Introduction
Neural Networks
Representation
Multiple Tasks
SMRL
Description
Experiments
Analysis
Optimizations
Discussion
Questions

Sequentially Learning Multiple Tasks in Artificial Neural Networks

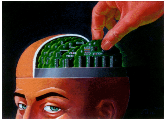
Preventing Catastrophic Interference with Meaningful Representations

Jordi Bieger

April 15, 2009

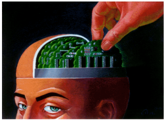
Programme

1. Introduction
   - Neural Networks
   - Representation
   - Multiple Tasks
2. Static Meaningful Representation Learning
   - Description
   - Experiments
   - Analysis
   - Optimizations
3. Discussion
4. Questions



Programme

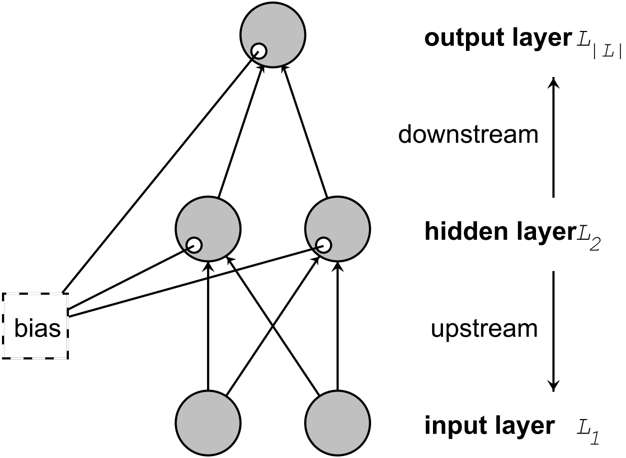
1. Introduction
   - Neural Networks
   - Representation
   - Multiple Tasks
2. Static Meaningful Representation Learning
   - Description
   - Experiments
   - Analysis
   - Optimizations
3. Discussion
4. Questions



Introduction

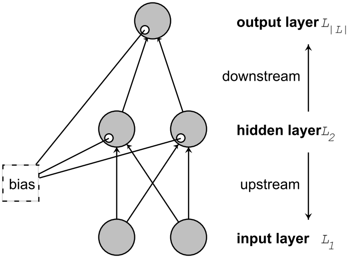
- Artificial Neural Networks (ANNs) are loosely based on neural mechanisms in the brain.
- Standard Multi-Layer Perceptrons (MLPs) fail to model the brain's ability to sequentially learn multiple tasks.
- I propose a simple solution called "Static Meaningful Representation Learning".

Multi-Layer Perceptrons

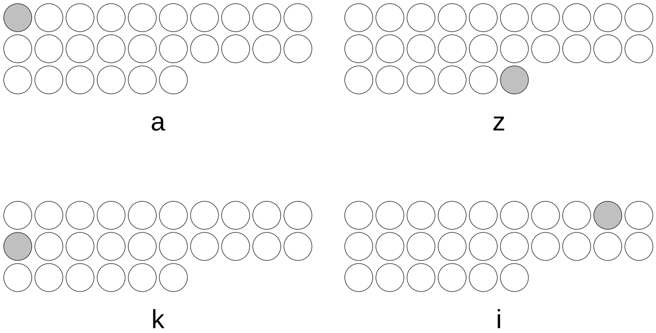


Representation

- Situation needs to be encoded into a representation the network “understands”.
- These representations are often arbitrary.

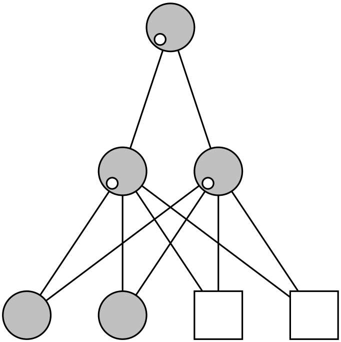


Representation



Multiple tasks

- Network needs to be told what task to do.
- Accomplished by adding extra task representation nodes.
- Action words have task relevant representations in human brains.
- Sequential learning causes catatrophic interference.



Tasks

Inputs		NONE	AND	NIF	1ST	JUST2	2ND	XOR	OR	NOR	IFF	¬2ND	¬JUST2	¬1ST	IF	NAND	ALL
−	−	−	−	−	−	−	−	−	−	+	+	+	+	+	+	+	+
−	+	−	−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
+	−	−	−	+	+	−	−	+	+	−	−	+	+	−	−	+	+
+	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+	−	+
binary		0000	0001	0010	0011	0100	0101	0110	0111	1000	1001	1010	1011	1100	1101	1110	1111
#		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

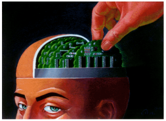
Programme

1 Introduction
- Neural Networks
- Representation
- Multiple Tasks

2 Static Meaningful Representation Learning
- Description
- Experiments
- Analysis
- Optimizations

3 Discussion

4 Questions



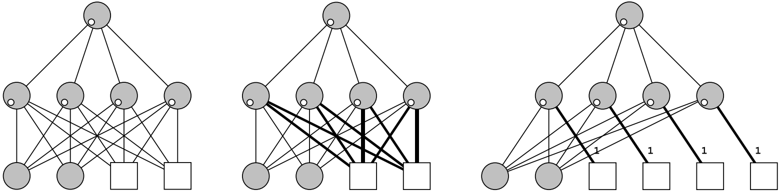
Introduction
Neural Networks
Representation
Multiple Tasks
SMRL
Description
Experiments
Analysis
Optimizations
Discussion
Questions

Static Meaningful Representation Learning

- Static Meaningful Representation Learning (SMRL)
  - Initial knowledge acquisition phase
  - Fix all weights in the network
  - Novelty learning phase
- Uses Parametric Bias (PB) nodes for learning meaningful task representations

Network Types

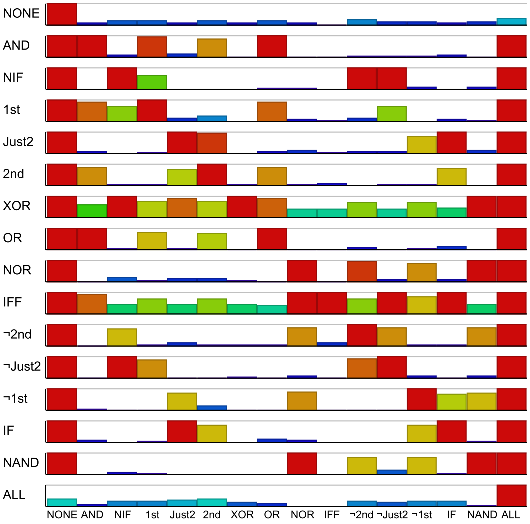
Implicit PB | Fixed Weight IPB | Explicit PB



Results

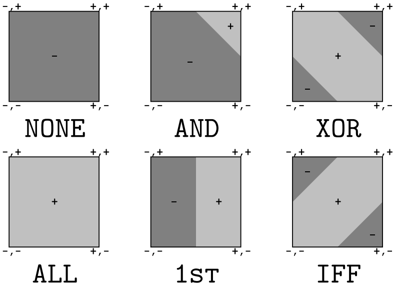
# Hidden Nodes	# PB Nodes	IPB	FWIPB	EPB
2	1	24.9%	26.4%	n/a
4	1	25.6%	26.0%	n/a
2	2	30.4%	31.6%	32.2%
4	2	36.3%	36.7%	n/a
6	2	38.6%	40.5%	n/a
6	6	-	-	54.9%
2+4	2	30.8%	31.6%	32.2%
4+4	2	34.0%	35.7%	n/a
4+4	4	-	-	38.7%
4	4	41.8%	42.3%	42.1%

Results


NONE
AND
NIF
1st
Just2
2nd
XOR
OR
NOR
IFF
¬2nd
¬Just2
¬1st
IF
NAND
ALL
NONE AND NIF 1st Just2 2nd XOR OR NOR IFF ¬2nd ¬Just2 ¬1st IF NAND ALL

Difficulty

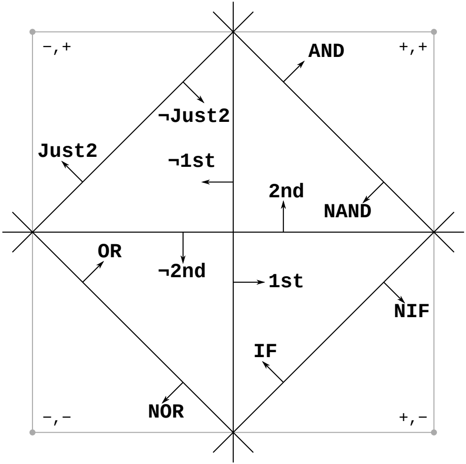
Network has to be “smart” enough.



Similarity

Similarity depends on effect of inputs on target outputs

- Parallel: arrows in the same direction
- Similar: arrows in roughly the same direction ($< 90^\circ$)



Results

Network Type	# Hidden Nodes	# PB Nodes	Difficulty	Parallel	Similarity	Prodigy	Overall
IPB	4	2	77	98	72	66	36
IPB	4	4	88	100	79	79	42
FWIPB	4	2	81	98	72	69	37
FWIPB	4	4	91	100	79	88	42
EPB	2	2	73	100	68	51	32
EPB	4	4	87	100	80	75	42

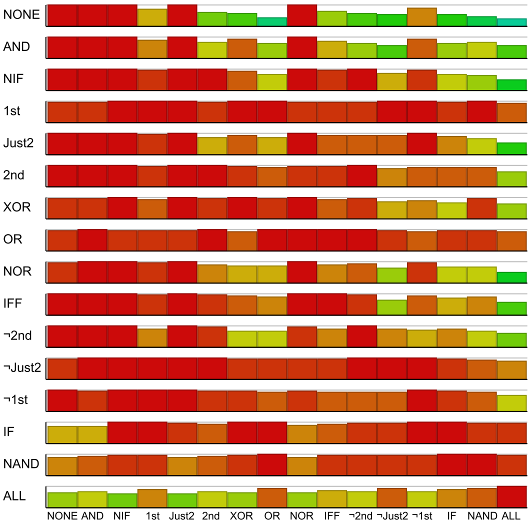
Arctangent

Network Type	# Hidden Nodes	# PB Nodes	Difficulty	Parallel	Similarity	Prodigy	Overall
IPB	4	2	83 (+6)	100 (+2)	70 (-2)	73 (+7)	37 (+1)
IPB	4	4	90 (+2)	100 (0)	75 (-4)	83 (+4)	40 (-1)
FWIPB	4	2	87 (+6)	100 (+1)	72 (0)	80 (+11)	38 (+1)
FWIPB	4	4	96 (+5)	100 (+0)	78 (-1)	93 (+5)	43 (+0)
EPB	2	2	77 (+4)	100 (0)	64 (-4)	57 (+6)	32 (0)
EPB	4	4	96 (+9)	100 (0)	80 (-1)	93 (+18)	44 (+2)

Gaussian

Network Type	# Hidden Nodes	# PB Nodes	Difficulty	Parallel	Similarity	Prodigy	Overall
IPB	4	2	74 (-3)	86 (-11)	78 (+6)	88 (+22)	73 (+37)
IPB	4	4	87 (-1)	90 (-10)	89 (+9)	97 (+18)	84 (+42)
FWIPB	4	2	74 (-6)	85 (-14)	78 (+6)	86 (+16)	74 (+37)
FWIPB	4	4	82 (-9)	91 (-9)	86 (+7)	92 (+4)	81 (+39)
EPB	2	2	65 (-8)	83 (-17)	69 (+1)	76 (+25)	60 (+28)
EPB	4	4	88 (+1)	95 (-5)	94 (+14)	97 (+22)	89 (+47)

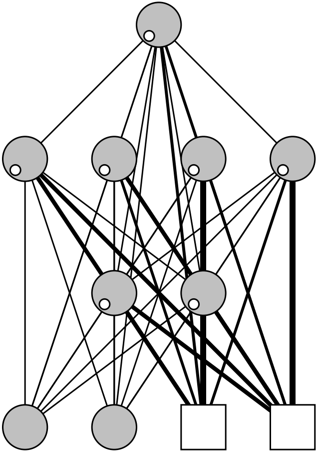
Gaussian



Sequentially Learning Multiple Tasks in Artificial Neural Networks
Jordi Bieger

Introduction
Neural Networks
Representation
Multiple Tasks
SMRL
Description
Experiments
Analysis
Optimizations
Discussion
Questions

Extra connections

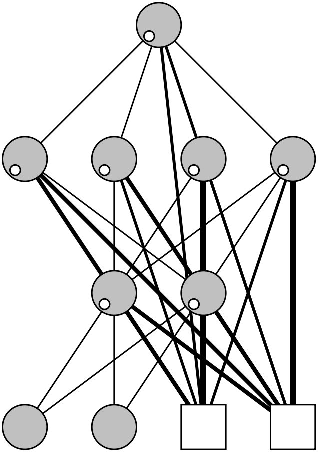


Introduction
Neural Networks
Representation
Multiple Tasks
SMRL
Description
Experiments
Analysis
Optimizations
Discussion
Questions

Extra connections

Network Type	# Hidden Nodes	# PB Nodes	Difficulty	Parallel	Similarity	Prodigy	Overall
IPB	4	2	83 (+6)	100 (+2)	70 (-2)	74 (+8)	39 (+3)
IPB	4+4	2	81 (+6)	99 (0)	71 (-3)	74 (+16)	39 (+5)
IPB	4	4	94 (+6)	100 (0)	74 (-5)	90 (+11)	43 (+1)
FWIPB	4	2	83 (+3)	100 (+1)	73 (+2)	73 (+4)	40 (+3)
FWIPB	4+4	2	85 (+5)	99 (+0)	76 (+0)	78 (+13)	41 (+5)
FWIPB	4	4	96 (+5)	100 (+0)	80 (+1)	93 (+5)	44 (+2)
EPB	2	3	84 (+11)	100 (0)	69 (+1)	70 (+19)	38 (+6)
EPB	4	5	91 (+4)	100 (0)	73 (-7)	85 (+10)	43 (+1)
EPB	4+4	9	95 (+9)	100 (0)	80 (-1)	91 (+19)	47 (+8)

Extra representation weights



Extra representation weights

Network Type	# Hidden Nodes	# PB Nodes	Difficulty	Parallel	Similarity	Prodigy	Overall
IPB	4	2	85 (+8)	98 (+0)	74 (+2)	78 (+12)	41 (+5)
IPB	4+4	2	80 (+5)	99 (-1)	74 (0)	70 (+12)	39 (+5)
IPB	4	4	97 (+9)	100 (0)	87 (+8)	97 (+18)	53 (+11)
FWIPB	4	2	84 (+4)	99 (+1)	77 (+5)	77 (+8)	41 (+5)
FWIPB	4+4	2	83 (+3)	99 (0)	78 (+3)	72 (+7)	40 (+5)
FWIPB	4	4	95 (+4)	100 (0)	90 (+11)	93 (+5)	49 (+7)
EPB	2	3	84 (+11)	100 (0)	74 (+6)	74 (+23)	41 (+9)
EPB	4	5	93 (+7)	100 (0)	89 (+9)	89 (+14)	53 (+11)
EPB	4+4	9	96 (+11)	100 (0)	92 (+11)	92 (+20)	59 (+20)

Programme

Sequentially Learning Multiple Tasks in Artificial Neural Networks

Jordi Bieger

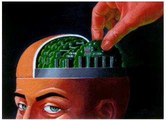
Introduction
- Neural Networks
- Representation
- Multiple Tasks

SMRL
- Description
- Experiments
- Analysis
- Optimizations

Discussion

Questions

1. Introduction
   - Neural Networks
   - Representation
   - Multiple Tasks
2. Static Meaningful Representation Learning
   - Description
   - Experiments
   - Analysis
   - Optimizations
3. Discussion
4. Questions



Introduction
Neural Networks
Representation
Multiple Tasks
SMRL
Description
Experiments
Analysis
Optimizations
Discussion
Questions

Evaluation

Advantages

- Simple
- Sequential learning of multiple tasks
- No catastrophic interference
- Meaningful representations

Disadvantages

- Weights don’t change
- Unproven for more interesting task domains

Conclusion

- Meaningful representations are powerful
- Combination with other techniques desirable
- Further research required

Introduction
Neural Networks
Representation
Multiple Tasks
SMRL
Description
Experiments
Analysis
Optimizations
Discussion
Questions

Programme

1 Introduction
- Neural Networks
- Representation
- Multiple Tasks

2 Static Meaningful Representation Learning
- Description
- Experiments
- Analysis
- Optimizations

3 Discussion

4 Questions