

# Intelligent Systems: Recognition and Reasoning

James L. Crowley

MoSIG M1  
Lecture 1

Second Semester 2020/2021  
2 Feb 2021

## Intelligence: Recognition and Reasoning

### Outline

The Science of Intelligent Systems.....	2
AI as a Scientific Discipline .....	2
Intelligence as a Description of Behavior: The Turing Test .....	3
Artificial Intelligence as a Scientific Domain .....	4
Intelligence as Knowledge and Reasoning. ....	5
Expert Systems .....	5
Kinds of Knowledge.....	6
Machine Learning .....	7
Artificial Neural Networks .....	8
Bayesian Approaches for Machine Learning .....	9
The Return of the Perceptron .....	10
Convolutional Neural Networks.....	10
Artificial Intelligence and Cognitive Science .....	12
Long Term Memory, Working Memory and Cognition .....	12
Course Overview .....	14
Grades.....	15

Class notes and exercises on the web:

<http://crowley-coutaz.fr/jlc/Courses/2020/MOSIG.SIRR/MoSIG.SIRR.html>

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 2

Winter Semester 2020-2021  
4 Feb 2021

## Learning and Evaluation for Pattern Recognition

### Outline

Notation .....	2
1. The Pattern Recognition Problem.....	3
Discriminant and Decision Functions.....	4
Machine Learning .....	5
Training and Validation.....	6
2. Two-Class Pattern Detectors .....	7
3. Performance Metrics for 2 Class Detectors .....	9
ROC Curves.....	9
True Positives and False Positives.....	10
Precision and Recall .....	11
F-Measure .....	12
Accuracy .....	12

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1

Lesson 3

Winter Semester 2021

9 February 2021

## **Bayes Rule with Probability Distributions and Densities**

Notation.....	2
Probability .....	3
Probability as Frequency of Occurrence .....	3
Axiomatic Definition of probability .....	4
Bayes' Rule.....	5
Probability Distribution Tables .....	6
Joint Probability Distributions Tables (PDTs).....	7
Conditional Probability Tables (CPTs).....	9
Histograms for Numerical Properties .....	11
Bayes Rule with a Ratio of Histograms .....	12
Number of samples required .....	13
Mean and Standard Deviation.....	14
Histograms with integer and real valued features .....	14
Histograms for Vectors of Properties.....	15
Probability Density Functions.....	16
Bayes Rule with probability density functions .....	16

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MOSIG M1

Winter Semester 2021

Lesson 4

11 February 2021

## **Non-Parametric Models for Bayesian Recognition**

Notation.....	2
Bayesian Classification .....	3
Classification with a Ratio of Histograms .....	4
Number of samples required.....	5
Variable Sized Histogram Cells .....	6
Kernel Density Estimators .....	7
K Nearest Neighbors .....	10
Probability Density Functions.....	11
Bayes Rule with probability density functions .....	12
The Central Limit theorem and Normal densities. ....	12
Univariate Normal Density Function.....	13
Biased and Unbiased Variance .....	15
Multivariate Normal Density Function.....	16

Bibliographical sources:

"Pattern Recognition and Machine Learning", C. M. Bishop, Springer Verlag, 2006.

"Pattern Recognition and Scene Analysis", R. E. Duda and P. E. Hart, Wiley, 1973.

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 5

Winter Semester 2021  
23 February 2021

## Gaussian Mixture Models, K-Means and EM

Notation.....	2
Probability Density Functions.....	3
Bayes Rule with probability density functions .....	3
The Central Limit theorem and Normal Densities.....	4
Multivariate Normal Density Function .....	4
Gaussian Mixture Models .....	7
A Sum of Independent Sources.....	7
Estimating Gaussian Mixture models from Training Data .....	8
K-Means Clustering .....	9
The Expectation Maximization Algorithm (EM) .....	11
Convergence Criteria .....	13
Using Gaussian Mixture Models with Baye's Rule .....	14

### Sources:

C. M. Bishop, "Pattern Recognition and Machine Learning", Springer Verlag, 2006.

Jeff Bilmes, A Gentle Tutorial of the EM Algorithm, Tech Report, Univ of Washington, 1998.  
(available for download from course website).

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 6

Winter Semester 2021  
25 February 2021

## Support Vector Machines and Kernel Methods

### Outline

Notation .....	2
The Margin for a Linear Classifier .....	3
Linear Discriminant Functions .....	3
Margin .....	3
Support Vector Machines .....	4
Hard-Margin SVMs - a Simple Binary Classifier. ....	5
Finding the Support Vectors .....	6
Soft Margin SVMs.....	8
Kernel Methods .....	10
Quadratic Kernels .....	10
Radial Basis Function Kernels .....	12
Kernel Functions for Symbolic Data .....	13

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 7

Winter Semester 2021  
2 March 2021

## Perceptrons and Gradient Descent

### Outline

Notation .....	2
Perceptrons .....	3
History .....	3
The Perceptron Classifier .....	3
The Perceptron Learning Algorithm.....	5
Artificial Neural Networks .....	6
The Artificial Neuron .....	7
Homogeneous Coordinate Notation .....	8
Gradient Descent .....	9
Loss (Cost) Function .....	9
Feature Scaling .....	11
Local Minima .....	11
Batch mode .....	13
Stochastic Gradient Descent.....	13

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 8

Winter Semester 2021  
4 March 2021

## Artificial Neural Networks

### Outline

Notation .....	2
Introduction .....	3
Key Equations .....	3
Artificial Neural Networks .....	4
The Multilayer Neural Network model .....	6
Initializing the weights .....	9
Backpropagation .....	10
Derivation of Backpropagation as gradient Descent. ....	13
General formula for the error term .....	17
Formula for multiple activations .....	18
Summary of Backpropagation .....	19



# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 9

Winter Semester 2021  
9 March 2021

## **Generative Networks: Auto-Encoders, Variational Autoencoders and Generative Adversarial Networks**

### **Outline:**

Notation.....	2
Key Equations .....	2
Generative Networks .....	3
Cross-Entropy and the Kullback-Leibler Divergence.....	4
Entropy.....	4
Cross entropy .....	6
Binary cross entropy .....	6
Categorical Cross Entropy Loss.....	7
The Kullback-Leibler Divergence .....	8
AutoEncoders.....	9
The Sparsity Parameter .....	10
Auto-Encoders Encode a Signal as Latent Variables.....	12
Variational Autoencoders .....	13
Generative Adversarial Networks.....	14
Generative Networks.....	14
GAN Learning as Min-Max Optimization.....	14

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 10

Winter Semester 2021  
11 March 2021

## **Network Programming Exercise: Recognizing Handwritten Digits using Neural Networks**

### **Outline:**

The MNIST Digits Dataset .....	2
Installing a miniconda Programming Environment.....	4
Python .....	4
Conda Python.....	4
Installing MiniConda .....	5
Installation of MiniConda on an Apple Macintosh.....	5
Installation of MiniConda on Linux: .....	5
Numpy.....	6
Matplotlib.....	6
Jupyter Notebooks.....	6
Keras .....	8
Keras Code Examples .....	9
A Simple Example of an MLP for MNIST digits.....	11
Keras Example: a simple CNN for MNIST Digits .....	13

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 11

Winter Semester 2021  
16 March 2021

## Convolutional Neural Networks

### Outline

Notation .....	2
The Mammalian Visual Cortex .....	3
Receptive Fields in the Visual Cortex .....	3
Convolutional Neural Networks. ....	6
Fully-Connected Networks .....	6
Early Convolutional Neural Networks: LeNet5 .....	6
Multiple Receptive Fields at each Layer .....	9
CNN Hyper-parameters .....	10
Pooling.....	10
Classic CNN Architectures .....	11
AlexNet.....	13
VGG - Visual Geometry Group.....	14
A Keras example of a simple CNN .....	15

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lesson 12

Winter Semester 2021  
18 March 2021

## Locating Patterns in Images

### Outline

Introduction .....	2
Computer Vision Tasks used in ML challenges .....	2
Benchmark Data Sets for Object Detection .....	3
Data sets for other Visual tasks .....	5
Generative Convolutional Networks .....	6
Generating images with deconvolution .....	6
DCGAN .....	7
Deconvolution with VGG16 .....	8
YOLO: You Only Look Once .....	11
The Yolo-1 Network .....	12
Training YOLO .....	14
Loss Function for YOLO .....	15
Limitations of YOLO-1 .....	15
YOLO-9000 (YOLOv2) .....	16

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lessons 13

Winter Semester 2020-2021  
23 March 2021

## Recurrent Neural Networks

### Outline

Notation .....	2
Recurrent Neural Networks .....	3
Finite vs Infinite impulse networks .....	4
Finite Impulse Recurrent Networks .....	4
Tokenizing Word Data .....	6
Folding and Unfolding.....	7
Forward Propagation Equations .....	7
Training .....	8
Long Short-Term Memory (LSTM) .....	10
The Four layers of an LSTM unit.....	12

### Sources

- 1) Goodfellow, I., Bengio, Y., and Courville, A., Deep learning. MIT press, 2016.
- 2) Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning internal representations by error propagation. In Rumelhart, D. E. and McClelland, J. L., editors, Parallel Distributed Processing, volume 1, pages 318–362. MIT Press.
- 3) Understanding LSTM Networks - Christopher Olah (<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>)

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1

Winter Semester 2021

Lecture 14

25 march 2021

## Symbolic Reasoning, Expert Systems and MYCIN

### **Outline:**

Expert Systems and the AI Revolution of the 1980s ...	2
Knowledge and Reasoning .....	3
Kinds of Knowledge .....	3
Knowledge Based system .....	4
Expert System = Inference Engine + Domain Knowledge .....	4
The MYCIN Expert System .....	5
MYCIN: An Antibiotics Therapy Advisor .....	6
The MYCIN architecture .....	6
Reasoning with Backward Chaining Rules .....	7
Domain Concepts (Facts) .....	8
Parameters: the attributes of facts .....	9
The MYCIN Confidence Factor .....	10
Independent Rules and the Combine Function .....	10
Co-routines: Findout and Monitor .....	11
Why did expert systems fail? .....	13

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1  
Lessons 15

Winter Semester 2020-2021  
30 March 2021

## Dichotomizers, CART and Random Forests

### Outline

Notation .....	2
Decision Trees .....	3
Some Background from Information Theory .....	5
Entropy and Information Gain.....	5
GINI index or Gini Impurity .....	7
Comparison of Entropy, Gini , and Classification Error .....	8
Iterative Dichotomizers .....	9
The ID3 Algorithm .....	9
Improved ID3: The C4.5 algorithm .....	11
Classification and Regression Trees (CART).....	13
Classifying observations CART Models .....	14
Random Forests .....	17

### Sources

T. Hastie, R. Tibshirani and J. Friedman, "Elements of Statistical Learning", Springer, 2001  
C. M. Bishop, "Pattern Recognition and Machine Learning", Springer Verlag, 2006.  
[https://en.wikipedia.org/wiki/Decision\\_tree\\_learning](https://en.wikipedia.org/wiki/Decision_tree_learning)

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1

Second Semester 2020/2021

Lesson 16

1 April 2021

## **Structured Knowledge Representations: Working Memory, Concepts and Relations**

Background from Cognitive Science .....	2
Short Term and Long Term Memory .....	2
Recall .....	3
Working memory .....	3
Perception is Active, Action is Perceptive .....	4
Spreading Activation.....	5
Attention.....	5
Chunking .....	6
Conceptual Knowledge .....	6
Concepts .....	6
Schema .....	7
Relations .....	8
Kinds of Relations.....	8
Predicates .....	9
Relations as N-Ary Predicates .....	9
Implicit vs Explicit representations for Relations.....	10
RDF and the Semantic Web.....	12

### **Bibliography:**

- 1) W. Kintsch, *Comprehension: A paradigm for cognition*, Cambridge university press, 1998
- 2) G.A. Miller, The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review*, 63(2), p.81.1956
- 3) J. R. Anderson, A spreading activation theory of memory, *Journal of Verbal Learning and Verbal Behavior*, Volume 22, Issue 3, Pages 261-295, June 1983
- 4) T. Berners-Lee, J. Hendler, J., and O. Lassila, (2001). The Semantic Web. *Scientific American*, 284(5), 28-37. 2001.



# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MoSIG M1

Second Semester 2020/2021

Lesson 17

6 April 2021

## **Knowledge Structures: Frames, Scripts and Situation Models**

Structured Knowledge Representations .....	2
Frames .....	2
Scripts.....	4
Situation Models .....	6
Situation Models as State Spaces .....	8
Narrative Reasoning.....	10
Narrative Reasoning with Situation Models .....	10
Events .....	11
Observability .....	11
Narratives .....	12
Entailment .....	12
Causal reasoning. ....	13
Reasoning about Causality with Narratives. ....	14
Problems with Structured Knowledge Representations...	15

### **Bibliography:**

- 1) M. Minsky, "A Framework for Representing Knowledge", MIT AI Lab no 306, 1974.
- 2) R. C. Schank and R. P. Abelson, (1977). Scripts, plans, goals, and understanding: An inquiry into human knowledge structures.
- 3) Tomkins, S. S. "Script theory: Differential magnification of affects." In Nebraska symposium on motivation. University of Nebraska Press, 1978.
- 4) P. N Johnson-Laird, "Mental models", MIT Press Cambridge, MA, USA, 1989.

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MOSIG M1

Second Semester 2020/2021

Lesson 18

8 April 2021

## Temporal Reasoning with Interval Relations

Temporal Logic .....	2
Temporal relations .....	3
Table of Transitivity .....	5
Constraint Propagation .....	7
Reference Intervals. ....	9
 Example Problem in Temporal Reasoning .....	 10

Background Reading:

James F. Allen: Maintaining knowledge about temporal intervals. In: Communications of the ACM. 26 November 1983. ACM Press. pp. 832–843

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MOSIG M1

Second Semester 2020/2021

Lesson 19

13 April 2021

## Planning and Problem Solving

The Intelligent Agent .....	2
General Problem Solver and Means Ends Analysis.....	2
The Rationality Principle .....	3
Planning as Search .....	4
Problem Spaces .....	4
Blocks World .....	5
Predicates .....	5
Actions .....	6
Comments on Blocks World and Search .....	7
Algorithms for Planning as Search .....	8
Algorithmic Complexity of Search .....	8
Nilsson's Conditions for Optimal Search .....	10
Cost and Optimality of Heuristic Search .....	11
Cost of Search vs Optimality of Result.....	11
Hierarchical Planning, Subgoals and Chunking .....	12
Subgoals .....	12
Hierarchy of states .....	13
Operators .....	13
Chunking .....	14
Example: Travel Planning.....	15

### Background:

Simon, H. A. (1981). *The sciences of the artificial*. Cambridge, Massachusetts: MIT Press.

Nilsson, N. J., (1998). *Artificial intelligence: a new synthesis*. Morgan Kaufmann. (This is an updated version of Nilsson's classic 1980 textbook).

Korf R. E. (1987), Planning as search: A quantitative approach, *Artificial Intelligence*, Vol 33, Issue 1, pp65-88.

# Intelligent Systems: Reasoning and Recognition

James L. Crowley

MOSIG M1

Second Semester 2020/2021

Lesson 20

15 April 2021

## Reasoning with Bayesian Networks

Evidential Reasoning .....	2
Bayesian Networks .....	4
Probability Distribution Tables .....	5
Joint Probability Distributions Tables.....	5
Conditional Probability Tables (CPT) .....	6
Conditional Independence .....	7
Independent Random Variables .....	7
Conditional Independence.....	7
Chain Rule.....	8
Factoring Distribution Tables with Bayesian Networks ....	9
Computing with Conditional Probability Tables .....	9
A Joint Distribution in Structured Form .....	11
Reasoning with Bayesian networks .....	12
Diagnostic Reasoning .....	12
Predictive reasoning .....	13
Intercausal Reasoning .....	13
Markov Blanket.....	14
Constructing a Bayesian Network. ....	15

### Sources:

1. Koller, D., and Friedman, N., Probabilistic graphical models: principles and techniques. MIT press, 2009.
2. NEIL, Martin, FENTON, Norman, and NIELSON, Lars. Building large-scale Bayesian networks. *The Knowledge Engineering Review*, 2000, vol. 15, no 3, p. 257-284.