Transparent Computational Intelligence (CI) Models for Health States Monitoring of Complex Systems

Vom Promotionsausschus der Technischen Universität Hamburg-Harburg

zur Erlangung des akademischen Grades Doktor-Ingenieur (Dr.-Ing.)
genehmigte Dissertation

von

Peter Olatunde Owotoki

aus

Nigeria

2007
Table of Contents

1. INTRODUCTION.................................................................- 1 -
   1.1 OVERVIEW ...............................................................- 1 -
   1.2 MOTIVATION ..............................................................- 5 -
   1.3 THESIS CONTRIBUTIONS AND PUBLICATIONS .......................- 9 -
   1.4 THESIS OUTLINE .......................................................- 11 -

2. BACKGROUND.................................................................- 13 -
   2.1 INTRODUCTION AND DEFINITION OF HEALTH STATE MONITORING ....- 13 -
   2.2 HEURISTICS RULES BASED APPROACH .................................- 16 -
      2.2.1 The knowledge representation (or lost in translation) problem - 17 -
      2.2.2 The knowledge elicitation problem ..............................- 17 -
   2.3 MODEL BASED REASONING APPROACH ..................................- 19 -
   2.4 PECULIARITIES AND CHALLENGES OF MONITORING IN A COMPLEX SYSTEM DOMAIN ......................................................- 21 -
      2.4.1 Integration and architectural challenge .......................- 25 -
      2.4.2 Behavioural modelling challenge ..............................- 25 -
      2.4.3 Onboard monitoring challenge ....................................- 26 -
   2.5 CASE FOR A MULTI-AGENT DATA MINING APPROACH ..................- 27 -
      2.5.1 Agent based modelling ..........................................- 27 -
      2.5.2 Alternatives to agent based modelling ......................- 27 -
         2.5.2.1 Networks e.g. $\pi$-nets approach ....................- 28 -
         2.5.2.2 Differential equations ....................................- 29 -
      2.5.3 Reactive monitoring agents for embodied engineered complex systems ......................................................- 29 -
   2.6 DATA MINED COMPUTATIONAL INTELLIGENCE MODELS .................- 32 -
      2.6.1 Computational intelligence models ............................- 33 -
         2.6.1.1 Artificial neural networks (ANN) or connectionist models - 34 -
         2.6.1.2 Fuzzy computing models .................................- 35 -
         2.6.1.3 Evolutionary computing models .........................- 35 -
         2.6.1.4 Logic and information theoretic based models ............- 36 -
      2.6.2 The standard data mining process model ......................- 36 -

3. TRANSPARENCY OF COMPUTATIONAL INTELLIGENCE MODELS FOR MONITORING......................................................- 39 -
   3.1 CASE FOR TRANSPARENCY ..............................................- 39 -
   3.2 IMPORTANT TRANSPARENCY RELATED WORK ...........................- 42 -
   3.3 TRANSPARENCY DEFINED ..............................................- 44 -
   3.4 TRANSPARENCY QUOTIENT .............................................- 46 -
3.5 TRANSPARENCY FACTORS ................................................................. 48
  3.5.1 Rules and explanations (R) ....................................................... 48
  3.5.2 Degree of confidence (k) .......................................................... 50
  3.5.3 Visualizations of principal features (V) ...................................... 51
3.6 IMPLEMENTATION OF TRANSPARENCY ........................................ 53
  3.6.1 The transparency enabled generalized exemplars Model (TEGE) .... 53
     3.6.1.1 Rules and Explanation (R) ............................................... 53
     3.6.1.2 Degree of Confidence (k) .............................................. 54
     3.6.1.3 Visualizations of Principal Features (V) ............................. 61
3.7 PRACTICAL TIDBITS FOR DEPLOYMENT OF TRANSPARENT CI MODELS FOR MONITORING ................................................................. 62

4. AGENT BASED ARCHITECTURE FOR ONBOARD MONITORING OF COMPLEX SYSTEMS ................................................................. 67
  4.1 OVERVIEW OF THE ARCHITECTURE ............................................. 67
  4.2 FUNCTIONAL COMPONENTS ............................................................. 71
  4.3 LOCAL MONITORING AGENTS ........................................................ 73
     4.3.1 Data acquisition and generation module (DAGM) ....................... 73
        4.3.1.1 Time driven data capture ........................................... 74
        4.3.1.2 Event driven data capture .......................................... 74
     4.3.2 Embedded CI model .............................................................. 75
        4.3.2.1 The input data (X_i) .................................................. 76
        4.3.2.2 The filter ................................................................. 76
        4.3.2.3 The classifier ............................................................ 77
        4.3.2.4 The output pattern .................................................... 78
     4.3.3 Message formatting and packaging module ............................. 78
     4.3.4 Local monitoring agents' interface ..................................... 81
  4.4 ANCILLARY COMPONENTS ............................................................... 83
     4.4.1 Message agglomeration components or GANGLIA ..................... 83
     4.4.2 Complex system monitoring director (CSMDIR) ......................... 84
     4.4.3 Human machine interface manager (HMIMNGR) ......................... 86
     4.4.4 Onboard data storage (DATASTORE) .................................... 87
  4.5 INTERFACE TO THE OFF BOARD DATA MINER .................................. 88
  4.6 OFF BOARD DATA MINER ............................................................... 89
     4.6.1 Specification of the health states ...................................... 89
     4.6.2 Specification of important parameters .................................. 89
     4.6.3 Collection of relevant data .............................................. 90
        4.6.3.1 Construction of data collection platform .......................... 90
     4.6.4 Off board model generation .............................................. 91
        4.6.4.1 Model selection ....................................................... 91
        4.6.4.2 Model fine-tuning .................................................... 91
  4.7 DISCUSSION AND CONCLUSION .................................................... 93

x
5. SELECTION OF COMPUTATIONAL INTELLIGENCE MODELS

5.1 INTRODUCTION

5.1.1 Learning bias of a computational intelligence model

5.1.2 The model selection problem

5.2 THEORETICAL BASIS FOR MODEL SELECTION

5.2.1 Ockham's razor or law of parsimony (Lex Parsimoniae)

5.2.2 Bayesian theorem approach

5.2.2.1 Bayesian model averaging (BMA)

5.2.2.2 Bayesian model comparison (BMC)

5.2.2.3 Bayesian information criterion (BIC) or Schwarz information criterion (SIC)

5.2.2.4 Akaike information criterion (AIC)

5.2.2.5 Bayes factor

5.2.3 Zero sum theorems

5.2.3.1 The no-free-lunch theorem for search and optimization

5.2.3.2 The conservation law of generalization performance

5.3 PERFORMANCE BOUNDS APPROACH FOR MODEL SELECTION

5.3.1 Probably approximately correct (PAC) learning model

5.4 META LEARNING APPROACH FOR MODEL SELECTION

5.4.1 Statistical and information theoretic task characterization

5.4.2 Task characterization via landmarking

5.4.3 Task characterization via the representation of an inductive learner

5.5 EMPIRICAL MEASURES AND METRICS FOR PERFORMANCE COMPARISON

5.5.1 Predictive accuracy (PA)

5.5.2 Confusion matrix

5.5.2.1 False negative rate

5.5.2.2 False positive rate

5.5.2.3 Precision

5.5.2.4 Recall or sensitivity

5.5.2.5 Specificity or true negative

5.5.3 Receiver operating characteristics (ROC) and area under ROC curve (AUC)

5.5.4 Cross Validation

5.6 LEONARDO – THE AUTOMATIC MODEL SELECTION WIZARD

5.6.1 LEONARDO’s architecture

5.6.2 The data layer and the knowledge bases

5.6.2.1 Knowledge base of classifiers

5.6.2.2 Tasks performance knowledge base (TPKB)

5.6.3 The LEONARDO application core

5.6.3.1 The pre-selection component

5.6.3.2 The ranking component

5.6.4 The user interface layer

5.6.4.1 The TPKB creator

5.6.4.2 The model selection wizard

5.6.4.3 The pre-selection UI
5.7 RESULTS OF MODEL SELECTION EXPERIMENTS WITH LEONARDO.

5.7.1 "Balance-scale" dataset

5.7.2 "Eucalyptus" dataset

5.7.3 "Ionosphere" dataset

5.7.4 "Lymph" dataset

5.7.5 "Sonar" dataset

5.7.6 Evaluation on "racks" dataset obtained from the Mercedes Benz factory in Hamburg Germany

5.8 DISCUSSION AND CONCLUSION

6. APPLICATIONS

6.1 INTRODUCTION

6.2 HEALTH STATE MONITORING FOR THE AIRCRAFT CABIN ILLUMINATION FUNCTIONAL COMPONENT

6.2.1 Limitations of current approaches for monitoring of cabin illumination system

6.2.1.1 The runtime mismatch disadvantage

6.2.1.2 The isolated - integrated mismatch

6.2.1.3 The non dynamic monitoring logic disadvantage

6.2.2 Prototype for illumination functional component data and CI model generation

6.2.2.1 Definition and enumeration of health states of interests

6.2.2.2 Specification of important contributing parameters

6.2.2.3 Prototype illumination component for data collection

6.2.2.4 Data pre-processing

6.2.2.5 Model selection

6.2.2.6 Model construction and fine tuning

6.2.2.7 Model deployment

6.3 MONITORING OF THE AIRBUS IDG FUNCTIONAL COMPONENT

6.3.1 The dataset

6.3.2 CI model selection, deployment and evaluation

6.4 FUNCTIONS BASED ON ARTIFICIALLY GENERATED DATASET AND BENCHMARK DATA

6.4.1 Generic health state monitoring with artificially generated dataset

6.4.1.1 Model selection

6.4.2 Classifier function with UCI sonar benchmark dataset

6.5 SIMULATOR OF THE ONBOARD CABIN MONITORING SYSTEM

6.5.1 Implementation of the simulated onboard functional components

6.5.2 Implementation of the simulated onboard ancillary components

6.6 DISCUSSION AND CONCLUSION
7. CONCLUSION

7.1 SUMMARY

7.2 OUTLOOK AND FUTURE WORK

APPENDIX

A. LIST OF COMPUTATIONAL INTELLIGENCE MODELS APPLIED IN THESIS

B. XML SCHEMA OF THE KNOWLEDGE BASE OF CLASSIFIERS

C. KNOWLEDGE BASE OF CLASSIFIERS ENTRIES FOR CI MODELS APPLIED IN
   THESIS

D. XML SCHEMA OF THE TASKS PERFORMANCE KNOWLEDGE BASE

E. FUNCTIONAL COMPONENTS OF AN AIRBUS CABIN SYSTEM

F. SAMPLE OF THE PROCESSED AIRBUS ILLUMINATION DATASET

LIST OF FIGURES

LIST OF TABLES

BIBLIOGRAPHY

CURRICULUM VITAE