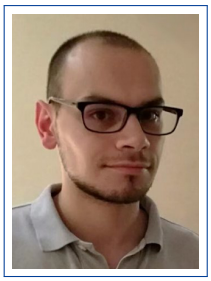
MATTHIEU ZIMMER

Postdoctoral Researcher in Deep Reinforcement Learning

Born on: May 30 1990
Nationality: French

Minhang District
Shanghai, China
☎ *+86 185 2139 8961*
✉ *contact@matthieu-zimmer.net*
https://matthieu-zimmer.net



RESEARCH ACTIVITIES

Keywords Deep Reinforcement Learning, Deep Learning, Transfer Learning, Multi-Agent, Machine Learning.

Publications ICML, IJCAI, IEEE RA-Letters, IROS, ICRA, etc.

Projects Fairness in reinforcement learning (Shanghai NFS), Neural logic reinforcement learning (Huawei), Reinforcement learning in robotics (Chinese University of Hong Kong)

Co-Supervision 3 PhD students, 5 Masters students and 12 undergraduate students.

Reviewer ICML, NeurIPS, TNNLS, ICLR, AAAI, CORL, IJCAI-PRICAI, IEEE RA-Letters, ICRA, IROS, ICDL.

EXPERIENCES

June 2018 – ongoing — **Postdoctoral Researcher**, Univ. Michigan-SJTU Joint Institute — Shanghai (CH)
Supervisor: Paul Weng (Assistant Professor UM-SJTU Joint Institute)
- Development of deep reinforcement learning algorithms (actor-critic, fairness, exploitation of symmetries, multi-agent, interpretability).
- Setting up and manage the team cluster (15 hosts, 26 GPUs) with its web services.

2014-2017 – 205 hours — **Teaching Assistant**, University of Lorraine - ENSEM — Nancy (FR)

2014 – 6 months — **Research Intern**, Pierre and Marie Curie University - ISIR — Paris (FR)
Topic: Deep reinforcement learning and neuroevolution

2013 – 3 months — **Research Intern**, Pierre and Marie Curie University - LIP6 — Paris (FR)
Topic: Advising in reinforcement learning

2012 – 5 months — **Research Intern**, University of Lorraine - INRIA — Nancy (FR)
Topic: Meta-learning with neural networks

EDUCATION

October 2014 – January 2018 — **PhD in Computer Science**, University of Lorraine — Nancy (FR)
Laboratory: LORIA (University of Lorraine, INRIA, CNRS)
Topic: Developmental Reinforcement Learning
Supervisors: Alain Dutech (Researcher INRIA), Yann Boniface (Associate Professor UL)

2012–2014 — **Master in Computer Science**, Pierre and Marie Curie University — Paris (FR)
Specialization: Artificial Intelligence and Decision - Intelligent Agents, Learning and Decision
Magna cum laude

SKILLS

Prog. Lang. C++, Java, Python, Octave

Machine Learn. PyTorch, Caffe, TensorFlow, scikit-learn

Clusters AWS, GCP, Grid5000, OAR

Languages French, English

DETAILED VERSION

EXPERIENCES

June 2018 – ongoing

Postdoctoral Researcher, UM-SJTU Joint Institute, *AAAL* — Shanghai - China

Supervisor: Paul Weng (Assistant Professor UM-SJTU Joint Institute)

Publications: Zimmer and Weng (2019b), Zimmer and Weng (2019a), Lin et al. (2019), Siddique et al. (2020), Lin et al. (2020), Huang et al. (2020), Huang et al. (2021), Zimmer* et al. (2021), Zimmer et al. (2021)

- Development and analysis of deep reinforcement learning algorithms (actor-critic, fairness, exploitation of symmetries, multi-agent, interpretability).
- Collaboration with Huawei on a neural logic reinforcement learning project.
- Collaboration with Prof. Juan Rojas of the Chinese University of Hong Kong who leads a robotic team. We proposed a way to exploit symmetry naturally present in robotics problems to improve the data efficiency of goal-based reinforcement learning algorithms.
- Co-supervision of 3 PhD students, 3 Masters students and 12 undergraduate students.
- Member of the Program Committee of IJCAI 2019 and IJCAI-PRICAI 2020. Reviewer for ICLR 2019, ICDL 2019-2020, ACML 2020, ICML 2019, ICRA 2020-2021, IEEE RA-Letters 2020-2021, IROS 2020, NeurIPS 2020, IEEE TNNLS 2020, AAAI 2021.
  `https://publons.com/researcher/1407553/zimmer-matthieu/peer-review/`
- Setting up and manage the AAAL team cluster (15 hosts, 292 threads, 26 GPUs) with OAR, IPSEC, NFS, NIS, KVM, sharelatex, wiki, etc.

Tools: C++, Python, OpenAI Gym, OpenAI Baseline, PyTorch, Tensorflow, Roboschool, Octave.

2014 – 6 months

Research internship, ISIR - UPMC, *AMAC* — Paris - France

Supervisor: Stéphane Doncieux (Professor UPMC-ISIR)

Publication: Zimmer and Doncieux (2017)

Study on transfer learning with reward shaping methods within a framework of lifelong learning. Developmental and evolutionary approach.The principle was to first use a direct policy search in the sensorimotor space, i.e. with no pre-defined discrete sets of states nor actions, and then extract from the corresponding learning traces discrete actions and identify the relevant dimensions of the state to estimate the value function. Once this is done, the robot can apply reinforcement learning to be more robust to new domains and, if required, to adapt faster than a direct policy search.

Tools: C++, C, OpenGL, ODE, Git, Python, Scikit-learn, Octave, Sferes, Bash, LaTex, FANN, OAR.

2013 – 3 months

Research internship, LIP6, *DECISION* — Paris - France

Supervisors: Paolo Viappiani (CNRS-LIP6), Paul Weng (UPMC-LIP6)

Publication: Zimmer et al. (2014)

Bibliographical research, reading articles and state of the art on the integration of knowledge from an expert during reinforcement learning.

An agent (the "teacher") advises another one (the "student") by suggesting actions the latter should take, while learning a specific task in a sequential decision problem; the teacher is limited by a "budget" (the number of times such advice can be given). Implementation in C++ of a new idea : the teacher is also learning, he learns to give advice to propitious moments to the student. He is learning how to teach better. We provided experimental results with the Mountain car domain, showing how our approach outperforms the state-of-the-art heuristics.

Tools: C++, C, Git, Torcs, Latex, OAR.

Source Code: https://github.com/matthieu637/smile

2012 – 5 months

Research internship, INRIA - LORIA, *Cortex & Maia* — Nancy - France

Supervisors: Yann Boniface (Associate Professor UL), Alain Dutech (Researcher INRIA-LORIA), Nicolas Rougier (Researcher INRIA-LORIA)

Meta-learning in neural networks.

Deepening ideas developed in articles of consciousness and meta-representations with multilayer perceptrons. How can they judge their own performances and improve them. Introduction to research, neural networks, latex and python.

A first neural network was learning a classification task, while a second one, called higher-order network, learned to bet if the prediction of the first network was correct from its hidden layer neurons. The higher-order network was indeed capable of learning such information, which meant that it can predict when the first network was going to fail. Thus, we proposed several architectures to combine the two networks in order to increase the overall prediction quality of the first network.

Tools: Python, Latex, Git.
Source Code: https://github.com/matthieu637/anne

2010 (3 weeks) **Summer internship**, Mathieu Perrein France — Waldweistroff - France
C# and WPF development using Microsoft Visual Studio.

EDUCATION

October 2014 – January 2018 **PhD in Computer Science**, University of Lorraine — Nancy - France

Laboratory	LORIA (University of Lorraine, INRIA, CNRS)
Topic	Developmental Reinforcement Learning (Zimmer, 2018)
Supervisors	Alain Dutech (Researcher INRIA–LORIA), Yann Boniface (Associate Professor UL–LORIA)
Reviewers	Olivier Pietquin (Professor University of Lille–Deepmind), Olivier Sigaud (Professor UPMC–INRIA)
Examiners	Isabelle Debled (Professor University of Lorraine), Celine Teulière (Associate Professor Institut Pascal)

Tools: C++, Git, OAR, OpenAI Gym, LaTex, Python, Jenkins, Caffe, Octave, ODE, Scikit-learn, FANN, OpenGL.
Source Code: https://github.com/matthieu637/ddrl, https://github.com/matthieu637/lhpo

2012–2014 **Master in Computer Science**, Pierre and Marie Curie University — Paris - France

Specialization	Artificial Intelligence and Decision
Research Training	Intelligent Agents, Learning and Decision

Magna cum laude

2008–2012 **Bachelor in Computer Science**, University of Lorraine — Nancy - France
Magna cum laude

2005–2008 **High School Diploma in Sciences**, Lycée Charlemagne — Thionville - France
Specialization: Mathematics

TEACHING EXPERIENCES

2019–2023 **Qualification - Associate Professor** — France
Section 27 - Computer Science

2014–2017 **Teaching Assistant**, *Engineer school* ENSEM — Nancy - France

162 hours Algorithms and Programming in Python (2nd year integrated preparatory cycle).
Course Manager: Jean-Philippe Mangeot
Practical work on python imperative programming (pyzo IDE): searching, sorting and small games (Reversi, Connect Four, ...). I wrote several practical work subject about data structures, Dijkstra, artificial intelligence and networking. I developed a first server (in Java) to create gaming party between two students, so they could challenge their artificial intelligence agent in a tournament determining their grades. During the last year, instead of comparing their agent on small games, we decided that the students had to create autonomous trading agents. Thus, I developed a second server (also in Java) to simulate a stock market exchange. In both cases, the students simply had to interface with the server in python, so they could focus on developing their artificial intelligence. The source code is available at github.com/matthieu637/cpp-2a-info.

8 hours Collaboration and Programming in Java (2nd year integrated preparatory cycle).
I designed and did seminars on Linux command-line, git, continuous integration, object-oriented design, unit testing, threads and synchronization in Java. To let students practice collaborative development, I set up a Jenkins instance communicating with a Github project.

35 hours Algorithmic and Java Object-Oriented Programming (1st year of engineering school).
Course Manager: Vincent Chevrier
Practical work on search algorithms, object-oriented design and Lego robot navigation.

2015
20 hours

Training on Teaching, University of Lorraine — Nancy - France

SUPERVISION OF RESEARCH ACTIVITIES

PhD Students

2020
ongoing

Zhaohui Jiang, SJTU, UM-SJTU Joint Institute — Shanghai - China
With Paul Weng and Claire Glanois.
PhD student in our group since 2020: Interpretable deep reinforcement learning with neuro-symbolic approaches.
We made a submission to KR 2021 (Zimmer et al., 2021).

2020
ongoing

Xuening Feng, SJTU, UM-SJTU Joint Institute — Shanghai - China
With Paul Weng and Claire Glanois.
PhD student in our group since 2019: Model-based and interpretable deep reinforcement learning.
We made a submission to KR 2021 (Zimmer et al., 2021).

2020
ongoing

Junqi Qian, SJTU, UM-SJTU Joint Institute — Shanghai - China
With Paul Weng.
PhD student in our group since 2018: Regularization methods and in deep reinforcement learning.
Junqi is also applying our algorithm (Zimmer and Weng, 2019b) to more domains while tacking into account the cost for optimizing of the hyperparameters.

Master Students

2020
ongoing

Jianyi Zhang, SJTU, UM-SJTU Joint Institute — Shanghai - China
With Paul Weng.
Master student in our group since 2019: Safe deep reinforcement learning with constraints.
We made a submission to KR 2021 (Zimmer et al., 2021).

2019-2021

Jiancong Huang, GUT and SJTU, UM-SJTU Joint Institute — China
With Paul Weng and Juan Rojas.
Master student in Mechanical Engineering: Hyperparameter Auto-tuning in Self-Supervised Robotic Learning.
We published a journal article in IEEE Robotics and Automation Letters (Huang et al., 2021) and our preliminary results in a NeurIPS workshop (Huang et al., 2020).

2019-2020

Yijiong Lin, GUT and SJTU, UM-SJTU Joint Institute — China
With Paul Weng and Juan Rojas.
Master student in Mechanical Engineering : Invariant Transform Experience Replay.
We propose to exploit the symmetries present in robotic tasks to improve the data efficiency of deep reinforcement learning algorithms. Experiments are performed in the robotic Fetch tasks from OpenAI Gym showing that we attain a 3 to 13 times speedup compared to the baseline.
We published a journal article in IEEE Robotics and Automation Letters (Lin et al., 2020) and our preliminary results in a NeurIPS workshop (Lin et al., 2019).
Yijiong Lin is now pursing a PhD degree with Prof. Nathan Lepora (University of Bristol-United Kingdom).

2018
ongoing

Umer Siddique, SJTU, UM-SJTU Joint Institute — Shanghai - China
With Paul Weng.
Master student in our group since 2018: Fair optimization in deep reinforcement Learning.
We published in ICML 2020 (Siddique et al., 2020) and ICML 2021 (Zimmer* et al., 2021) for the single-agent case and multi-agent case respectively.
Umer Siddique is willing to pursue a PhD with our group.

2017
5 months

Achille Fedioun, University of Lorraine, LORIA — Nancy - France
With Alain Dutech and Yann Boniface.
End-of-studies internship (Master Computer Science and Engineer school): Reinforcement learning with continuous state and action spaces using model-free actor-critic algorithms with deep neural networks.
Achille had to compare the features of two new algorithms (Qprop and ACER) with ours. He extended one of our actor-critic agents with off-policy multi-step replay using the Retrace algorithm in C++. As experimental validation, he used the cluster of the lab to train deep neural networks on the half-cheetah environment.

2015
5 months

Nicolas Lefebvre, University of Lorraine, LORIA — Nancy - France

With Alain Dutech and Yann Boniface.
End-of-studies internship (Master Cognitive Science): Reinforcement learning with continuous state and action spaces using model-free actor-only algorithms.
Nicolas had to explore if the Power algorithm could be used with neural networks instead of dynamic movement primitives. He developed the Power algorithm inside our C++ framework using Gaussian mixture policies. He experimentally validated his agent on the acrobot environment (double inverted pendulum).

Undergraduate Students

2018-2020

Research Projects of Undergraduate Students, SJTU, UM-SJTU JI — Shanghai - China

With Paul Weng.

- 5 students, Fall 2019-Spring 2020, 8 months, Experimental Evaluation of deep reinforcement learning algorithms on HPC over Atari games and PyBullet environments
  (PRP - Chenmin Hou, Zhengjie Ji, Shuhyi Zhu, Siwei Ye and Run Peng)
- 1 student, Summer 2019, 4 months, Improving DQN with dynamic discount factor
  (VE490 - Xinyang Ren)
- 1 student, Summer 2019, 4 months, Displaying the landscape of deep neural networks for deep reinforcement learning
  (VE490 - Yifei Zhang)
- 4 students, Spring 2019, 4 months, Deep reinforcement learning for UAV control
  (PRP - Xingyue Qian, Yunfan He, Chen Zhikai and Gaopeng Song)
- 1 student, Spring 2019, 4 months, Model-based reinforcement learning with PILCO on Roboschool
  (VE490 - Zhenyuan Zhang)

COMPUTER SKILLS

Languages C++, Java, Python, Octave, C, Prolog, C#, OCaml

Web J2EE, PHP, JavaScript, AJAX, HTML, CSS, WordPress, Laravel

Libraries Boost, SFML, CEGUI, GLib, Apache Commons, JFlex, Java CUP, Jenkins, OpenCV

Machine Learn. Caffe, PyTorch, TensorFlow, scikit-learn, OpenAI Baselines

Simulators ODE, TORCS, OpenAI Gym, OpenAI Roboschool

Storage PostgreSQL, Oracle, MySQL, SQLite, XML (Schema, DTD, XPath)

Utilities KDevelop, Eclipse, NetBeans, Microsoft Visual Studio, CodeBlocks, LaTeX, Git, PyCharm, Spyder

OS. ArchLinux, Debian, Ubuntu

Other Computer cluster (Grid5000, AWS, GCP) around 300 years of computations, Shell Bash, Csh, UML, Lua

LANGUAGES

French Mother tongue

English Fluent

INTERESTS

Development Isometric 2D game in team, Dynamic website with applet-server, Server management

Others Free software, Self Hosting, Hardware and Blockchains

Awards Travel Grant to attend the IRCN Course in Neuro-Inspired Computation at Tokyo 2019
NeurIPS 2019: Learn to Move (top 20) - 1100$ GCP credits

Sport Badminton (8 years)

PUBLICATIONS

International Journals

Jiancong Huang, Juan Rojas, **Matthieu Zimmer**, Hongmin Wu, Yisheng Guan, and Paul Weng. "Hyperparameter Auto-tuning in Self-Supervised Robotic Learning". In: *IEEE Robotics and Automation Letters* . 2021.
PDF

Yijiong Lin, Jiancong Huang, **Matthieu Zimmer**, Juan Rojas, and Paul Weng. "Invariant Transform Experience Replay". In: *IEEE Robotics and Automation Letters* . 2020.
PDF Video

Matthieu Zimmer and Stephane Doncieux. "Bootstrapping Q-Learning for Robotics from Neuro-Evolution Results". In: *IEEE Transactions on Cognitive and Developmental Systems* . 2017.
PDF Video

International Conferences

Matthieu Zimmer*, Claire Glanois*, Umer Siddique, and Paul Weng. "Learning Fair Policies in Decentralized Cooperative Multi-Agent Reinforcement Learning". In: ***International Conference on Machine Learning***. 2021.
PDF

Umer Siddique, Paul Weng, and **Matthieu Zimmer**. "Learning Fair Policies in Multi-Objective Deep Reinforcement Learning with Average and Discounted Rewards". In: ***International Conference on Machine Learning***. 2020.
PDF

Matthieu Zimmer and Paul Weng. "An Efficient Reinforcement Learning Algorithm for Learning Deterministic Policies in Continuous Domains". In: *Distributed Artificial Intelligence*. Sept. 2019.
PDF

Matthieu Zimmer and Paul Weng. "Exploiting the sign of the advantage function to learn deterministic policies in continuous domains". In: ***International Joint Conferences on Artificial Intelligence***. Aug. 2019.
PDF Slides Poster

Matthieu Zimmer, Yann Boniface, and Alain Dutech. "Developmental Reinforcement Learning through Sensorimotor Space Enlargement". In: *The 8th Joint IEEE International Conference on Development and Learning and on Epigenetic Robotics*. Sept. 2018.
PDF Slides Video Blog

Matthieu Zimmer, Yann Boniface, and Alain Dutech. "Neural Fitted Actor-Critic". In: *ESANN - European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. Apr. 2016.
PDF Poster

Work in progress

Claire Glanois, Paul Weng, **Matthieu Zimmer**, Li Dong, Tianpei Yang, Hao Jianye, and Liu Wulong. "Survey on Interpretable Reinforcement Learning". In: . 2021.
In submission to Journal of Artificial Intelligence Research.

Matthieu Zimmer, Xuening Feng, Claire Glanois, Zhaohui Jiang, Jianyi Zhang, Paul Weng, Hao Jianye, Li Dong, and Liu Wulong. "Differentiable Logic Machines". In: 2021. arXiv: 2102.11529.
In submission to KR 2021.

International Workshops

Jiancong Huang, Juan Rojas, **Matthieu Zimmer**, Hongmin Wu, Yisheng Guan, and Paul Weng. "Hyperparameter Auto-tuning in Self-Supervised Robotic Learning". In: *Deep Reinforcement Learning Workshop - NeurIPS*. 2020.

PDF Video

Yijiong Lin, Jiancong Huang, **Matthieu Zimmer**, Juan Rojas, and Paul Weng. "Towards More Sample Efficiency in Reinforcement Learning with Data Augmentation". In: *Robot Learning: Control and Interaction in the Real World - NeurIPS workshop*. Dec. 2019.
PDF

Matthieu Zimmer, Yann Boniface, and Alain Dutech. "Off-Policy Neural Fitted Actor-Critic". In: *Deep Reinforcement Learning Workshop, NIPS 2016*. Dec. 2016.
PDF Poster

Matthieu Zimmer, Yann Boniface, and Alain Dutech. "Toward a data efficient neural actor-critic". In: *European Workshop on Reinforcement Learning*. Dec. 2016.
PDF Poster

Matthieu Zimmer, Paolo Viappiani, and Paul Weng. "Teacher-student framework: a reinforcement learning approach". In: *AAMAS Workshop Autonomous Robots and Multirobot Systems*. 2014.
PDF Slides

National Conferences

Matthieu Zimmer, Yann Boniface, and Alain Dutech. "Vers des architectures acteur-critique neuronales efficaces en données". In: *Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes*. Aug. 2016.
PDF Slides Video

Theses and Various Reports

Matthieu Zimmer. "Apprentissage par renforcement développemental". PhD thesis. University of Lorraine, Jan. 2018.
PDF Slides Video

Matthieu Zimmer. "Construction Automatique d'état et d'actions en Apprentissage par Renforcement". MA thesis. University Pierre and Marie Curie, 2014.
PDF Slides Video

Matthieu Zimmer, Yann Boniface, Alain Dutech, and Nicolas Rougier. "Dans quelle mesure un système apprenant peut prendre conscience de ses performances et altérer son comportement". Research Report. 2012.
PDF

SOFTWARES AND REPRODUCIBLE RESEARCH

DLM **Differentiable Logic Machines**
Duration $\geq$ 10 months, Url=`https://github.com/matthieu637/dlm`
Python 100%
290 commits (hidden to public), 50 branches, around 18000 lines of code, contribution: 85%

Interpretable reinforcement learning with first-order logic policies in Pytorch.
Publication: Zimmer et al. (2021).

DFRL **Distributed Fair deep Reinforcement Learning**
Duration = 1 year, Url=`https://github.com/matthieu637/dfrl`
Python 95%, Shell 5%
228 commits (hidden to public), around 4000 lines of code, contribution: 99%

Deep reinforcement learning libray for fair policies in multi-agent scenarios with Tensorflow.
Publication: Zimmer* et al. (2021, p. ICML 2021).

DDRL **Deep Developmental Reinforcement Learning**
Duration = 4 years, Url=`https://github.com/matthieu637/ddrl`
C++ 86.7%, C 6.2%, CMake 3.4%, Python 2.5%, Shell 1.2%
842 commits, around 50000 lines of code, contribution: 99%

Deep reinforcement learning library in Caffe with new environments.
Publications: Zimmer and Weng (2019b, IJCAI 2019) Zimmer et al. (2018, 2016a).

LHPO **Lightweight asynchronous and distributed hyperparameter optimization**
Duration $\geq$ 6 years, Url=`https://github.com/matthieu637/lhpo`
Bash 60.7%, Octave 35.8%, Python 3.5%
182 commits, around 3500 lines of code, contribution: 100%

This software is used most of the graduate students in our group.
It has been used in the following works: Zimmer et al. (2021), Zimmer* et al. (2021), Siddique et al. (2020), Zimmer and Weng (2019a,b), Zimmer et al. (2018), Zimmer and Doncieux (2017), and Zimmer et al. (2016a)

SMILE **SeMi-supervIsed Learning agEnt**
Duration = 1 year, Url=`https://github.com/matthieu637/smile`
C++ 51.5%, TeX 22.3%, C 16.8%, Octave 5.0%, Shell 3.1%, CMake 1.3%
154 commits, around 7000 lines of code, contribution: 100%

Publication: Zimmer et al. (2014).

ACRONYMS

AAAI Association for the Advancement of Artificial Intelligence
AAAL Artificial Agent Autonomous Learning (Team name)
ACML Asian Conference on Machine Learning
AMAC Architectures et Modèles pour l'Adaptation et la Cognition (Team name)
CNRS National Center for Scientific Research, France
DQN Deep Q Network (Reinforcement Learning Algorithm)
GUT Guangdong University of Technology
HPC High-Performance Computing
ICDL International Conference on Development and Learning and on Epigenetic Robotics
ICLR International Conference on Learning Representations
ICML International Conference on Machine Learning
ICRA International Conference on Robotics and Automation
IJCAI International Joint Conference on Artificial Intelligence
INRIA National Institute for Research in Computer Science and Automation, France
IRCN International Research Centre for Neuro-intelligence
ISIR Institut des Systèmes Intelligents et de Robotique, France
JI Joint Institute
KR International Conference on the Principles of Knowledge Representation and Reasoning
LIP6 Laboratoire Informatique de Paris 6, France
LORIA Laboratoire Lorrain de Recherche en Informatique et ses Applications, France
NeurIPS Neural Information Processing Systems
ODE Open Dynamic Engine (Physic Engine)
TNNLS Transactions on Neural Networks and Learning Systems
UL University of Lorraine
UM-SJTU University of Michigan-Shanghai Jiao Tong University
UPMC University Pierre and Marie Curie, France

Last update June 29, 2021