biological model

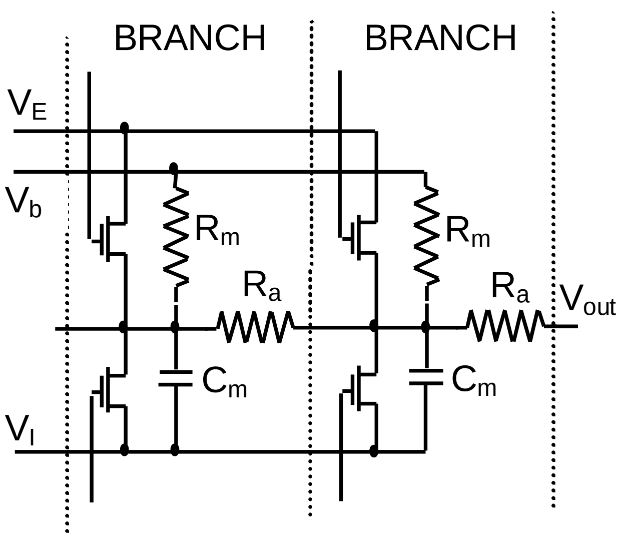
R_m = Neuron membrane resistance
C_m = Neuron membrane capacitance
R_a = Serial axial cytoplasm resistance
V_b = Resting neuron membrane voltage
V(X,t) = Neuron membrane voltage at location X and time t as a result of a current density I(X,t)



$$R_m * C_m * \frac{\partial V(X,t)}{\partial t} = \frac{R_m}{R_a} * \frac{\partial V^2(X,t)}{\partial X^2} - V + R_m * I(X,t)$$

Analog circuit for "bottom-up" design

R_m = Neuron membrane resistance
C_m = Neuron membrane capacitance
R_a = Serial axial cytoplasm resistance
V_b = Resting neuron membrane voltage
V_E = Excitation voltage (i.e., $V_E > V_b$)
V_I = Inhibitory voltage (i.e., V_I=GND)



- Output is a transient output voltage spike
- Learning involves varying inhibitions and excitations to minimize error between actual and desired output

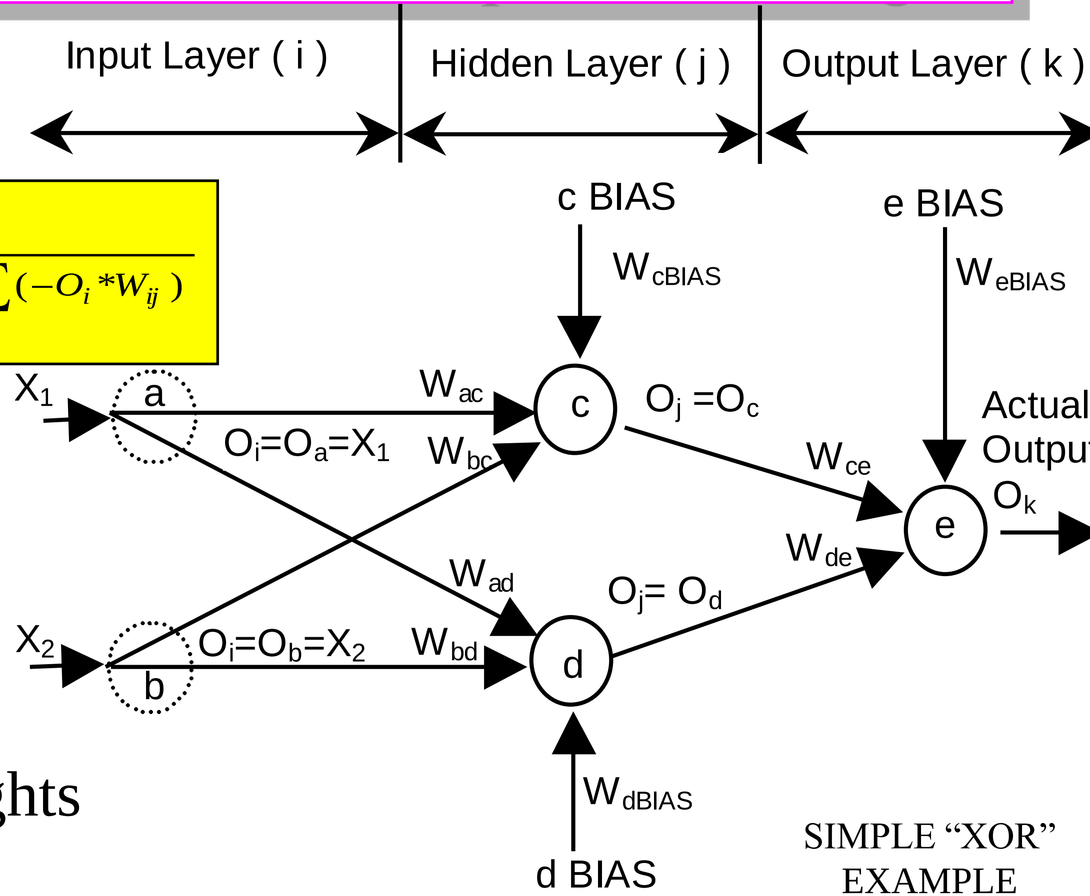
Mathematical models considered for “top-down” neurocomputer

- Backpropagation
- MADALINE III
- Hopfield
- BOLTZMANN MACHINE,
- BAM (Bi-directional Associative Memory)
- NEOCOGNITRON

Layered Neural Network for "top-down" design

Sigmoid transfer function:

$$O_j = \frac{1}{1+e^{-(jBIAS * W_{jBIAS}) + \sum_i (-O_i * W_{ij})}}$$



Learn by changing weights as a function of error:

$$\Delta W_{jk} = \eta * [(d_k - O_k) * O_k * (1 - O_k)] * O_j$$

$$\Delta W_{kBIAS} = \eta * [(d_k - O_k) * O_k * (1 - O_k)] * kBIAS$$

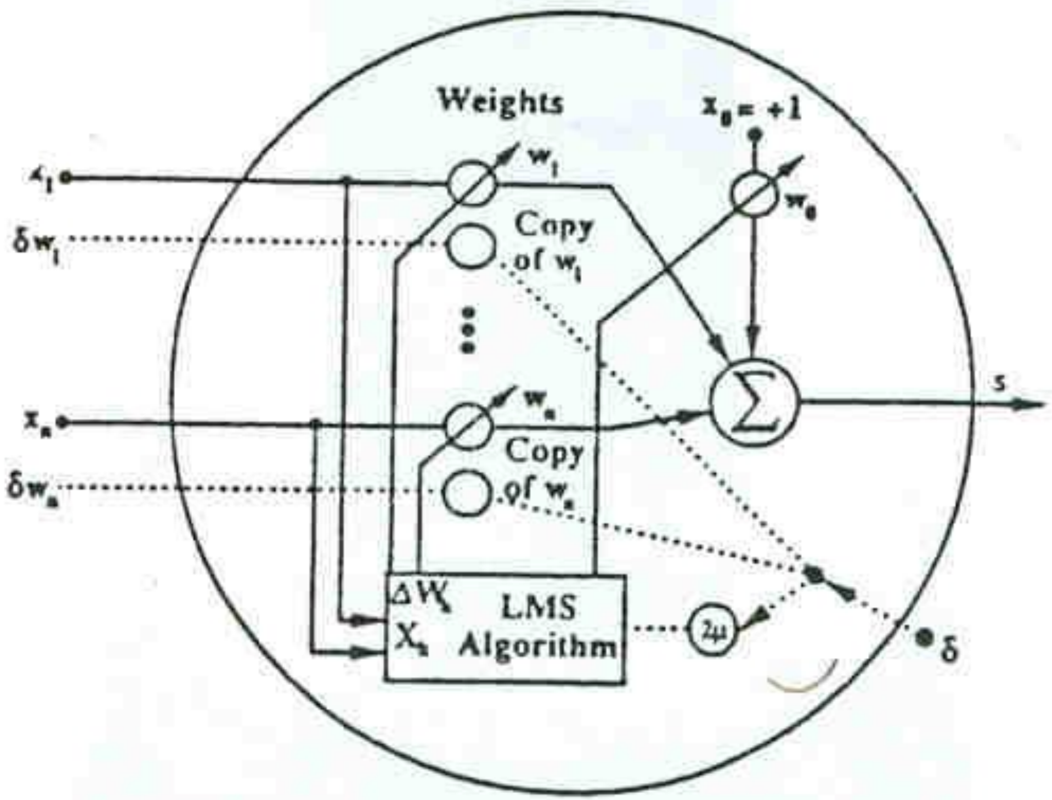
$$\Delta W_{ij} = \eta * \left(O_j * (1 - O_j)\right) * \sum_k \left[(d_k - O_k) * O_k * (1 - O_k) * W_{jk}\right] * O_i$$

$$\Delta W_{jBIAS} = \eta * \left(O_j * (1 - O_j)\right) * \sum_k \left[(d_k - O_k) * O_k * (1 - O_k) * W_{jk}\right] * jBIAS$$

SIMPLE "XOR" EXAMPLE

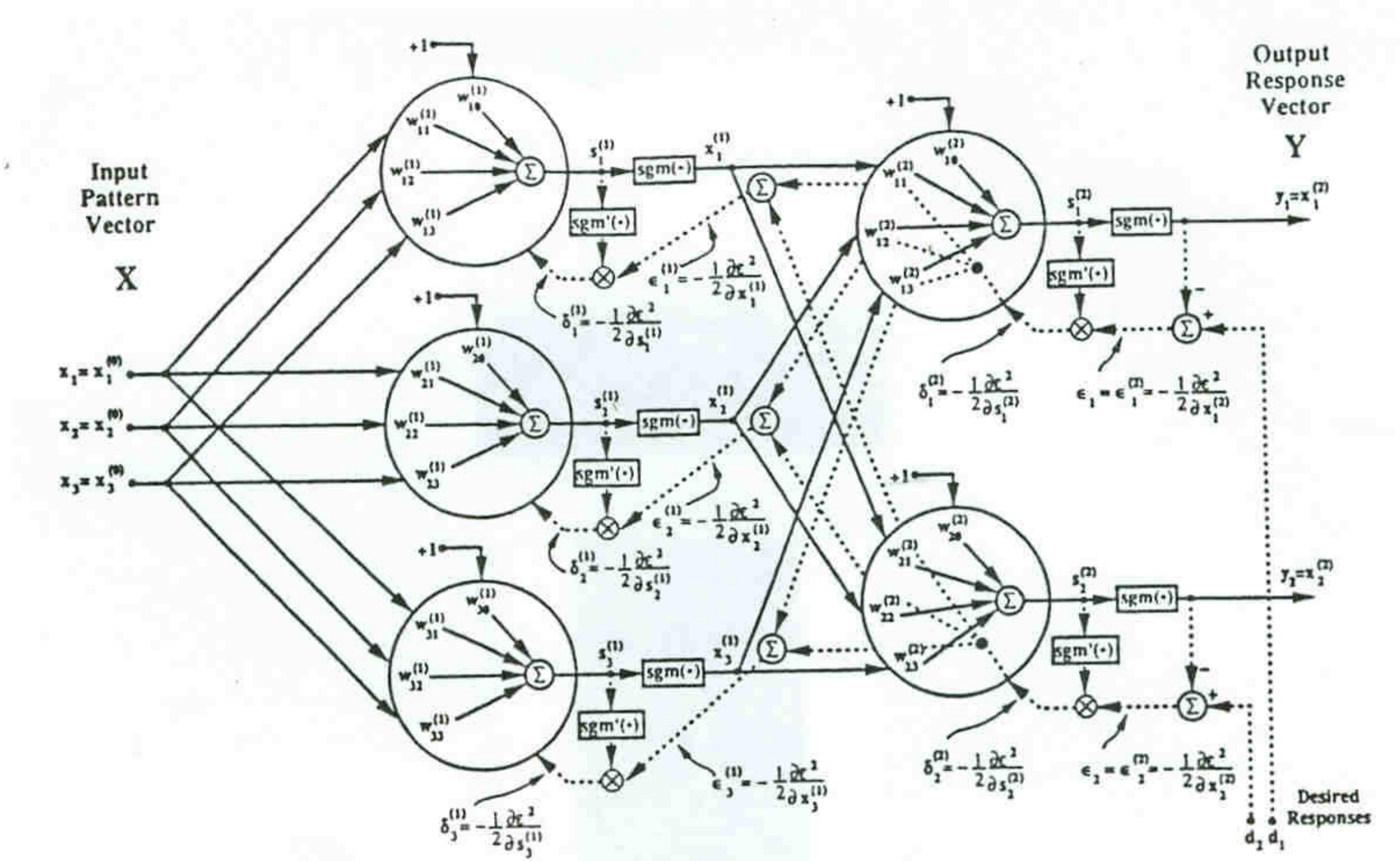
Input/Output Exemplars

X_1	X_2	Desired Output (d_k)
0	0	**0**
0	1	**1**
1	0	**1**
1	1	**0**



Detail of linear combiner and associated circuitry

From HANDOUT including Appendix A: "Historical Perspectives" excerpt from: Wunderlich, J.T. (1991). **A vector-register Neural-network microprocessor with on-chip learning.** *Masters Thesis*, Pennsylvania State University.



From HANDOUT including Appendix A: “Historical Perspectives” excerpt from: Wunderlich, J.T. (1991). **A vector-register Neural-network microprocessor with on-chip learning.** *Masters Thesis*, Pennsylvania State University.

Possible backprop neurocomputer chip implementations

- Discrete analog components for all computations
- Digital circuits for all computations except transfer function implemented as serial or parallel analog circuits
- *Analog circuits can suffer from drift, fabrication inconsistencies, and conversion delays*
- Digital circuits for all computations including transfer function implemented as serial or parallel look-up tables
  - *Parallel high-precision look-up tables use much real-estate*
- Parallel vector-register digital circuits for all computations including a polynomial approximation of the transfer function (CHOOSEN APPROACH)
  - Scalable, fast, single-chip with on-chip learning
  - 1991 J. Wunderlich Master’s Thesis and Patent Disclosure document

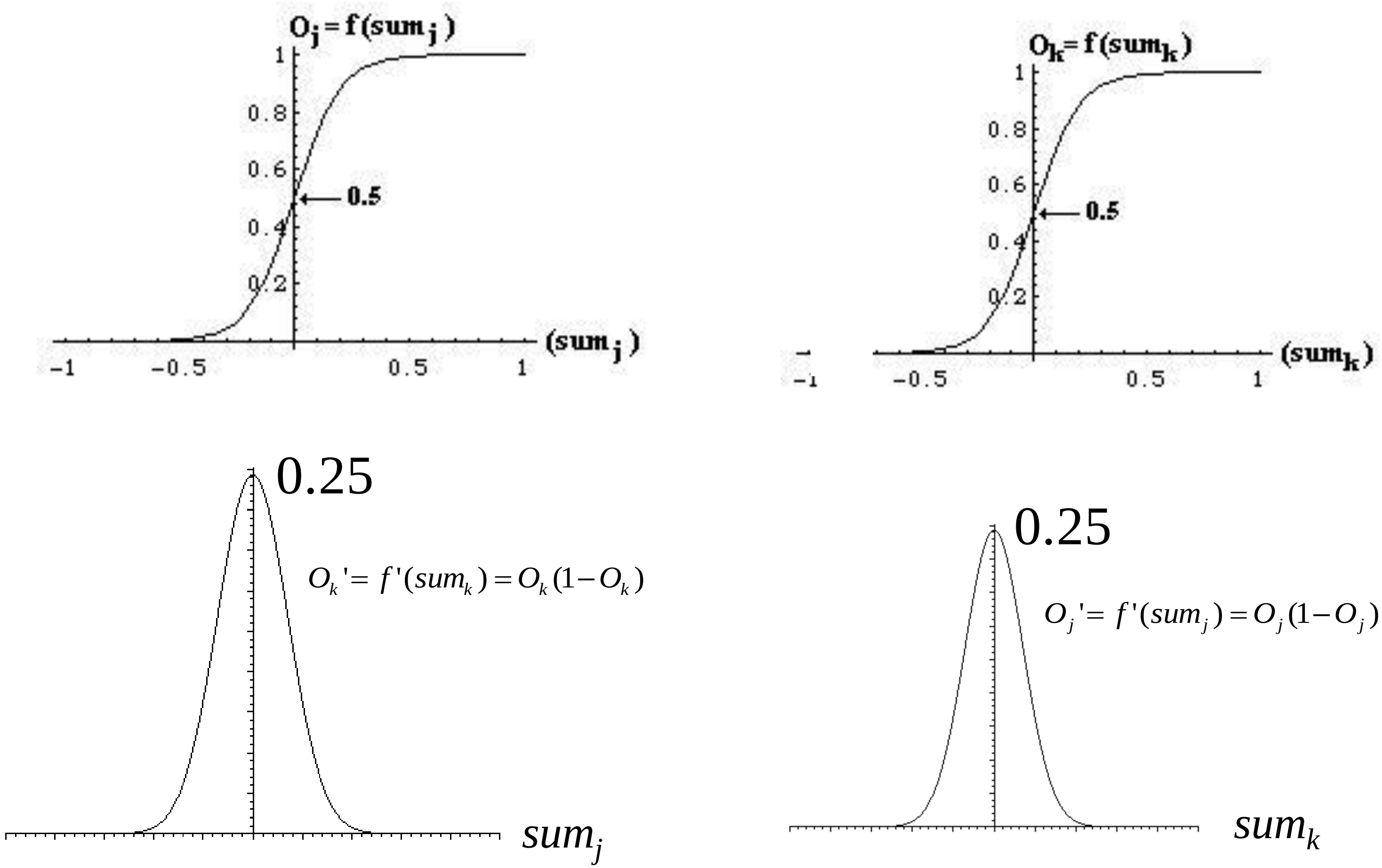
Polynomial approximation techniques considered

1991 J. Wunderlich Master's Thesis and Patent Disclosure document

- Taylor Series
- Hermite
- Divided Difference
- Cubic Spline
  - *This eliminated since piecewise requires separate calculations for each part of spline*

First recall that learning is most noticeable (and mathematically magnified) when neuron input summation is very small

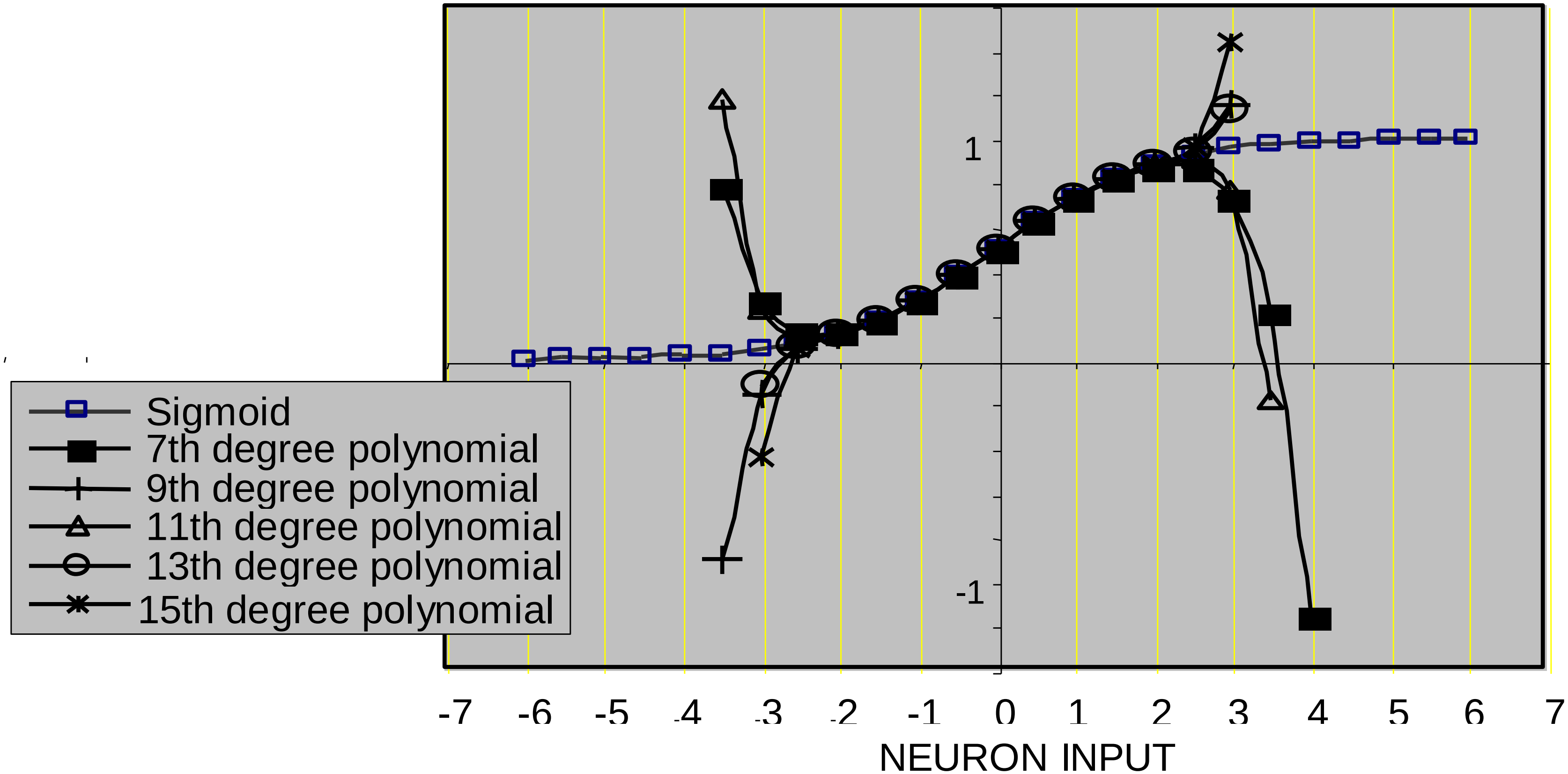
So we must maintain the highest of precision in a range of summated inputs centered around zero. The allowable width of the range is found via NN simulations concurrent to the numerical method techniques used to find a polynomial approximation of the Sigmoid transfer function



1991 J. Wunderlich Master's Thesis and Patent Disclosure document

Neuron output for Taylor polynomial approximations of sigmoid:
$1/(1+e^{-x})$ expanded about $x_0 = 0$

$$P_{Taylor}(x) = f(x_0) + f'(x - x_0) + f''(x_0)\left[\frac{(x-x_0)^2}{2!}\right] + \ldots\ldots\ldots + f^n(x_0)\left[\frac{(x-x_0)^n}{n!}\right]$$

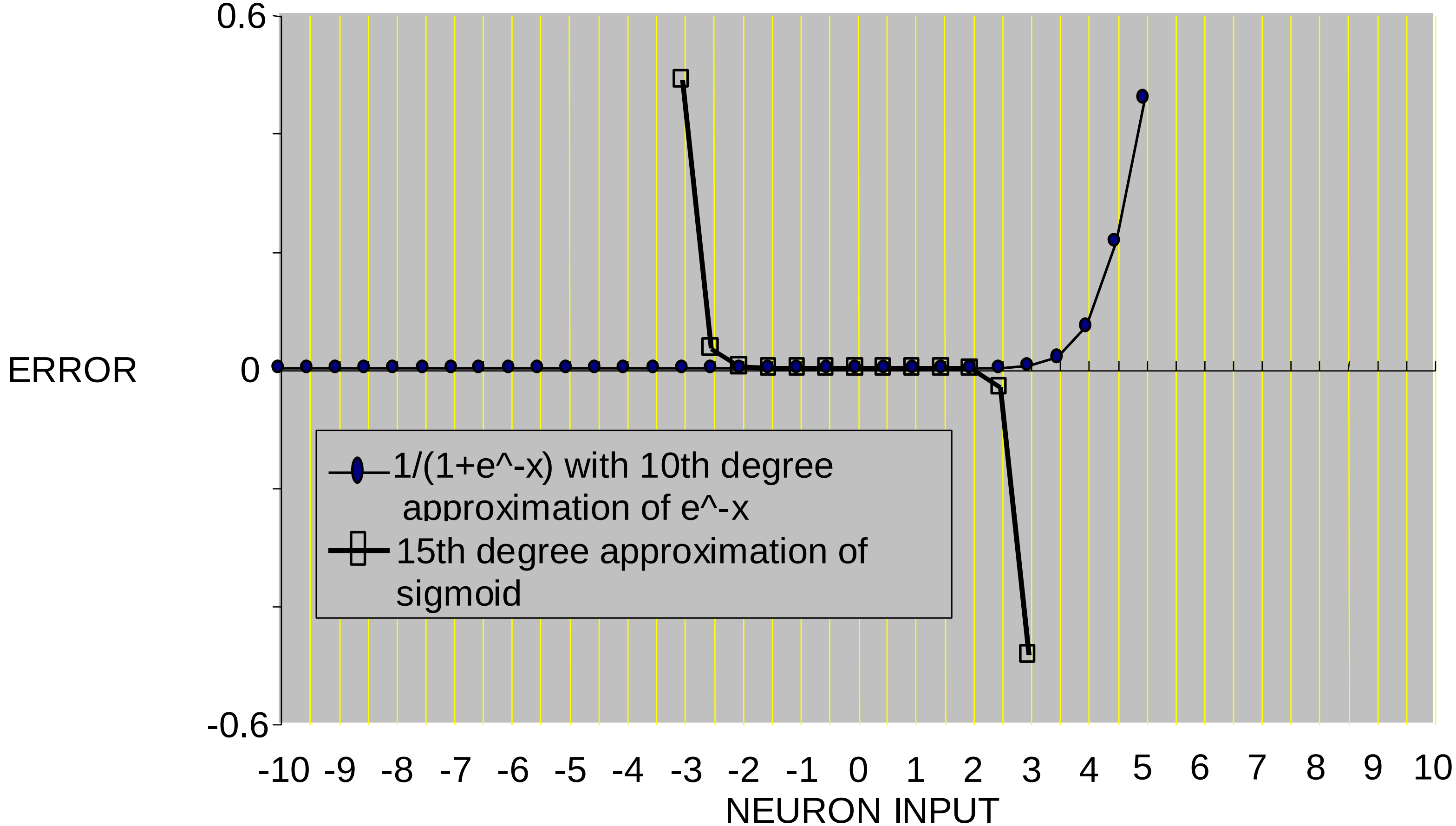


1991 J. Wunderlich Master's Thesis and Patent Disclosure document

Error for Taylor series polynomial approximations of sigmoid: $1/(1+e^{-x})$

$$Error_{Taylor} = f(x) - P_{Taylor}(x) = (x - x_0)^{(n+1)} \left[\frac{f^{(n+1)}(\xi(x))}{(n+1)!} \right]$$

for some number $\xi(x)$ between x and x_0



1991 J. Wunderlich Master's Thesis and Patent Disclosure document

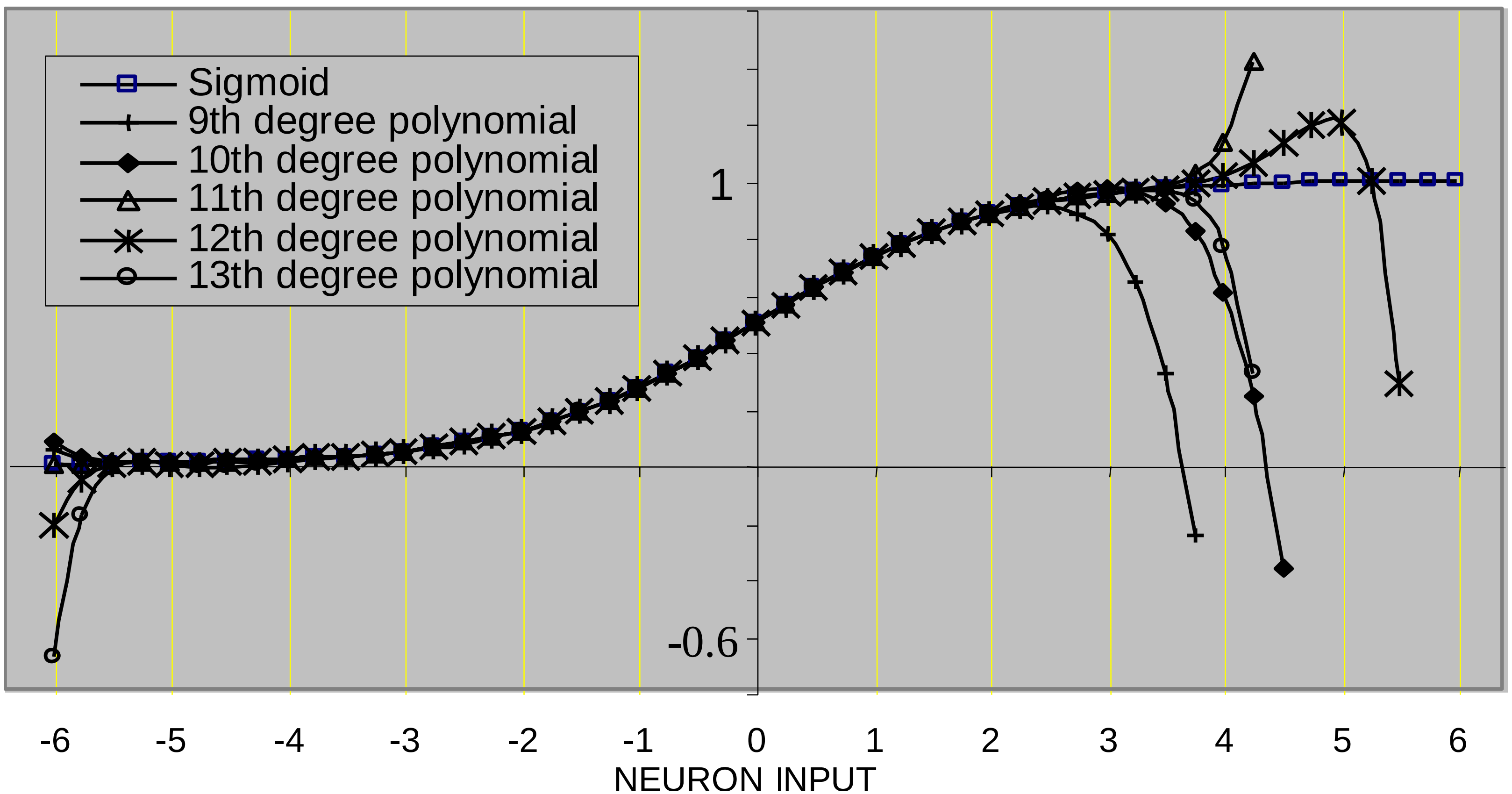
Neuron output for Hermite polynomial approximations of sigmoid: $1/(1+e^{-x})$ with evaluation points at x = [-6 to +6] at 0.25 intervals

$$P_{Hermite}(x) = \sum_{j=0}^{n} f(x)H_{n,j}(x) + \sum_{j=0}^{n} f'(x)\hat{H}_{n,j}(x)$$

$$H_{n,j}(x) = \left[1 - 2\left(x - x_j\right)L'_{n,j}(x_j)\right]L^2_{n,j}(x)$$

$$\hat{H}_{n,j}(x) = \left(x - x_j\right)L^2_{n,j}(x)$$

$$L_{n,j}(x) = \prod_{\substack{i=0 \\ i \neq j}}^{n} \frac{\left(x - x_i\right)}{\left(x_j - x_i\right)}$$

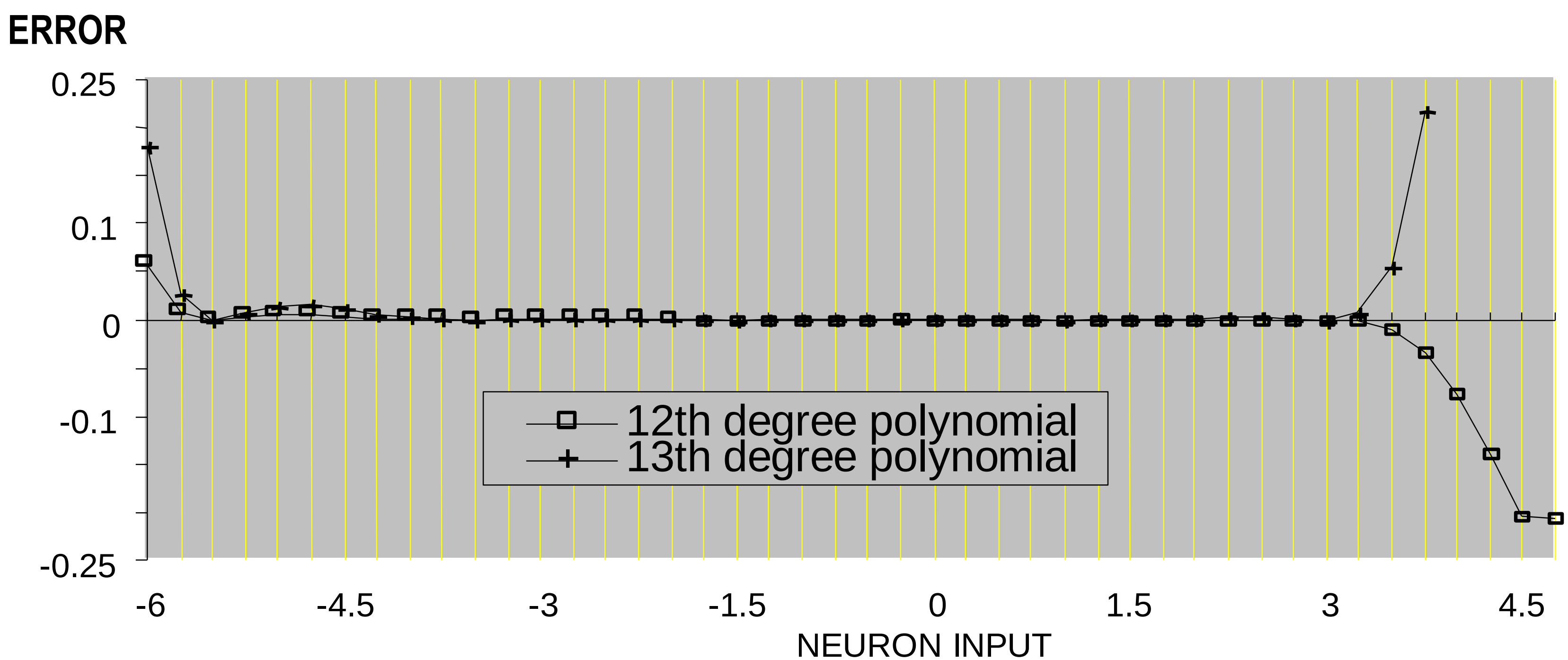


1991 J. Wunderlich Master's Thesis and Patent Disclosure document

Error for 12th and 13th degree Hermite polynomial approximation of sigmoid: $1/(1+e^{-x})$

$$Error_{Hermite} = f(x) - P_{Hermite}(x) = \left[\frac{\prod_{i=0}^{n} (x - x_i)^2}{(2n+2)!} \right] f^{(2n+2)}(\xi)$$

for some number between adjacent points in $(x_0, x_1, \ldots, x_n)$



1991 J. Wunderlich Master's Thesis and Patent Disclosure document

Neuron output with Divided-Difference polynomial approx of sigmoid: $1/(1+e^{-x})$ with evaluation points at x = [- 6, - 5, - 4, - 3, - 2, - 1, -0. 5, 0.5, 1, 2, 3, 4, 5, 6]

$$P_{DivDiff}(x) = a_0 + a_1(x-x_0) + a_2(x-x_0)(x-x_1) + \ldots + a_n(x-x_0)(x-x_1)\cdots(x-x_{n-1})$$

$$a_0 = f[x_0] = f(x_0)$$

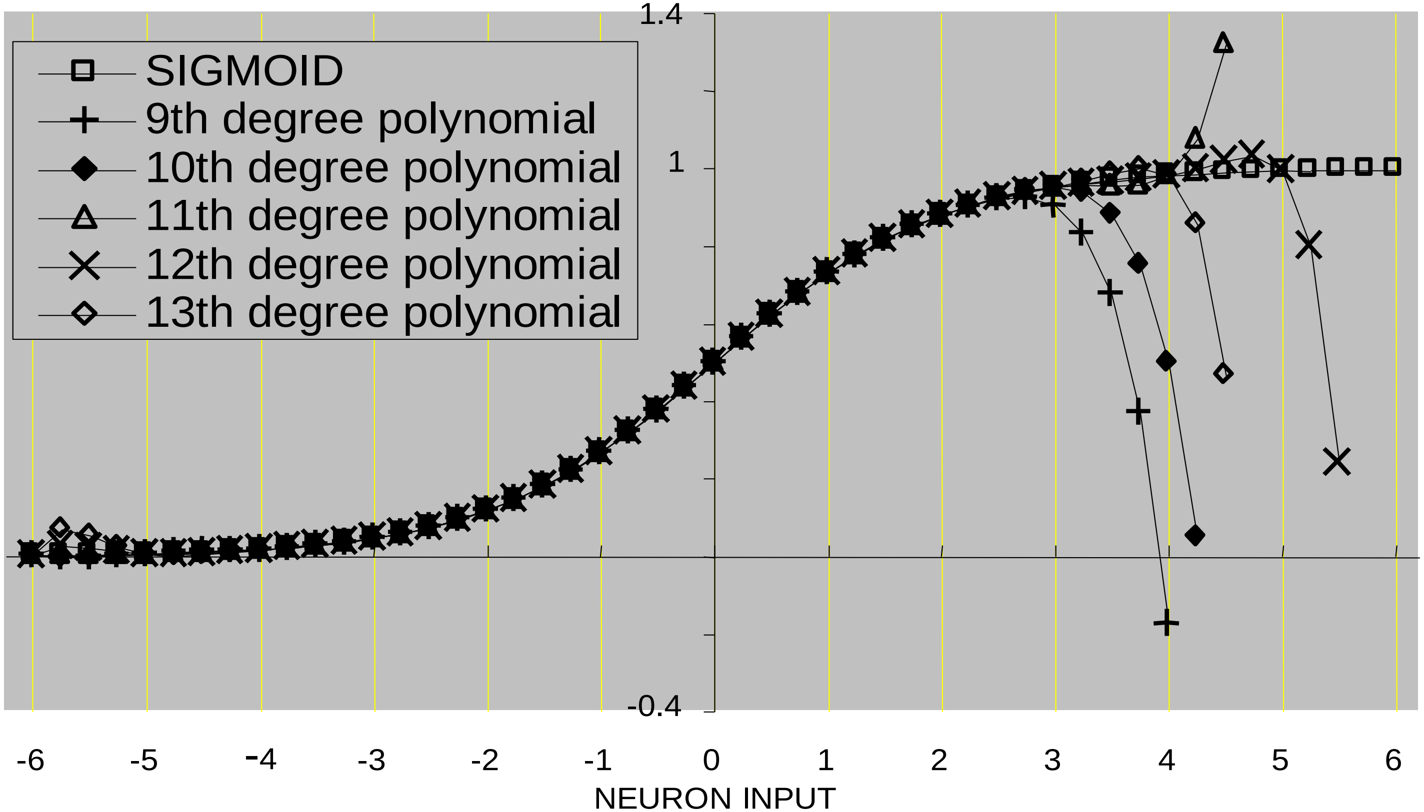
$$a_1 = f[x_0, x_1] = \frac{f[x_1] - f[x_0]}{(x_1 - x_0)}$$

$$a_2 = f[x_0,x_1,x_2] = \frac{f[x_1,x_2] - f[x_0,x_1]}{(x_2-x_0)} = \frac{1}{x_2-x_0}\left[\frac{f(x_2)-f(x_0)}{x_2-x_1} - \frac{f(x_2)}{x_1-x_0}\right]$$

$$= \frac{f(x_0)}{(x_0-x_1)(x_0-x_2)} + \frac{f(x_1)}{(x_1-x_0)(x_1-x_2)} + \frac{f(x_2)}{(x_2-x_0)(x_2-x_1)}$$

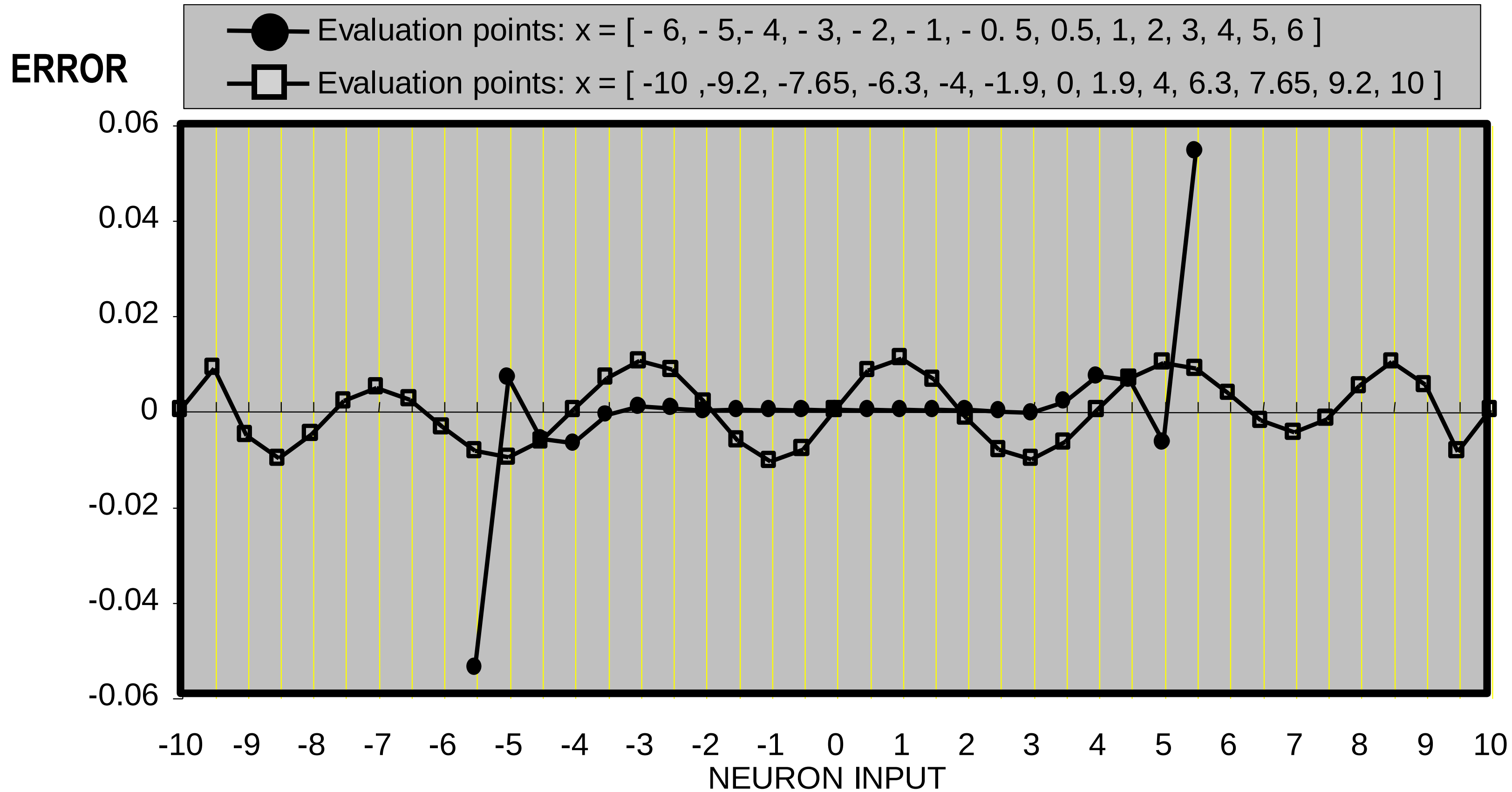
$$a_n = f[x_0, x_1, \ldots x_n]$$

$$= \frac{f(x_0)}{(x_0-x_1)\cdots(x_0-x_n)} + \frac{f(x_1)}{(x_1-x_0)\cdots(x_1-x_n)} + \ldots + \frac{f(x_n)}{(x_n-x_0)(x_n-x_{n-1})}$$



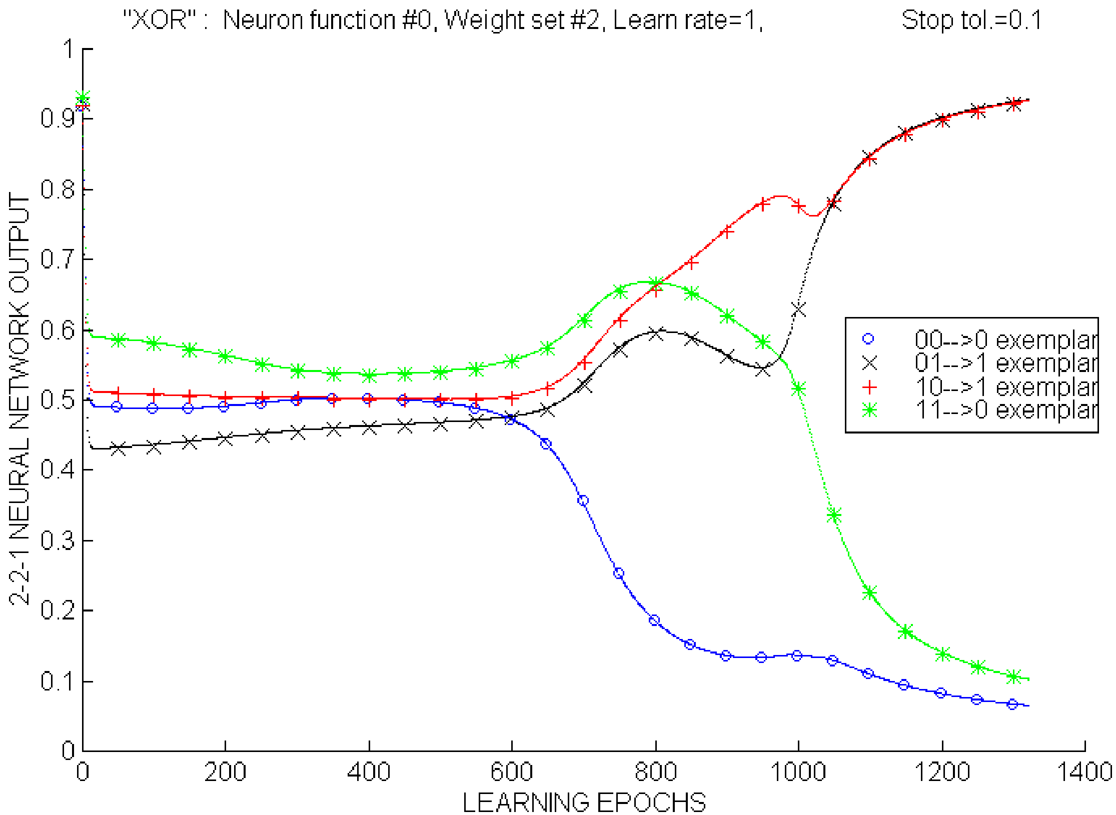
1991 J. Wunderlich Master's Thesis and Patent Disclosure document

Approximation error for 12th degree
Divided-Difference polynomial approximations
of sigmoid: $1/(1+e^{-x})$



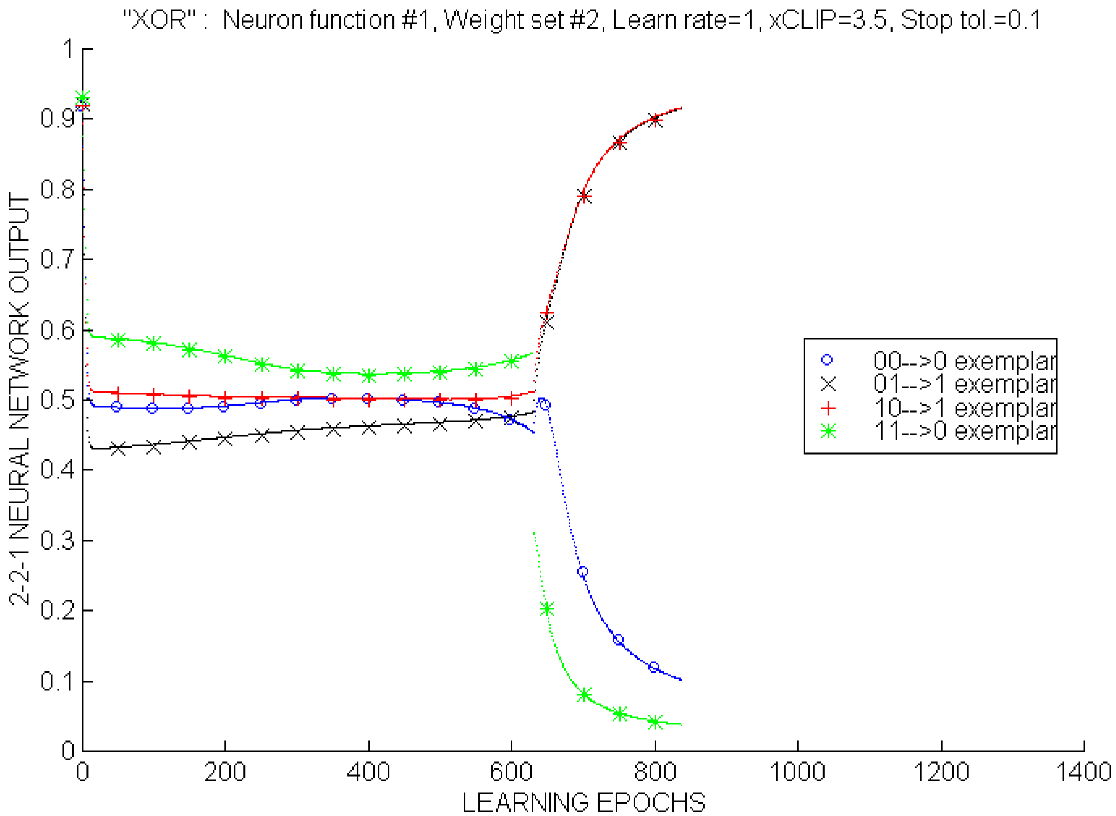
1991 J. Wunderlich Master's Thesis and Patent Disclosure document

Backpropagation learning "XOR" using standard sigmoid



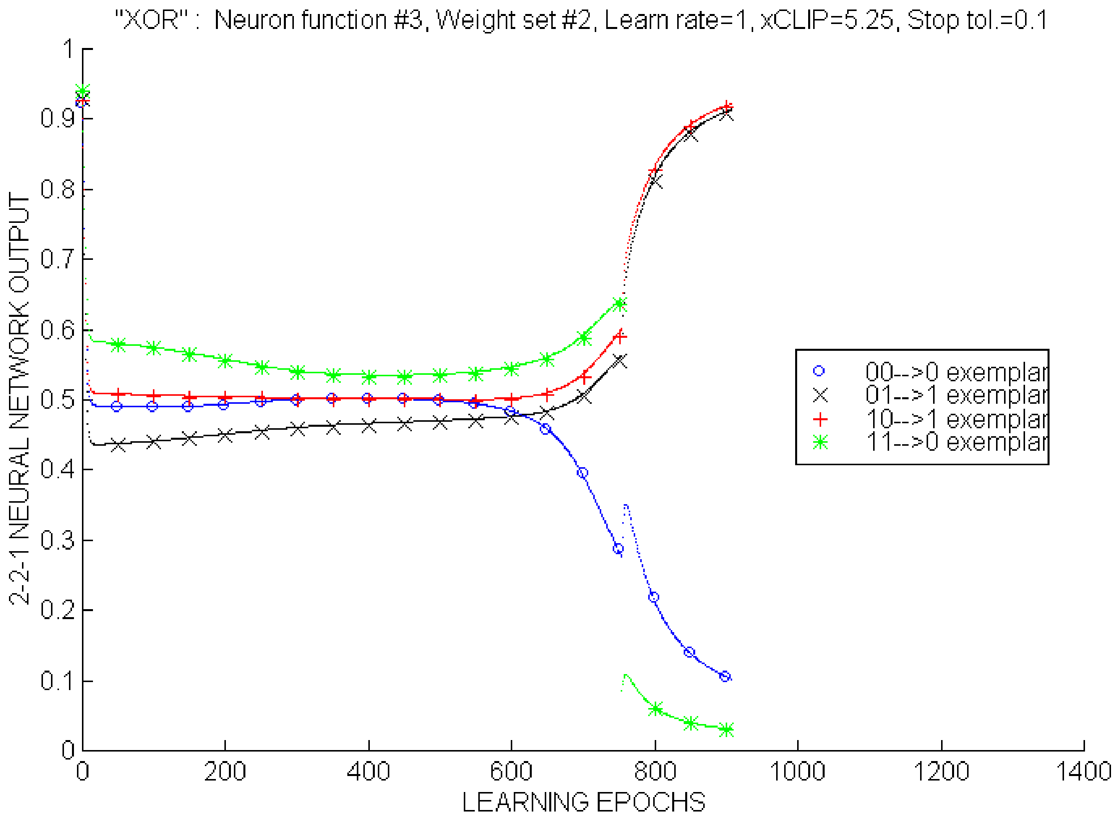
1991 J. Wunderlich Master's Thesis and Patent Disclosure document

Backpropagation learning "XOR" using 10th degree Taylor polynomial approximation of e^{-x} in sigmoid: $1/(1+e^{-x})$. Expanded about $x_0 = 0$, and "clipped" outside of domain -3.5 >X <3.5



1991 J. Wunderlich Master's Thesis and Patent Disclosure document

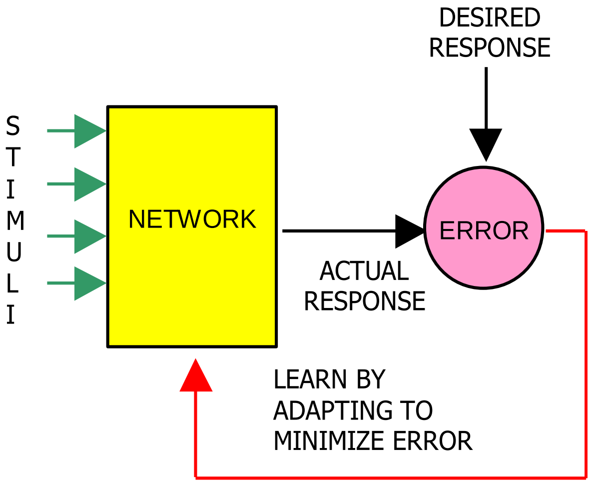
Backpropagation learning “XOR” using 12th degree Divided-Difference polynomial approximation of sigmoid, and "clipping" outside of domain -5.25 >X <5.25 to demonstrate robustness of methodology (i.e., 12th degree poly has same accuracy for -10 > X >10)

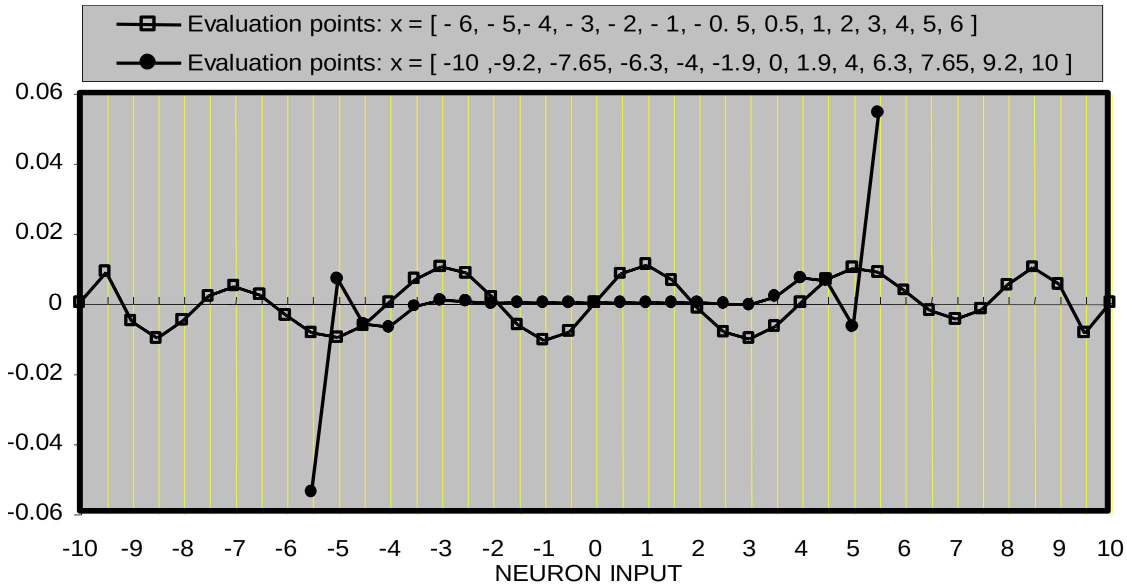


1991 J. Wunderlich Master’s Thesis and Patent Disclosure document

Top-down Design

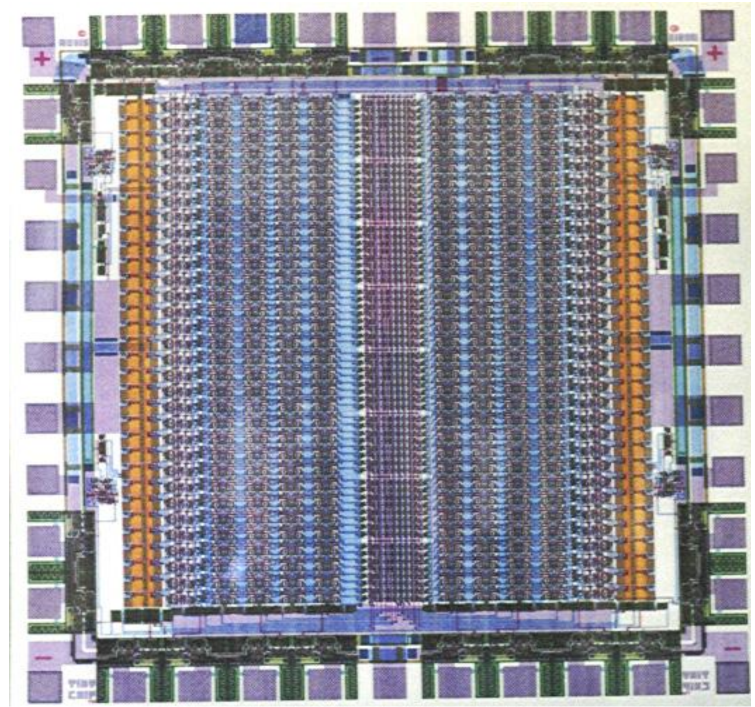
- This top-down designs is:
  - **Vector-register**
  - **Entirely digital**
  - **Fully parallel** by using a polynomial approximation of sigmoid to allow parallel **on-chip learning**
  - (U.S. Patent Disclosure recorded in 1991 for J. Wunderlich design)
  - Transistor-count estimate to implement 100 Neurons is 1,000,000

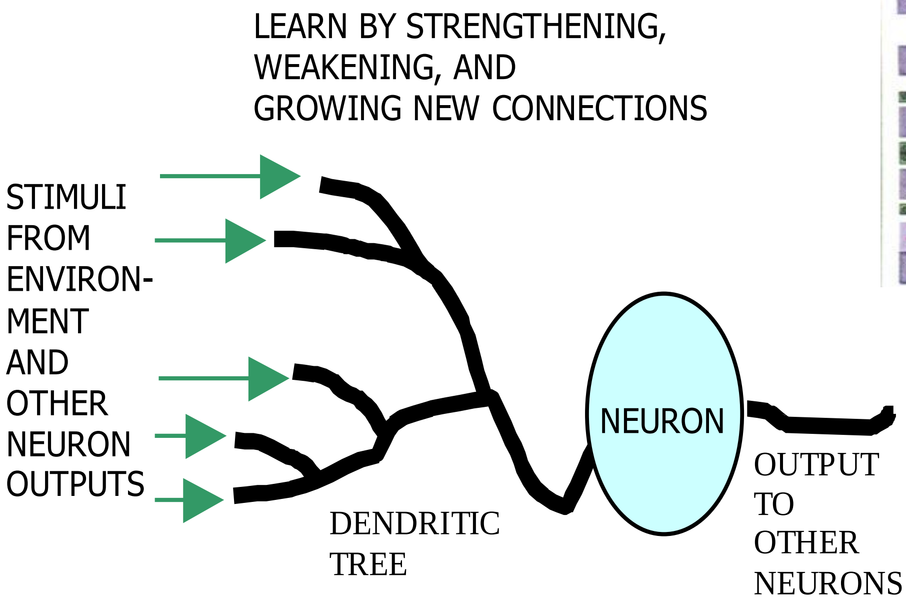




1991 J. Wunderlich Master's Thesis and Patent Disclosure document

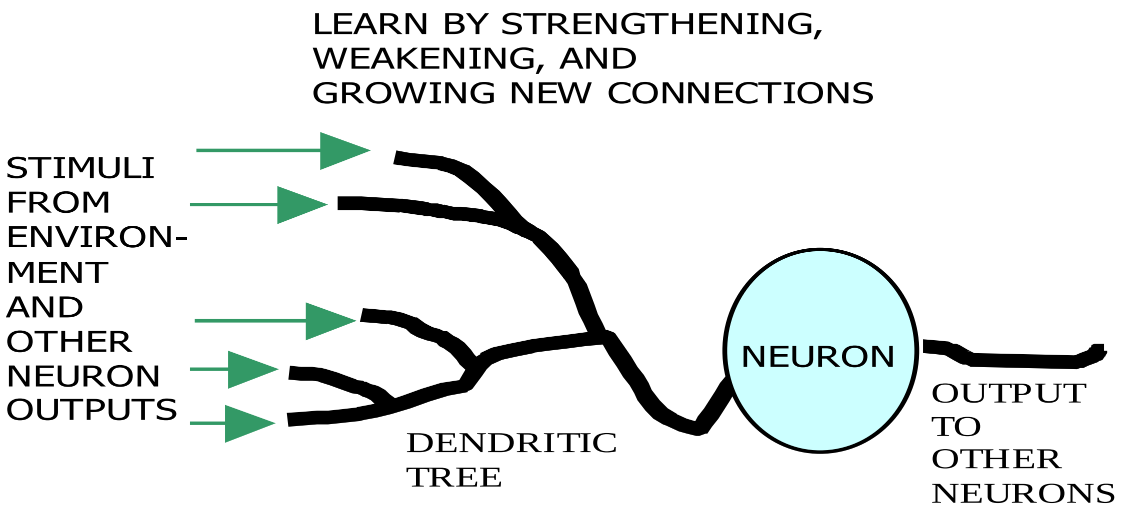
“Bottom-up” neurocomputer
(artificial dendritic tree VLSI chip)
by Wunderlich and Elias, 1992

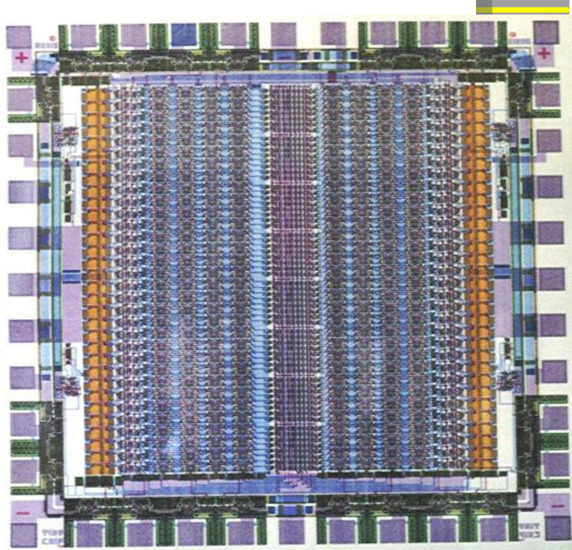




Bottom-up
more difficult to implement

- *Useful* bottom-up designs difficult to realize
  - 3D biological brains vs. 2D IC's
    - Connection density problems
  - In biological learning, connections not only strengthened or weakened, but are **grown**





Read more at:

- Wunderlich, J.T. (2004). **Top-down vs. bottom-up neurocomputer design**. In *Intelligent Engineering Systems through Artificial Neural Networks, Proceedings of ANNIE 2004 International Conference, St. Louis, MO.* H. Dagli (Ed.): Vol. 14. (pp. 855-866). ASME Press.

- Wunderlich, J.T. (20xx). **Two single-chip neurocomputer designs; one bottom-up, one top-down**. *Draft of a book chapter*

WUNDERLICH FUTURE RESEARCH

- Merging bottom-up and top-down techniques for pre-processing sensory data
  - Visual
  - Auditory
  - Olfactory
- And using top-down techniques for higher reasoning
  - including combining Neural Networks with symbolic AI programming
- Add machine *intelligence* into robots ....

Mid-1990's J. Wunderlich IBM Research

IBM S/390 Hardware Development Lab, Researcher & Hardware Development Engineer (Advisory-Level)

- Helped develop Symmetric Multi-Processor (SMP) mainframe-supercomputer architectures (jointly developed with IBM Germany) by engineering systems-level software and part of an operating system (SAK) to "stress" features and force hardware failures through pseudo-random generation of correlated machine states and operating scenarios.
- Machines included 20 CPU's divisible into 15 logical partitions and scalable to 512 processors via a dynamic optical interconnect (IBM Parallel Sysplex).
- Engineered software to run in three environments: VLSI circuit simulation, initial hardware test, and manufacturing.
- New 64-bit processing (address and data) required simulating 64-bit arithmetic and virtual-addressing to test simulated 64-bit prototype architectures using 32-bit machines. Prototypes were released as "IBM eServer zSeries" (now called zEnterprise).
- My research included microprocessor branch-prediction verification strategies in a multiprocessor environment; and theory for hardware verification with seven correlated random number generators. My development projects included writing 20,000 lines of high-level language (PL/X) and S/390 assembly code including operating system application interfaces (API's). My RNG API code was also translated into C for IBM AS/400 minicomputers and RS-6000 (AIX type UNIX) workstations (the predecessor of POWER7 supercomputers like "Watson") requiring my supervising an engineer in Austin TX via the IBM intranet.
- Other projects included verification programs for cache coherency, virtual addressing, space-switching, linkage control, and 125 new IEEE floating-point instructions (to supplement IBM Hex floating-point). All 1400 IBM S/390 machine instructions were tested (including vector-register instructions for add-on vector-register unit). A patent process was initiated for my random number theory and API's.

2011 IBM Watson

- Watch four-minute video: https://www.youtube.com/watch?v=WFR3lOm_xhE
- J. Wunderlich related IBM Research in the mid-1990's
  - IBM S/390 supercomputer research (New York) ported to IBM RS6000 workstations (Austin, Texas) – predecessor to POWER7 that **Watson** runs on
    - Supervised an Austin Texas Engineer via the IBM Intranet
- In 2011 Watson was a Special-Purpose Machine built to play **Jeopardy**
  - Like **IBM "Deep-Blue" --** Special-Purpose Machine built to play **Chess** that beat world-champion Garry Kasparov in 1996
    - An IBM SP2 MPP Supercomputer built by the IBM "Power- Parallel" group in the same center as the IBM S/390 SMP Supercomputer Development Lab that Dr Wunderlich worked in. The two groups often shared in activities.
- Present applications for Watson include Cloud computing, Healthcare, Education, and Weather Foresting

2011 IBM Watson

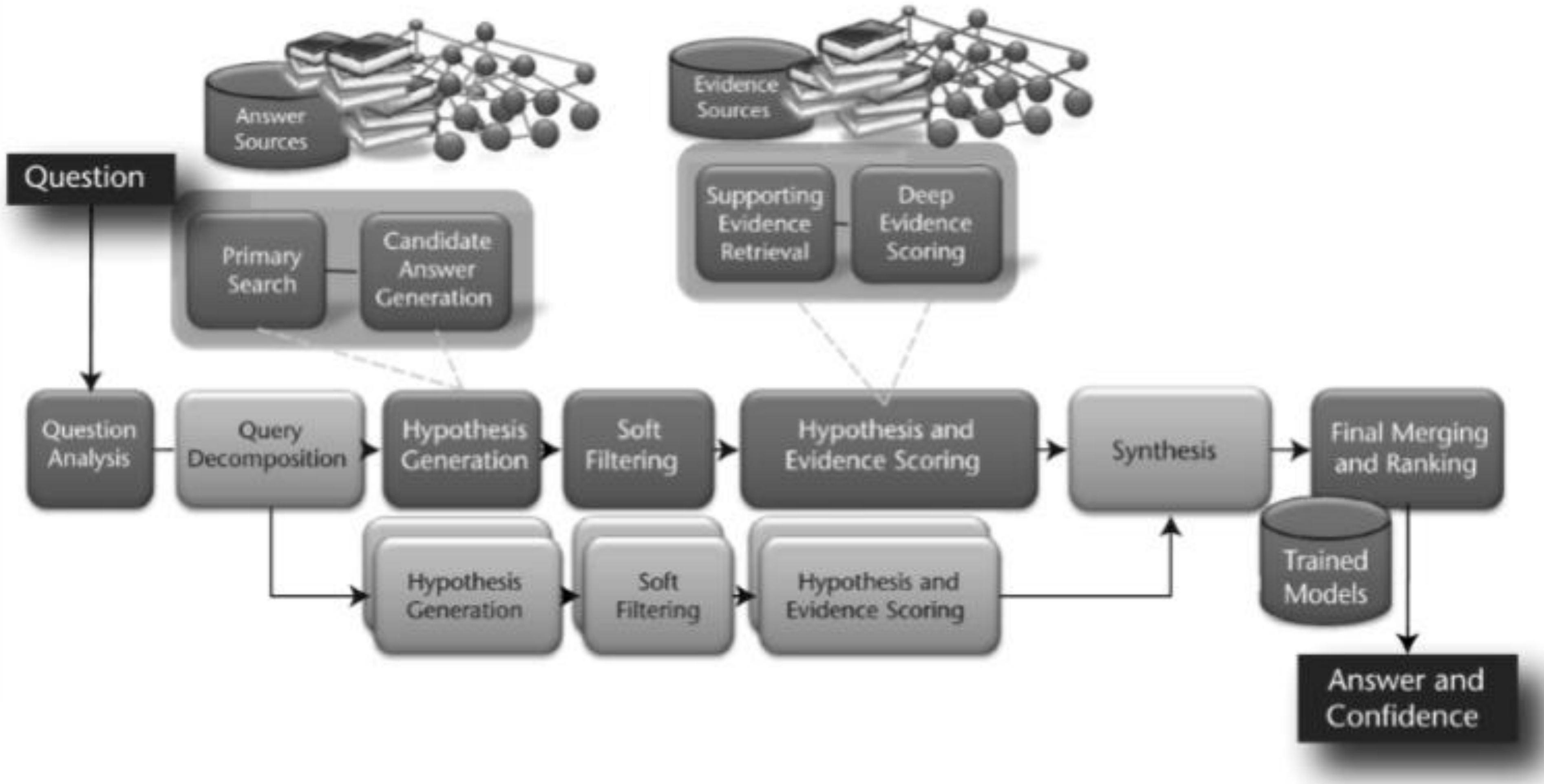
- Natural Language Processing
  - Understanding Context
    - **Disambiguating language** (understanding *which* meaning of a word in a sentence)
  - Somewhat understanding puns and other kinds of wordplay
- Knowledge Representation
  - Problem Definition
  - **Pattern Matching**
- Data Mining
- **Confidence** and Probability Theory
- **Machine Learning** (adaptability)
- MPP (Massively Parallel Processing) hardware

- In 2011 Watson was **not connected to the Internet** when it played Jeopardy. It instead had 200 million pages of documents on four terabytes of disc space including an entire copy of Wikipedia; and for a short time the "Urban Dictionary" (removed because Watson was cursing)

2011 IBM Watson

From 2010 AI Magazine "*Building Watson: An Overview of the DeepQA Project*"

http://www.aaai.org/ojs/index.php/aimagazine/article/view/2303/2165



2011 IBM Watson

From 2010 AI Magazine "*Building Watson: An Overview of the DeepQA Project*"
http://www.aaai.org/ojs/index.php/aimagazine/article/view/2303/2165

Excluded Question Types.

The *Jeopardy* quiz show ordinarily admits two kinds of questions that IBM and Jeopardy Productions, Inc., agreed to exclude from the computer contest: audiovisual (A/V) questions and Special Instructions questions. A/V questions require listening to or watching some sort of audio, image, or video segment to determine a correct answer. For example:

Category: Picture This
(Contestants are shown a picture of a B-52 bomber)
Clue: Alphanumeric name of the fearsome machine seen here.
Answer: B-52

Special instruction questions are those that are not "self-explanatory" but rather require a verbal explanation describing how the question should be interpreted and solved. For example:

Category: Decode the Postal Codes
Verbal instruction from host: We're going to give you a word comprising two postal abbreviations; you have to identify the states.
Clue: Vain
Answer: Virginia and Indiana

2011 IBM Watson

From 2010 AI Magazine "*Building Watson: An Overview of the DeepQA Project*"
http://www.aaai.org/ojs/index.php/aimagazine/article/view/2303/2165

The Domain

As a measure of the *Jeopardy* Challenge's breadth of domain, we analyzed a random sample of 20,000 questions extracting the lexical answer type (LAT) when present. We define a LAT to be a word in the clue that indicates the type of the answer, independent of assigning semantics to that word. For example in the following clue, the LAT is the string "maneuver."

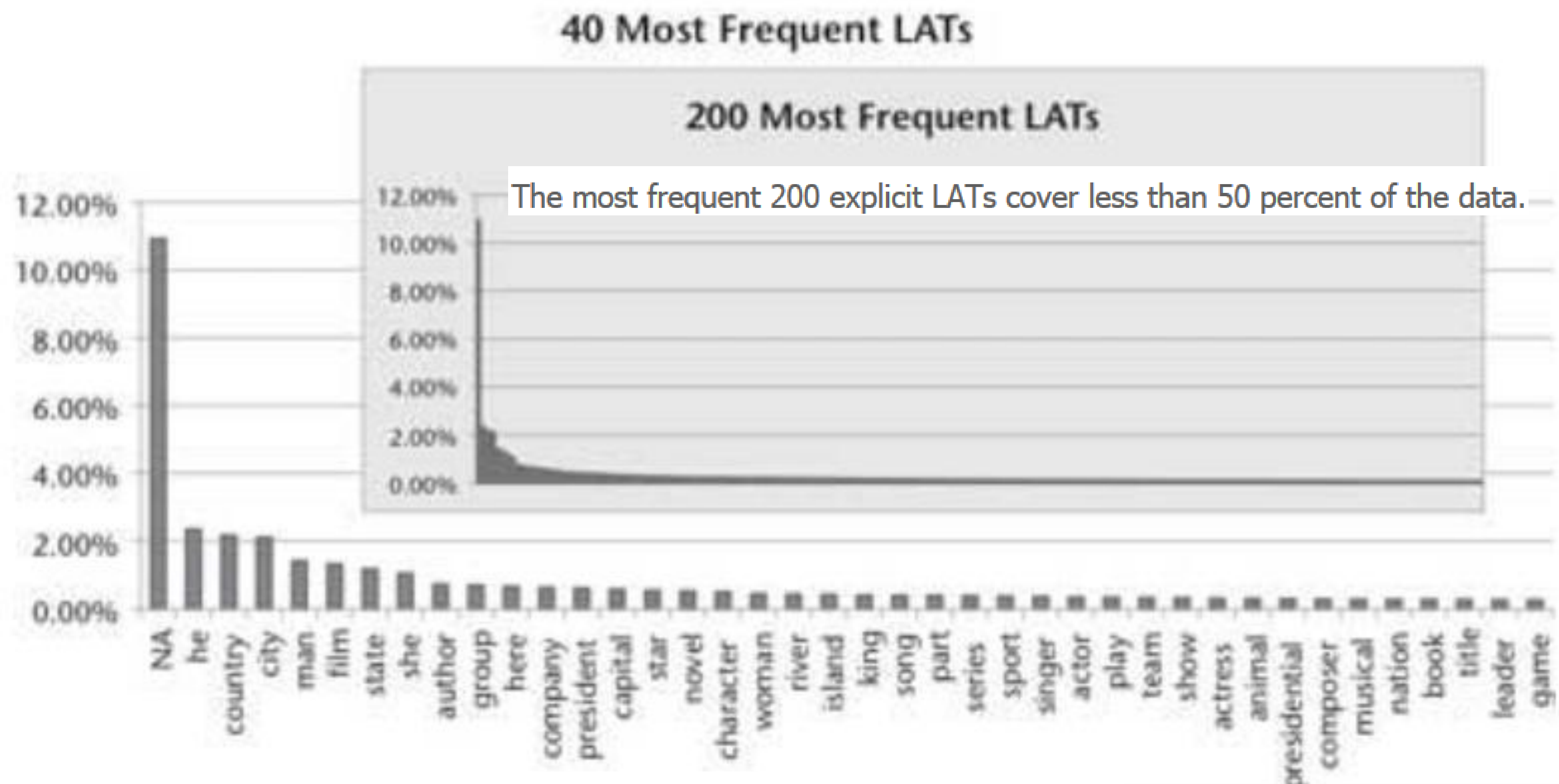
Category: Oooh….Chess
Clue: Invented in the 1500s to speed up the game, this maneuver involves two pieces of the same color.
7 *Answer:* Castling

About 12 percent of the clues do not indicate an explicit lexical answer type but may refer to the answer with pronouns like "it," "these," or "this" or not refer to it at all. In these cases the type of answer must be inferred by the context. Here's an example:

Category: Decorating
Clue: Though it sounds "harsh," it's just embroidery, often in a floral pattern, done with yarn on cotton cloth.
Answer: crewel

2011 IBM Watson

From 2010 AI Magazine "*Building Watson: An Overview of the DeepQA Project*"
http://www.aaai.org/ojs/index.php/aimagazine/article/view/2303/2165



Designer's ***trying*** to make Watson not so Application Specific.

. Our clear technical bias for both business and scientific motivations is to create general-purpose, reusable natural language processing (NLP) and knowledge representation and reasoning (KRR) technology that can exploit as-is natural language resources and as-is structured knowledge rather than to curate task-specific knowledge resources.

2011 IBM Watson

From 2010 AI Magazine "*Building Watson: An Overview of the **DeepQA** Project"*

http://www.aaai.org/ojs/index.php/aimagazine/article/view/2303/2165

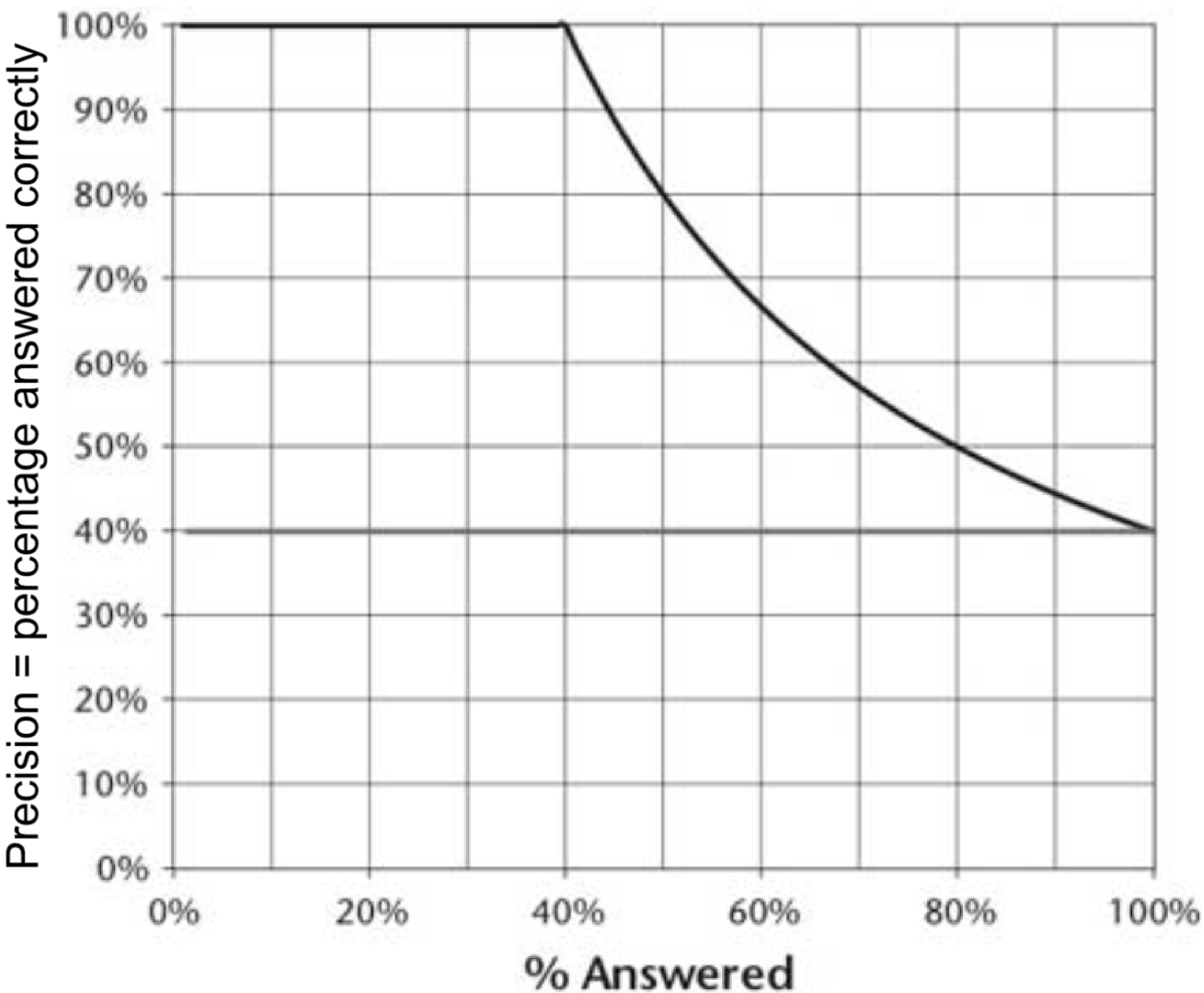


Figure 2. Precision Versus Percentage Attempted.
Perfect confidence estimation (upper line) and no confidence estimation (lower line).

*Programmers could "**Tune**" this to be more or less Aggressive in % attempted vs Precision to compete with the known ratio for best Jeopardy players*

Watson would eventually TUNE ITSELF

2011 IBM Watson

From 2010 AI Magazine "*Building Watson: An Overview of the DeepQA Project*"

http://www.aaai.org/ojs/index.php/aimagazine/article/view/2303/2165

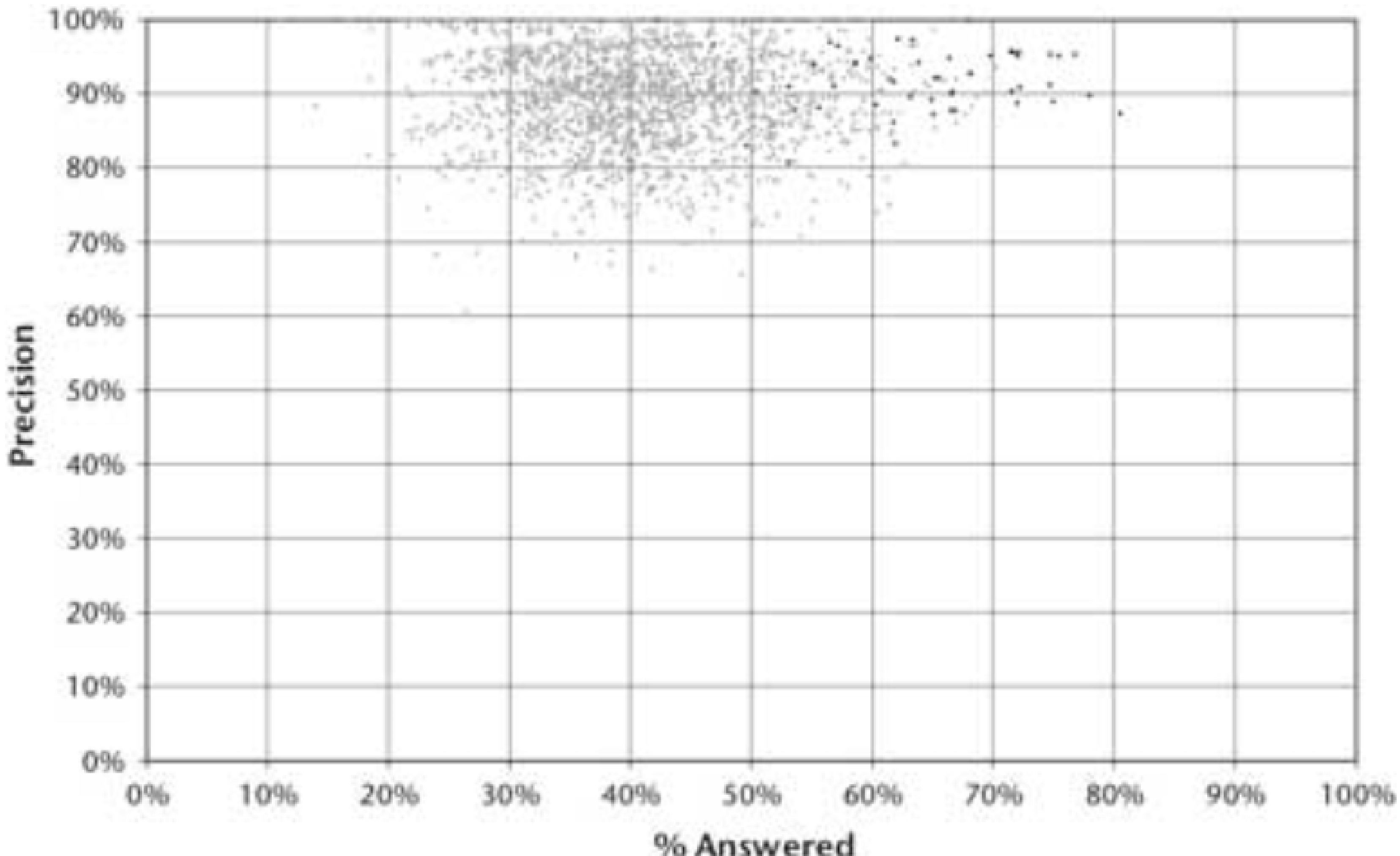
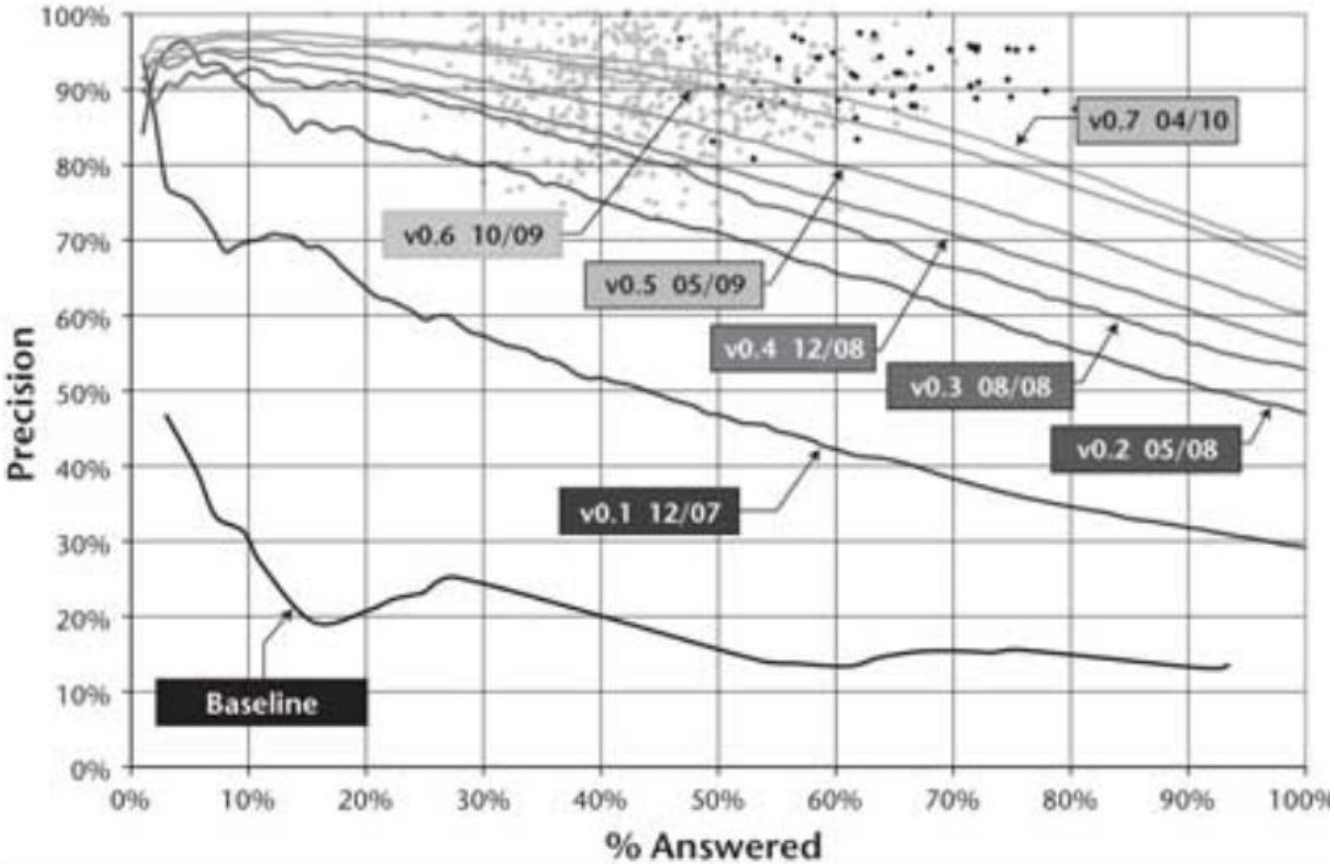


Figure 3. Champion Human Performance at Jeopardy.

2011 IBM Watson

From 2010 AI Magazine "*Building Watson: An Overview of the DeepQA Project*"
http://www.aaai.org/ojs/index.php/aimagazine/article/view/2303/2165



*Watson's performance, and therefore "**Confidences**" increased over time*

After approximately 3 years of effort by a core algorithmic team composed of 20 researchers and software engineers with a range of backgrounds in natural language processing, information retrieval, machine learning, computational linguistics, and knowledge representation and reasoning, we have driven the performance of DeepQA to operate within the winner's cloud on the *Jeopardy* task, as shown in figure 9. Watson's results illustrated in this figure were measured over blind test sets containing more than 2000 *Jeopardy* questions.

2011 IBM Watson

From ***"Final Jeopardy,"*** by Stephen Baker, **2012**, Mariner Books Publishing:

- Some initial problems:
  - Developed a **speech defect** of adding "D' to words ending in "N"; like "What is Pakistan**d**"
  - **<u>No-common-sense</u> wagering** on a "Daily Double" -- like when it bet only $5 when loosing $12,400 to $6,700 because one heuristics (rule) told it to not bet much if it had close to only half as much as opponent; The reasoning of the rule was to have enough to catch up in "Final Jeopardy" where contestants all wager before given a final question
- Watson could **<u>build overall confidence</u>** *(and therefore increase it's aggressiveness of play)* if it had just raced through an entire category
- Watson works best with hard-facts unencumbered by **<u>humor</u>, <u>slang</u>, or <u>cultural references</u>**
- Watson, like a Google search, **can't make <u>inductive leaps</u>** like Charles Darwin

2011 IBM Watson

From ***"Smart Machines, IBM's Watson and the Era of Cognitive Computing,"***

by John E. Kelly III and Steve Hamm **2013**, Columbia University Press:

- Watson will eventually interpret **images**, numbers (it had problems with Roman Numerals), **voices**, and other **<u>sensory information</u>**
  - ***Neural Networks are well-suited for this***
    - ***Do preprocessing, then feed to a cognitive core-brain***

- "**<u>Big Data</u>**"

  The digital Universe is growing ~60% per year with **social media**, **sensor networks, and huge warehouses of business, scientific, and government records on-line**
  - **<u>Coevolution of Computer Science and Medicine</u>**
    - Billions of combinations of variables in **human genome** results in **<u>15 to 20% of medical diagnoses inaccurate or incomplete</u>**
  - **<u>Urban Design & Planning</u>**
    - Requires understanding inner workings of a city
    - **<u>Human navigation</u>**: At a busy intersection we instantly identify people, vehicles, buildings, streets, and sidewalks; and see how they interrelate

2011 IBM Watson

From ***"The Second Machine Age,"*** by Erik Brynjolfsson and Andrew McAfee **2014**, W Norton & Son Publishing:

- "**Dr. Watson**" **matching** peer-reviewed medical literature to **patient symptoms, medical histories, and test results to formulate diagnosis and treatment**
  - Would take human 160 hours per week to do Watson's reading of Medical literature
  - IBM partners with Memorial Sloan-Kettering Cancer Center
  - Watson **augments a physician's clinical expertise and judgment**
- Watson **not good at "Thinking outside the Box"** (**Ideation, Creativity, Innovation**) … it would be lousy at playing Wheel of Fortune
- **Humans needed for idiosyncrasies and special cases** .. think about the risks of a 100% driverless car

2011 IBM Watson

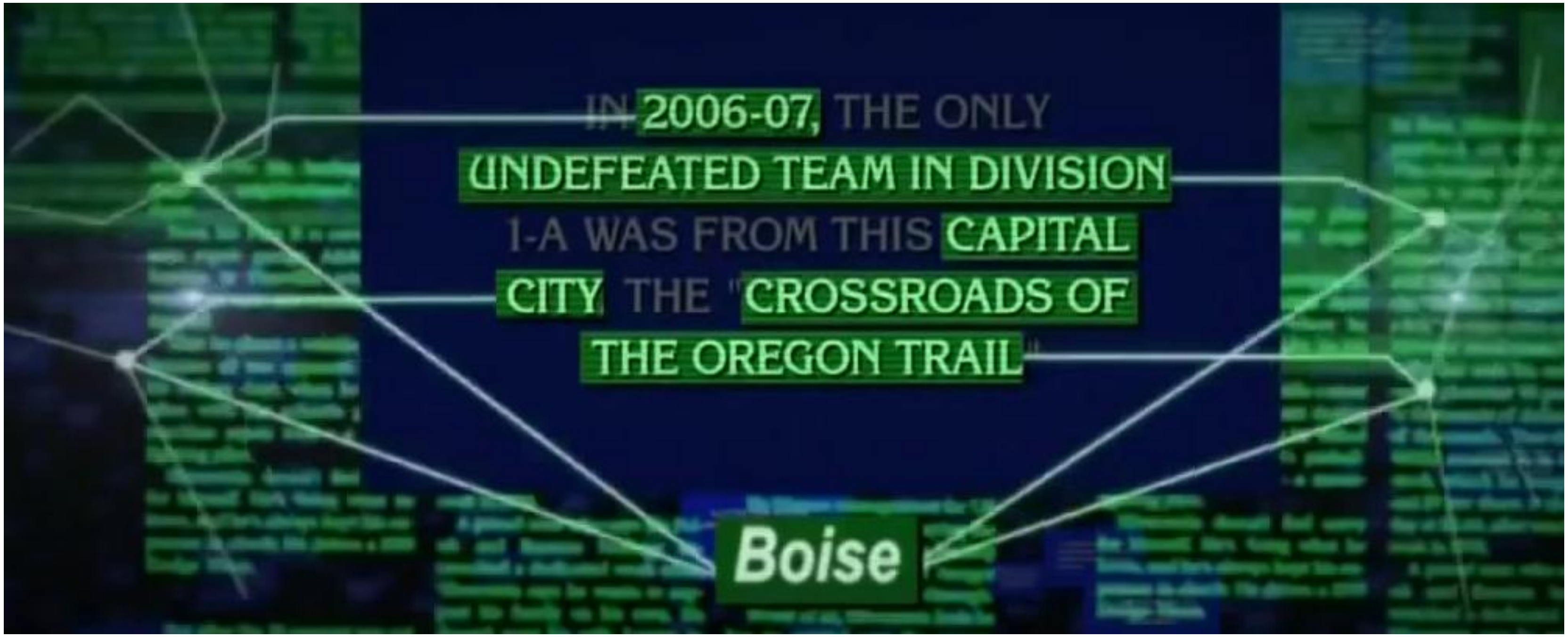
From: ***"IBM Watson: Smartest machine ever built,"*** **2015**, PBS NOVA episode, https://www.youtube.com/watch?v=3zQI-LMcDnA

- Understanding a jeopardy question is difficult (also understanding categories)
- HCI **H**uman **C**ompute **I**nteraction is difficult
  - 100's of practice games with humans
- Parsing sentences to find correct meaning of a **double-meaning** sentence
- Jeopardy different than **well-defined rules of Chess**; also:
  - **Humans play chess more conceptually** (Control center, flank opponent)
  - IBM Deep Blue just did exhaustive search of all possibilities
- Compared to six million rules for human **Common Sense** in the software "Psych," IBM wanted Watson to be more flexible
- **2800 CPU's**
- **Disc Storage** included Bible, World Book Encyclopedia, all of Wikipedia, much of New York Times archive, the internet movie database, many books, plays, etc
- Since deaf (receives questions by text), **couldn't initially hear other answers**
- Unlikely to understand **overall meaning** in plays, parables, etc

2011 IBM Watson

From: ***"IBM Watson: Smartest machine ever built,"*** **2015**, PBS NOVA episode, https://www.youtube.com/watch?v=3zQI-LMcDnA

- Watson giving 10's of thousands of old jeopardy questions with correct answers
  - Watson looks for patterns
  - Then Watson looks for supporting evidence



2011 IBM Watson

From: ***"IBM Watson: Smartest machine ever built,"* 2015**, PBS NOVA episode,
https://www.youtube.com/watch?v=3zQI-LMcDnA

- Then weigh the evidence, on average, and calculate a **confidence** for all possible answers

This original version of Watson was an advanced example of:

1) ***Natural Language Processing***
2) ***Statistical Analysis***

Watson, competing on the game show Jeopardy. The bars at the bottom show its confidence in each answer. If no answer passes the confidence threshold (the white line), Watson doesn't respond.

From 2015 ARS TECHNICA: ***"Debugging the Myths about Artificial Intelligence"***
http://arstechnica.com/information-technology/2015/12/demystifying-artificial-intelligence-no-the-singularity-is-not-just-around-the-corner/

2014 IBM Watson

VIDEO: **"IBM Watson: How it Works,"** IBM: https://www.youtube.com/watch?v=_Xcmh1LQB9I

CONCEPTS:

- Observation/Evaluation/Decision-Making
- Unstructured Data (80% of the current Data on Earth)
- Natural Language Processing
- Context
- Intent
- Inferences

METHODOLOGY:

- "Corpos" body of relevant literature
- Curate Content
- "Ingestion" preprocessing (indexing & organizing)
- Machine Learning
  - "QuestionAnswer" pairs (by experts) for "Ground Truth"
- Continuous learning
- Evidence-based recommendations
- Yield of new inferences and patterns
- Hypothesis' generation / evidence search / confidence
  - From weighted evidence scores
- Data Analytics to glean insights
  - Create inspirations for Human Experts to augment their decisions

2015 IBM Watson

Reference: ***"IBM Pushes Deep Learning with a Watson Upgrade, "*** **2015**, MIT Technology Review, https://www.technologyreview.com/s/539226/ibm-pushes-deep-learning-with-a-watson-upgrade/

"Deep learning involves training a computer to recognize often complex and abstract patterns by feeding large amounts of data through successive networks of artificial neurons, and refining the way those networks respond to the input "

"Combining disparate strands of AI research could become an important trend in coming years"

"Applying learning from one area, such as vision, to another, such as speech, is known as a multimodal approach. It could make future AI systems far more useful and could yield fundamental insights into the nature of intelligence."

2016 IBM Watson

Watch all of this Oxford University Video:

https://www.youtube.com/watch?v=rXVoRyIGGhU

2009 Asimo's "Intelligence"

Paraphrase:

*"**Intelligence is a technology and a strategy** for robust and flexible problem solving in complex environments under constraints…. designed for autonomous adaptation to a changing world……rapidly and stably self-organize successful behaviors in response to an unpredictably changing world.*

Brain computing is not in local processing or learning algorithm but a way the brain organizes ……rather than modeling isolated subsystems, large-scale computational models of complete functional blocks at interacting levels of complexity have to be investigated. "

Prof. Dr. Edgar Koerner
Honda Research Institute Europe GmbH

SOURCE: http://www.jsps-bonn.de/fileadmin/veranstaltungen/2009_Abstr_Koerner.pdf

2009 Asimo's "Intelligence"

(1) Control of growth by gene-regulatory Networks
(2) Cortical architectures for storing experience
(3) Behavior-based allocation of resources for predictive visual scene analysis
(4) Global behavior control for autonomous interaction

Nested control loops for reflexes, attention, learning from sensory experience, and predictions for interaction.

> ***"This enables ASIMO to learn to recognize objects through interaction with humans, to learn associations between acoustic and visual objects, as well as to associate sound with behavioral concepts for interaction, and to demonstrate first steps of prediction driven behavior."***

Prof. Dr. Edgar Koerner
Honda Research Institute Europe GmbH

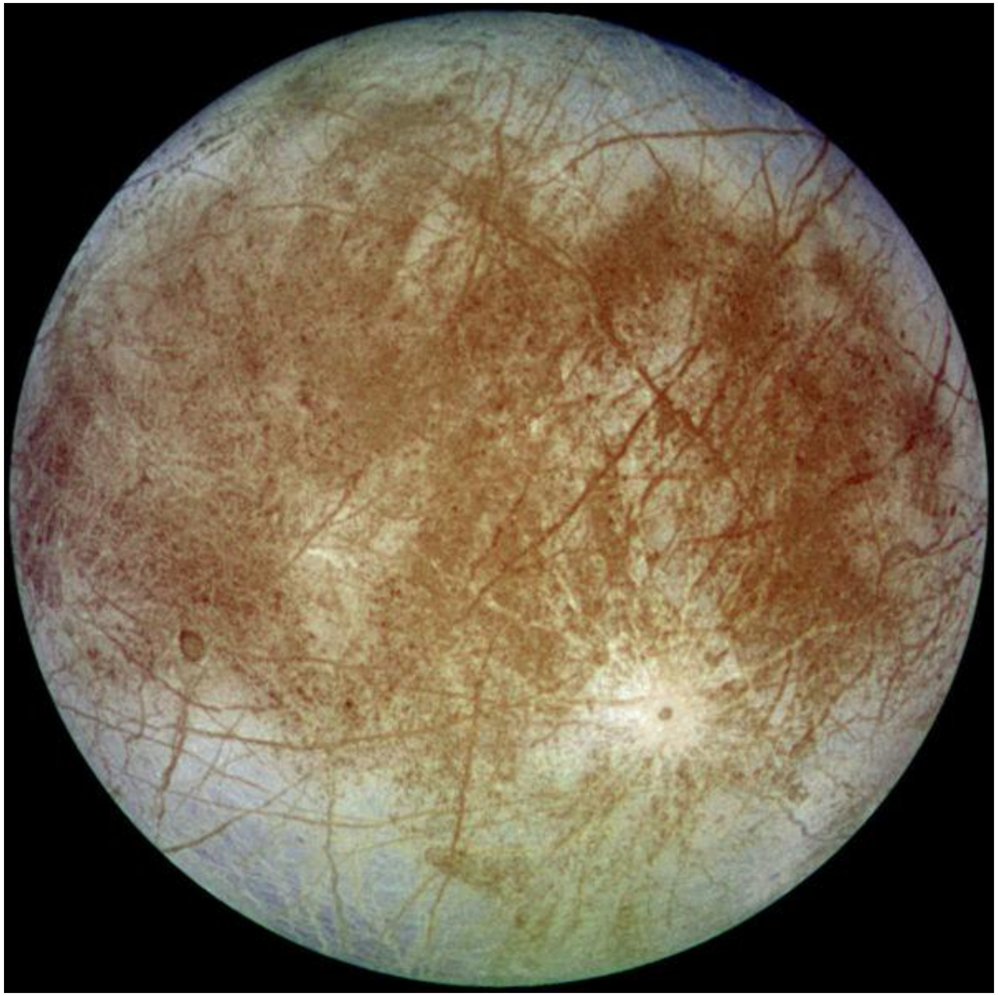
SOURCE: http://www.jsps-bonn.de/fileadmin/veranstaltungen/2009_Abstr_Koerner.pdf

Future space missions?

Much Intelligence may likely be needed

VIDEO: http://www.youtube.com/watch?v=Q3C5sc8b3xM

Europa Rover

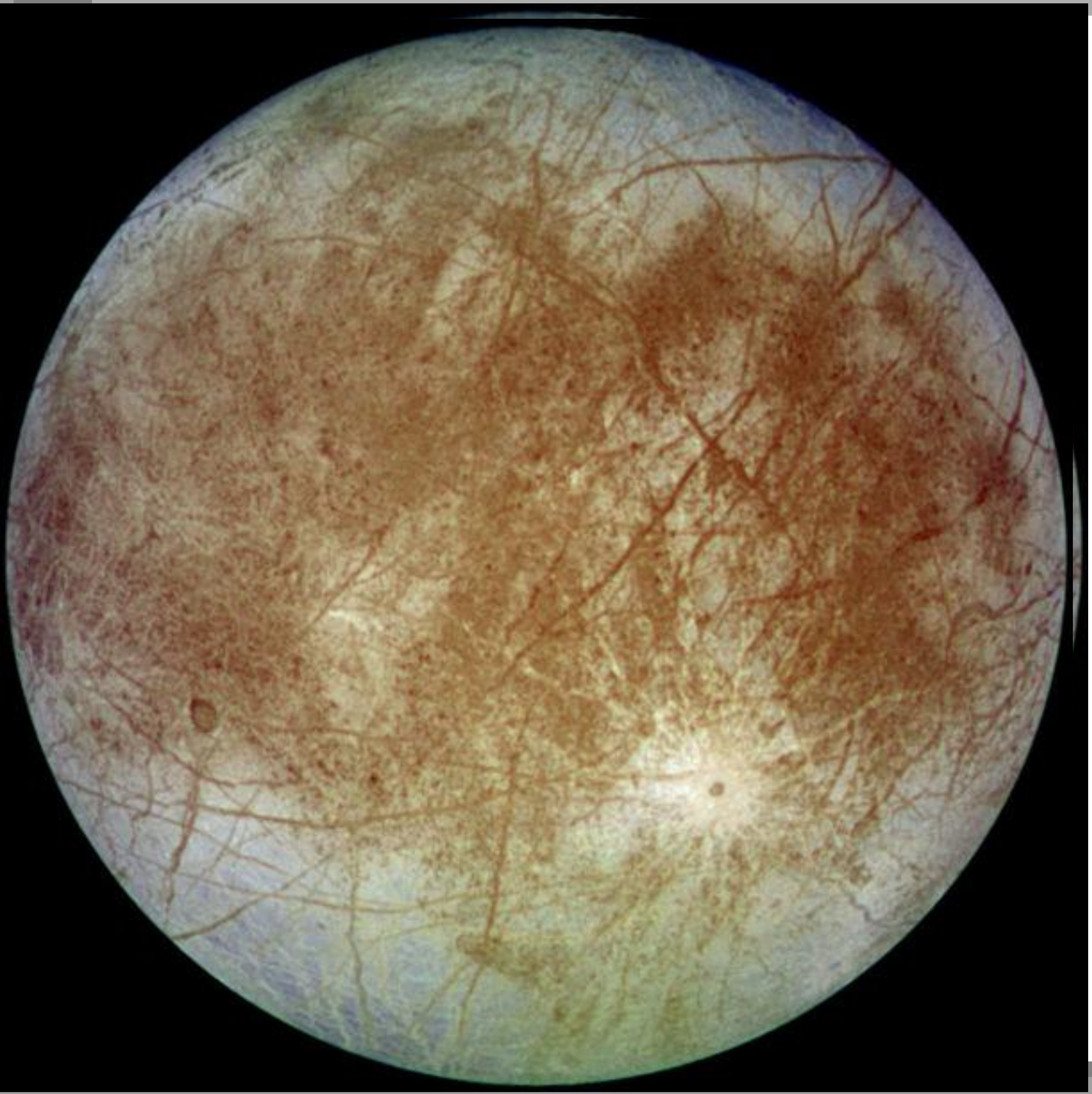


Europa is a moon of Jupiter covered in ice

Future NASA/ESA mission is to get to liquid ocean beneath the ice

Image from: http://photojournal.jpl.nasa.gov/catalog/PIA00502

Europa Rover



Europa
Is much further from Earth than Mars

And therefore much more <u>autonomous</u> robotics will be needed

Image from: http://photojournal.jpl.nasa.gov/catalog/PIA00502

Europa (Cracks in the ice – and an ocean below)

Image from: http://www.mapaplanet.org/explorer/help/data_set.html

Wunderlich, J.T. (2011). Designing robot autonomy: how tightly should we hold the leash? in proceedings of *The 5th Int'l Conference on Design Principles and Practices,* Rome, Italy.

PAPER: http://users.etown.edu/w/wunderjt/ROME_2011_PAPER9_434_Reading.pdf
TALK: http://users.etown.edu/w/wunderjt/ROME_2011_10.pdf