

Supplemental Material: A Classification Study on Testing and Verification of AI-based Systems

Emanuele De Angelis¹, Guglielmo De Angelis², and Maurizio
Proietti³

¹*IASI-CNR*, Rome, Italy, emanuele.deangelis@iasi.cnr.it

²*IASI-CNR*, Rome, Italy, guglielmo.deangelis@iasi.cnr.it

³*IASI-CNR*, Rome, Italy, maurizio.proietti@iasi.cnr.it

Data Set

This archive contains the whole set of documents that have been considered in the classification study reported in the manuscript titled: “A Classification Study on Testing and Verification of AI-based Systems”. Specifically, the file:

- `Classification_study_complete_table.ods`

reports the raw data about the 915 collected and processed according to the methodology described in Section 3 of the manuscript, while

- `Classification_study_primaryStudies_table.pdf`

contains the 78 resulting primary studies that are also detailed in the following.

Acknowledgments

This work has been partially supported by the project OPENNESS (Grant GeCoWEB A0375-2020-36616) founded by Regione Lazio through Research Program POR FESR LAZIO 2014-2020 – GRUPPI DI RICERCA 2020. The authors are members of the Italian Research Group: INdAM-GNCS.

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation			
								Yes/No	Method	Tool	Data
Zhu H., Bayley I.	Discovering boundary values of feature-based machine learning classifiers through exploratory datamorphic testing	2022	Testing	Exploratory Datamorphic	ML classification/clustering models	Generation of Pareto fronts (borders between classes defined by the ML-model UT).	Feature-based Classification	TRUE	Experiments	Morphy (https://github.com/hongzhu6129/ExploratoryTes tAI.git)	TRUE
Herbold S., Haar T.	Smoke testing for machine learning: simple tests to discover severe bugs	2022	Testing	Combinatorial Smoke	ML classification/clustering algorithms	Debugging	Feature-based Classification; Clustering	TRUE	Experiments; PoC	Atoml (https://github.com/sherbold/at com/)	TRUE
Bayani D., Mitsch S.	Fanoos: Multi-resolution, Multi-strength, Interactive Explanations for Learned Systems	2022	Verification	Reachability Analysis	ML-model agnostic	Formal/Probabilistic Explainability	Inverted Double Pendulum; CPU Usage	TRUE	PoC	Fanoos (https://github.com/DBayani/Fanoos)	
Maino C., Mastropietro A., Sorrentino L., Busto E., Misul D., Spessa E.	Project and Development of a Reinforcement Learning Based Control Algorithm for Hybrid Electric Vehicles	2022	Testing	Simulation & Monitoring	Q-learning RL-models	Energy Management System (EMS) Performance Evaluation	Real-time EMS of Hybrid Electric Vehicles	TRUE	Experiments		
Marta D., Pek C., Melsion G.I., Tumova J., Leite I.	Human-Feedback Shield Synthesis for Perceived Safety in Deep Reinforcement Learning	2022	Verification	Simulation & Monitoring	Deep RL-models	Safety in HRI	Human-Robot Interaction	TRUE	Experiments; PoC		
Xu L., Towey D., French A.P., Benford S., Zhou Z.Q., Chen T. Y.	Using metamorphic relations to verify and enhance Artcode classification	2021	Testing	Metamorphic	RF/SVM classifiers	Correctness of classification: Separation and Occlusion MRS	Image w/Artcode Classification	TRUE	Experiments		
Shea-Blymyer C., Abbas H.	Algorithmic ethics: Formalization and verification of autonomous vehicle obligations	2021	Verification	Model Checking	AI-model agnostic	Dominance Act Utilitarianism formulas	Autonomous Cyber-Physical Systems: Autonomous vehicles	TRUE	PoC	MC-DAU (https://github.com/sabotagelab/MC-DAU) Jupyter Notebook source code (https://doi.org/10.5281/zenodo.)	TRUE
Santos S., Silveira B., Durelli V., Durelli R., Souza S., Delamaro M.	On Using Decision Tree Coverage Criteria forTesting Machine Learning Models	2021	Testing	White-box	Decision Trees	Decision Tree Coverage/Boundary Value Analysis -adequate Test Case generation	Agnostic	TRUE	Experiments		TRUE

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation		
								Yes/No	Method	Data
Rehman F.U., Izurieta C.	A hybridized approach for testing neural network based intrusion detection systems	2021	Testing	Statistical Metamorphic Mutation	Neural Networks	Correctness of classification: w.r.t Shifting the features by constant k MR	Neural Network-based Intrusion Detection Systems	TRUE	Experiments	TRUE
Eliyahu T., Kazak Y., Katz G., Schapira M.	Verifying learning-augmented systems	2021	Verification	Constraint Reasoning Model Checking	Deep RL-models	Safety; Liveness	Aurora Congestion Controller; Pensieve Video Streamer; DeepRM Resource Manager	TRUE	Experiments	TRUE
Ma Y., Towey D., Chen T.Y., Zhou Z.Q.	Metamorphic testing of fake news detection software	2021	Testing	Metamorphic	ML-model agnostic	Correctness of classification w.r.t.	Fake News Detection	TRUE	Experiments	TRUE
Abarca A.I.R., Broersen J.	A deontic stit logic based on beliefs and expected utility	2021	Hybrid	Model Checking	AI-model agnostic	Explaining agent's actions	Ethical behavior of AI			
Steging C., Renooij S., Verheij B.	Discovering the rationale of decisions: Towards a method for aligning learning and reasoning	2021	Testing	Dataset alteration	Black-box ML-models	Accuracy; XAI: Rationale discovery	Legal	TRUE	Experiments	TRUE
Lwakatare L.E., Range E., Crnkovic I., Bosch J.	On the Experiences of Adopting Automated Data Validation in an Industrial Machine Learning Project	2021	Testing	Dataset validation	ML-training dataset	User-defined data quality properties	HW Telecommunication Devices Faults			
Usman M., Noller Y., Pasareanu C.S., Sun Y., Gopinath D.	NEUROSPF: A Tool for the Symbolic Analysis of Neural Networks	2021	Testing	Symbolic Execution Model Checking	Neural Networks	Coverage-based test input generation and execution; Generating adversarial examples and robustness analysis	Image Classification	TRUE	Experiments	TRUE
Chandrasekaran J., Lei Y., Kacker R., Richard Kuhn D.	A combinatorial approach to explaining image classifiers	2021	Testing	Combinatorial	Deep Neural Networks	XAI: Counterfactual explanation	Image Classification	TRUE	Experiments	TRUE

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation		
								Yes/No	Method	Data

Spieker H., Gotlieb A.	Adaptive metamorphic testing wit	2020	Testing	Metamorphic	ML-model agnostic	Selection of MRS	Image Classification; Object detection	TRUE	Experiments	Tetraband (http://github.com/helges/tetra band)
Arora H., Bielski J., Eisenstadt V., Langenhan C., Ziegler C., Althoff K.-D., Dengel A.	Consistency Checker: An automatic constraint-based evaluator for housing spatial configurations	2021	Verification	Model Checking	Deep Neural Networks	Consistency	Spatial configurations of housing construction			
Gao K., Wang J., Wang B., Wang R., Jia J.	UAV Test Data Generation Method based on CycleGAN	2021	Testing	Dataset alteration	ML-testing dataset	Quality of generated tests: Peak signal-to-noise ratio; Structural similarity	Unmanned Aerial Vehicles	TRUE	Experiments	
Lu Y., Sun W., Sun M.	Mutation Testing of Reinforcement Learning Systems	2021	Testing	Mutation	RL-models	Mutation score (quantitative eval. of the test environment); behavior differences between original and mutated agents	Agnostic	TRUE	PoC	
Pore A., Corsi D., Marchesini E., Dall'alba D., Casals A., Farinelli A., Fiorini P.	Safe Reinforcement Learning using Formal Verification for Tissue Retraction in Autonomous Robotic-Assisted Surgery	2021	Verification	Interval Analysis	Deep RL-models	Safety	Robot-assisted Minimally Invasive Surgery	TRUE	Experiments	
Termine A., Primiero G., D'Asaro F.A.	Modelling Accuracy and Trustworthiness of Explaining Agents	2021	Verification	Model Checking	Opaque ML-models	Accuracy; Trustworthiness of post-hoc explanations	Agnostic	TRUE	PoC	
Pistol I.C., Arusoaie A.	Analysing state-based models for AI problems	2021	Verification	Constraint Reasoning	AI-model agnostic	State-based models validation wrt. Valid Final State (VFS); Path to a Final State (PFS) properties (enumeration of solutions)	Agnostic	TRUE	PoC	TRUE
Riley J., Calinescu R., Paterson C., Kudenko D., Banks A.	Utilising Assured Multi-Agent Reinforcement Learning within safety-critical scenarios	2021	Verification	Quantitative verification	Reinforcement Learning-enabled Multi Agent Systems	Probabilistic Computation Tree Logic (PCTL); properties: Safety; Reliability; Performance constraints	MAS patrolling problem	TRUE	Experiments	

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation		
								Yes/No	Method	Data

Gros T.P., Höller D., Hoffmann J., Klauck M., Meerkamp H., Wolf V.	DSMC Evaluation Stages: Fostering Robust and Safe Behavior in Deep Reinforcement Learning	2021	Verification	Quantitative verification	Deep RL-models	(Local) Robustness; Safety	Racetrack: abstraction of autonomous driving	TRUE	PoC	
Amir G., Schapira M., Katz G.	Towards Scalable Verification of Deep Reinforcement Learning	2021	Verification	Constraint Reasoning Model Checking	Deep RL-models	Safety; Liveness	Aurora Congestion Controller; Pensieve Video Streamer; DeepRM Resource Manager	TRUE	Experiments	whiRL 2.0 (#">https://zenodo.org/record/4769612#). YoUcmzlbx8w)
Wang B., Hou Z., Zhang G., Shi J., Huang Y.	Tree ensemble property verification from a testing perspective	2021	Hybrid	Model Checking	Tree ensembles	Correctness of predictions: user-defined constraint-based properties	Agnostic	TRUE	Experiments	
Mqirmi P.E., Belardinelli F., León B.G.	An abstraction-based method to check multi-agent deep reinforcement-learning behaviors	2021	Verification	Quantitative verification	Reinforcement Learning-enabled Multi Agent Systems	weak fragment of Probabilistic Computation Tree Logic (wPCTL); properties: Safety; Optimality	Guarded flag collection	TRUE	Experiments	
Perez Morales D., Kitamura T., Takada S.	Coverage-Guided Fairness Testing	2021	Testing	Combinatorial	ML-model agnostic	Fairness	Decision-making algorithms; Classifiers	TRUE	Experiments	
Nakajima S.	Software Testing with Statistical Partial Oracles: - Application to Neural Networks Software -	2021	Testing	Metamorphic Statistical	Deep Neural Networks	Correctness of prediction	Image classification	TRUE	Experiments	
Riley J., Calinescu R., Paterson C., Kudenko D., Banks A.	Reinforcement learning with quantitative verification for assured multi-agent policies	2021	Verification	Quantitative verification	Reinforcement Learning-enabled Multi Agent Systems	Safety; Reliability; Performance constraints	MAS patrolling problem	TRUE	Experiments	
Wang Y., Sha W., Pan Z., Luo Y.	Research on Autonomous Driving Perception Test Based on Adversarial Examples	2021	Testing	Boundary Value	Deep Neural Networks	Correctness of classification/detection	Autonomous driving	TRUE	Experiments	

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation		
								Yes/No	Method	Data

Park S.-H., Lee I.-G.	Effective Voice Fuzzing Method for Finding Vulnerabilities in AI Speech Recognition Devices	2020	Testing	Fuzz	Deep Neural Networks	Correctness of classification	Speech recognition devices	TRUE	Experiments	
Chakraborty J., Majumder S., Yu Z., Menzies T.	Fairway: A way to build fair ML software	2020	Testing	Simulation & Monitoring	ML-model agnostic	Fairness; Fairness metrics: Equal Opportunity Difference; Average Odds Difference	Binary classifier	TRUE	Experiments	Fairway (https://github.com/Joymallyac/Fairway)
Corsi D., Marchesini E., Farinelli A., Fiorini P.	Formal Verification for Safe Deep Reinforcement Learning in Trajectory Generation	2020	Verification	Interval analysis	Deep RL-models	Safety	Robotics	TRUE	Experiments	
Liu W.-W., Song F., Zhang T.-H.-R., Wang J.	Verifying ReLU Neural Networks from a Model Checking Perspective	2020	Verification	Model Checking	ReLU Neural Networks	ReLU temporal logic; properties: interval; local robustness; reachability	Agnostic	TRUE	Experiments	
Santos S.H.N., Da Silveira B.N.C., Andrade S.A., Delamaro M., Souza S.R.S.	An Experimental Study on Applying Metamorphic Testing in Machine Learning Applications	2020	Testing	Metamorphic	ML classification/clustering models	Supervised ML classifier MRs	Breast Cancer Diagnostic	TRUE	Experiments	(Google Colab. notebook)
Zhu H., Bayley I.	Exploratory datamorphic testing of classification applications	2020	Testing	Exploratory Datamorphic	ML classification/clustering models	Generation of Pareto fronts Accuracy of Classification; Effectiveness of Classification	Classification	TRUE	Experiments	
Moreira D., Furtado A.P., Nogueira S.	Testing acoustic scene classifiers using Metamorphic Relations	2020	Testing	Metamorphic	ML-technique agnostic	ML classifier MRs	Acoustic Scene Classification	TRUE	Experiment	
Trenquier H., Ishikawa F., Tokumoto S.	Attribute-based Granular Evaluation for Performance of Machine Learning Models	2020	Testing	Dataset alteration Attribute-based	ML-model agnostic	Attribute Selection, Performance of the Prediction; Accuracy in the Classification	Image Classification	TRUE	Experiment	

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation		
								Yes/No	Method	Data

Zhu H., Bayley I., Liu D., Zheng X.	Automation of Datamorphic Testing	2020	Testing	Datamorphic	ML-model agnostic	Datamorphic Framework and its implementation	Classification	TRUE	PoC	Morphy (https://github.com/hongzhu6129/MorphyExamples) MonotonicityChecker (https://github.com/arnabsharma91/MonotonicityChecker)	TRUE
Sharma A., Wehrheim H.	Higher income, larger loan? monotonicity testing of machine learning models	2020	Testing	Constraint Reasoning	Black-box ML-models	Test input generation for checking Monotonicity	Classification	TRUE	Experiments		TRUE
Bose S., Barao T., Liu X.	Explaining AI for Malware Detection: Analysis of Mechanisms of MalConv	2020	Testing	Dataset alteration	Neural Networks	Explainability: decision boundaries between classes	Cybersecurity; Malware Detection	TRUE	Experiments		
Liem C.C.S., Panichella A.	Oracle Issues in Machine Learning and Where to Find Them	2020	Testing	Information Entropy Semantic Analysis	ML-technique agnostic	Oracle Assessment	Image Classification	TRUE	Experiments		
Trujillo M., Linares-Vásquez M., Escobar-Velásquez C., Dusparic I., Cardozo N.	Does Neuron Coverage Matter for Deep Reinforcement Learning?: A Preliminary Study	2020	Testing	Simulation & Monitoring White-box	Deep Neural Networks	Evolutionary Neuron Coverage	Mountain car problem	TRUE	PoC		
Martins J., Barbosa R., Lourenco N., Robin J., Madeira H.	Online Verification through Model Checking of Medical Critical Intelligent Systems	2020	Verification	Model Checking	Rule-based Classifiers	Detection of gray area of the decision space	Medical Critical Intelligent Systems	TRUE	Experiments		
Sharma A., Wehrheim H.	Automatic Fairness Testing of Machine Learning Models	2020	Testing	Property-based Constraint Reasoning	Black-box ML-models	Test input generation for checking Fairness	Classification	TRUE	Experiment	fairCheck (https://github.com/arnabsharma91/fairCheck)	TRUE
Ma P., Wang S., Liu J.	Metamorphic testing and certified mitigation of fairness violations in NLP models	2020	Testing	Metamorphic	AI-model agnostic	Fairness	NLP: CNN model for Sentiment Analysis	TRUE	Experiments		TRUE

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation		
								Yes/No	Method	Data

Liu Z., Xiong L., Liu Y., Lespérance Y., Xu R., Shi H.	A modal logic for joint abilities under strategy commitments	2020	Verification	Model Checking	AI-model agnostic	Alternating-time Dynamic Logic with Joint Abilities (JAADL)	Agnostic			
Carr S., Jansen N., Topcu U.	Verifiable RNN-based policies for pomdps under temporal logic constraints	2020	Verification	Quantitative Verification	Recurrent Neural Networks	Linear-time Temporal Logic formulas	Control policies in sequential decision-making problems	TRUE	Experiments	
Bacci E., Parker D.	Probabilistic Guarantees for Safe Deep Reinforcement Learning	2020	Verification	Quantitative Verification	Deep RL-models	Safety: bounds on the probability of safe operation; regions of correct behaviour	Deep RL controllers: Pendulum; Cartpole	TRUE	PoC	
Lamrani I., Banerjee A., Gupta S.K.S.	Toward operational safety verification via hybrid automata mining using I/O Traces of AI-Enabled CPS	2020	Hybrid	Simulation & Monitoring	AI-model agnostic	Safety	CPSs: artificial pancreas control system	TRUE	Experiments	
[No author name available]	A verified compositional algorithm for AI planning	2019	Verification	Interactive Theorem Proving	Algorithm for AI Planning	Correctness	Agnostic			
Kazak Y., Barrett C., Katz G., Schapira M.	Verifying deep-rl-driven systems	2019	Verification	Constraint Reasoning	Deep RL-models	Safety	Pensieve Video Streamer; DeepRM Resource Manager; Custard Internet Congestion Controller	TRUE	Experiments	TRUE
Zhang S., Yan G., Li Y., Liu J.	Evaluation of Judicial Imprisonment Term Prediction Model Based on Text Mutation	2019	Testing	Mutation	Neural Networks	Robustness	Judicial Imprisonment Term Prediction	TRUE	Experiments	
Zhu H., Magill S., Xiong Z., Jagannathan S.	An inductive synthesis framework for verifiable reinforcement learning	2019	Verification	Constraint Reasoning	RL-models	Runtime Shielding	Control- and Cyberphysical-system	TRUE	Experiments	

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	EMPIRICAL EVALUATION		
								YES/NO	Method	Tool

Saha P., Kanewala U.	Fault detection effectiveness of metamorphic relations developed for testing supervised classifiers	2019	Testing	Metamorphic	Supervised ML-models	Effectiveness of MRs	kNN classifiers	TRUE	Experiments		
Zhu H., Liu D., Bayley I., Harrison R., Cuzzolin F.	Datamorphic testing: A method for testing intelligent applications	2019	Testing	Datamorphic Metamorphic	ML-model agnostic	Correctness w.r.t. MRs	Face recognition applications	TRUE	Experiments		
Sato T., Aguirre A., Barthe G., Gaboardi M., Garg D., Hsu J.	Formal verification of higher-order probabilistic programs	2019	Verification	Quantitative Verification	Higher-order probabilistic programs	Probabilistic Logic specifications	Agnostic	TRUE	PoC		
Chakraborty S., Meel K.S.	On testing of uniform samplers	2019	Testing	Conditional sampling	Probabilistic generators	Almost-uniform generator: generator distribution - uniform distribution distance	Agnostic	TRUE	Experiments	Barbarik (https://github.com/meelgroup/barbarik) PAMC (http://faculty.sist.shanghaitech.edu.cn/faculty/songfu)	
Song F., Zhang Y., Chen T., Tang Y., Xu Z.	Probabilistic alternating-time μ -calculus	2019	Verification	Quantitative Verification	Stochastic Multi Agent Systems	Probabilistic Alternating μ -Calculus	Genetic Regulatory Networks	TRUE	Experiments		
Rabanser S., G�nnemann S., Lipton Z.C.	Failing loudly: An empirical study of methods for detecting dataset shift	2019	Testing	Dataset alteration	ML datasets	Dataset shift detection	Image classification datasets	TRUE	Experiments	failing-loudly (https://github.com/steverab/failing-loudly)	TRUE
Guidotti D., Leofante F., Pulina L., Tacchella A.	Verification and Repair of Neural Networks: A Progress Report on Convolutional Models	2019	Verification	Constraint Reasoning	Convolutional Neural Networks	Robustness	Image classification	TRUE	Experiments		TRUE
Nakajima S., Chen T. Y.	Generating Biased Dataset for Metamorphic Testing of Machine Learning Programs	2019	Testing	Metamorphic	Neural networks	(Follow-up) dataset generation	Image classification	TRUE	PoC		

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation		
								Yes/No	Method	Tool

Törnblom J., Nadjm-Tehrani S.	An Abstraction-Refinement Approach to Formal Verification of Tree Ensembles	2019	Verification	Model Checking	Tree ensembles	Robustness	Vehicle Collision Detection; Digit Recognition	TRUE	PoC; Experiments	VoTE (https://github.com/John-tornblom/VoTE/releases/tag/v0.2.1)
Platzter A.	The logical path to autonomous cyber-physical systems	2019	Verification	Model Checking	Cyber-physical Systems	Differential dynamic logic; Safety; Safe learning	Cyber-Physical Systems			
Dreossi T., Fremont D.J., Ghosh S., Kim E., Ravanbakhsh H., Vazquez-Chanlatte M., Seshia S.A.	VeriFAL: A Toolkit for the Formal Design and Analysis of Artificial Intelligence-Based Systems	2019	Hybrid	Simulation & Monitoring Fuzz Model Checking	Cyber-physical Systems	Temporal-logic falsification; model-based fuzz testing; dataset augmentation; counterexample analysis; hyper/model-	Cyber-physical systems	TRUE	PoC	VeriFAL (https://github.com/BerkeleyLearnVerify/VeriFAL)
Nakajima S.	Dataset Diversity for Metamorphic Testing of Machine Learning Software	2019	Testing	Metamorphic	Neural Networks	Test Input Dataset Diversity; Behavioral oracle MRS	Image Classification	TRUE	Experiments	
Amortila P., Bellemare M.G., Panangaden P., Precup D.	Temporally extended metrics for markov decision processes	2019	Verification	Quantitative Verification	RL-models	Safety	Agnostic			
Ma S., Liu Y., Lee W.-C., Zhang X., Grama A.	MODE: Automated neural network model debugging via state differential analysis and input selection	2018	Testing	Differential Analysis	Neural Networks	Debugging	Image Classification	TRUE	Experiments	TRUE
Belardinelli F., van der Hoek W., Kuijler L.B.	Second-order propositional modal logic: Expressiveness and completeness results	2018	Verification	Model Checking	AI-model agnostic	Second-order Propositional Epistemic Logic (SOPEL)	Agnostic			
Mallozzi P., Pardo R., Duplessis V., Pelliccione P., Schneider G.	MoVEMo: A structured approach for engineering reward functions	2018	Verification	Model Checking Simulation & Monitoring	RL reward functions	Correctness of the reward model	Autonomous Driving Systems	TRUE	Experiments	

Authors	Title	Year	Approach	Technique	Subject	Property/Goal	Application Domain	Empirical Evaluation		
								Yes/No	Method	Tool

Pathak S., Pulina L., Tacchella A.	Verification and repair of control policies for safe reinforcement learning	2018	Verification	Quantitative Verification	RL-models	Safety; Model Repair	Robotics	TRUE	Experiments	
Fulton N., Platzer A.	Safe reinforcement learning via formal methods: Toward safe control through proof and learning	2018	Verification	Simulation & Monitoring	RL-models	Differential dynamic logic; Safety	Cyber-physical Systems	TRUE	Experiments	
Junges S., Jansen N., Katoen J.-P., Topcu U., Zhang R., Hayhoe M.	Model checking for safe navigation among humans	2018	Verification	Quantitative Verification	RL-models	Safety	Autonomous Systems w/Uncontrollable Agents	TRUE	Experiments	TRUE
Dutta S., Jha S., Sankaranarayanan S., Tiwari A.	Output range analysis for deep feedforward neural networks	2018	Verification	Interval Analysis	Deep Neural Networks	Reachability Analysis	NN-based control systems	TRUE	Experiments	TRUE
Mason G., Calinescu R., Kudenko D., Banks A.	Assured reinforcement learning with formally verified abstract policies	2017	Verification	Quantitative Verification	Reinforcement Learning-enabled Multi Agent Systems	Probabilistic Computational Tree Logic (PCTL)	Guarded Flag Collection; Autonomous Assistance for Dementia Sufferer	TRUE	Experiments	
Cimatti A., Micheli A., Roveri M.	Validating domains and plans for temporal planning via encoding into infinite-state linear temporal logic	2017	Verification	Model Checking	Temporal Plans	Linear Temporal Logic (LTL) modulo Rational Arithmetic (RA) Satisfiability/Validity	Temporal Planning	TRUE	Experiments	TRUE

NUXMV
 (<https://es.fbk.eu/people/amicheli/resources/aaa117>)