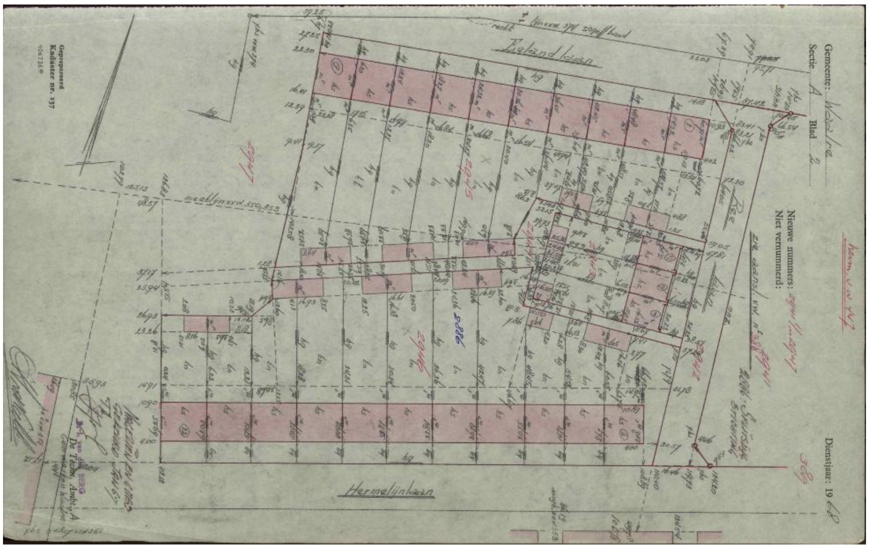
INTERNSHIP: LINE AND MEASUREMENT DETECTION IN FIELD SKETCHES USING DEEP LEARNING

Keywords: Image analysis, Deep Learning, Python, Tensorflow, Keras, Python.

Sioux LIME recently finished a feasibility study for the Dutch Kadaster. In this project we have primarily looked at extracting various information elements (in an automated manner) from so-called field sketches (“veldwerken”).



Example field sketch.

In this internship you further investigate the use of machine learning and more specifically deep learning for specific information extraction tasks. In more detail we would like to focus this internship on algorithm development (and implementation) for two tasks: (1) measurement number detection and (2) line detection. We have sufficient data available, though not labeled (yet).

We are particularly interested in using deep learning models since we believe this has clear advantages in this case, so using this type of models is a requirement. However, such a model can also form a component of a meta-model or combined with rules/heuristics to achieve the goals.

On a field sketch there are many measurement numbers written. These numbers are typically written close and orthogonal to a line. The reading direction is at an arbitrary angle, which complicates final interpretation. We believe that detection and de-rotation of the measurement numbers should be done before feeding them to an off-the-shelf OCR/ICR tool.

For line detection, we have implemented an advanced heuristic ourselves, not using machine learning at all. Can we also do this using machine learning and how to best do this? There are different line types that should be distinguished also. We are very interested in also exploring this within this internship. We have found literature on detecting road networks in aerial imagery and believe that these techniques might successfully be modified and applied to this case.

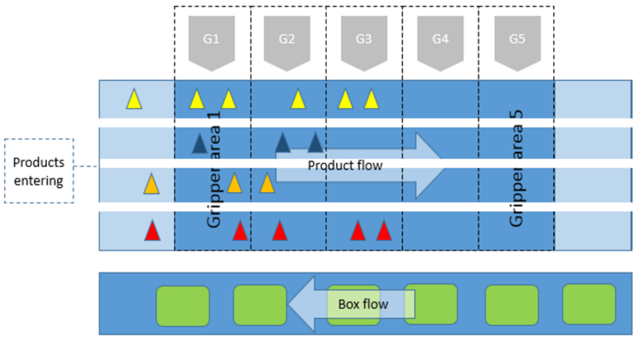
In terms of technology to use in this internship: Python combined with Jupyter Notebooks for most of the programming. For Machine Learning, Keras and Tensorflow are deep learning frameworks that will likely be used. We are also interested in exploring Pytorch.

Interested? Please contact us via stage@siouxlime.nl.

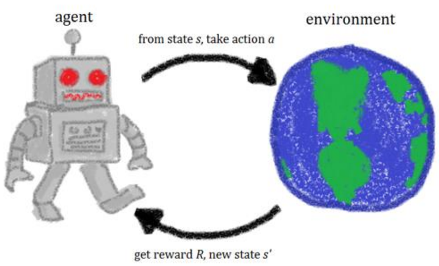
INTERNSHIP: SCHEDULING USING REINFORCEMENT LEARNING

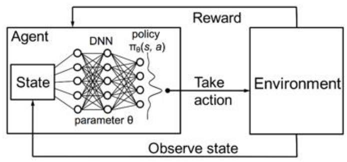
Keywords: Reinforcement Learning, Scheduling, Optimization, Markov Decision Processes, Deep Learning, Python, Tensorflow

LIME has carried out two packaging machine optimization projects for a specific client. This concerned the optimization of two slightly different machine types. The setup of such a packaging machine is that both the packaging boxes and the products flow on conveyor belts. Alongside the belts, robots are placed that can pick and place products. The algorithm controls the actions of the robots, determining in which order the products should be packed and which product should go in which box.



After having developed two distinct solutions, the natural question comes up: can we develop a more generic and potentially even better solution to this kind of scheduling problems? A single solution framework that can easily handle a wide range of different machine instances. Using Reinforcement Learning to tackle these scheduling problems could be part of such a solution. Another alternative could be to use an integer programming approach.





In a recent internship, we took the first steps towards applying reinforcement learning to this type of scheduling problems. We got some first insights and successes, but were not yet able to arrive at a trained model capable of near-optimal scheduling of a realistic case. We believe this is possible and in this internship you will aim for closing the gap.

As such, you will first familiarize yourself with the reinforcement learning basics and the work that has been done so far. The work done so far consists of a Tensorflow implementation of the famous A3C algorithm within the OpenAI Gym with bespoke choices for the state and action space representation as well as the neural network architecture. Going forward, it makes sense to consider re-using or building on the previous work, but we are open to new implementations. As such, you are free to develop and implement whatever you think will work best to learn the agent how to best pick-and-place the products in boxes.

Interested? Please contact us via stage@siouxlime.nl.

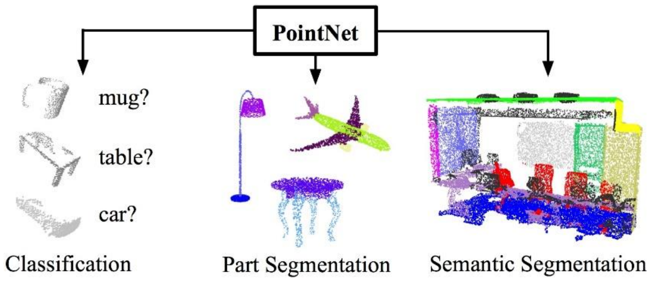
INTERNSHIP: 3D DEEP LEARNING

Keywords: Deep Learning, 3D data, Computer Vision, Segmentation, Object detection, Python, Tensorflow

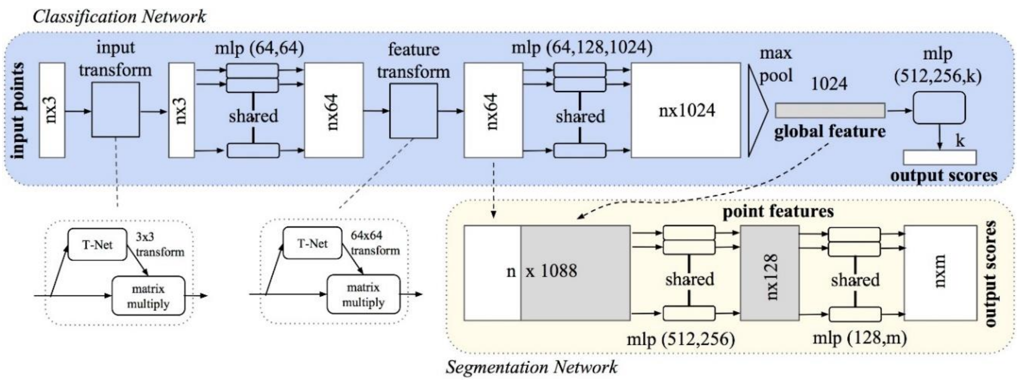
Deep Learning has made huge steps over the last years and is successfully applied in different application domains. Computer vision is probably the domain with the most successful deep learning applications so far. The state-of-the-art is still progressing at a high-pace.

In this internship we explore the state-of-the-art in deep learning for 3D image data. Machine learning techniques for this type of data have great potential as our world is inherently three-dimensional and even four-dimensional when considering the temporal domain.

3D data can be represented in different formats such as multi-view RGB(D) images, volumetric, polygonal mesh, point cloud or primitive-based CAD models. As a special case, we also want to look into so-called 2.5D data (also called "2D+Z"). A related question is also how one effectively analyzes 4D data (3D plus time dimension). The extra dimension introduces large computational and memory overhead that quickly is a bottleneck for applying deep learning successfully to these type of data.



Specific tasks for 3D Deep Learning that we are interested in are classification, object detection and semantic segmentation. Recent literature suggests various deep learning based approaches: volumetric CNNs, multi-view CNNs, spectral CNNs, feature-based DNNs, point cloud DNNs, ...



After having familiarized yourself with deep learning, your goals for this internship are:

- Explore and understand the state-of-the-art models for 3D Deep Learning.
- Study the pros and cons of suggested approaches on a number of aspects.
- Apply state-of-the-art models to datasets (training and testing).
- Potentially: develop a new model or improve an existing model.

There are several open datasets at your disposal that you can work on. Technology-stack to be used: Python, Jupyter Notebook, TensorFlow, Keras and PyTorch.

Interested? Please contact us via stage@siouxlime.nl.

Internship proposals 2018

Lense Swaenen (lswaenen@limebv.nl)

Real-time non-linear optimization of time-optimal trajectories

Keywords: Optimization, Automatic differentiation

The Eindhoven region hosts many manufacturers of high-tech machines. To have a high throughput, components of such machines need to make very precise movements as fast as possible. The limiting factors are often the dynamic constraints on the actuators, such as maximum velocity, acceleration and higher order derivatives. Completing a movement subject to such constraints as fast as possible is a challenging mathematical optimization problem. For the simple 1D case, there is a simple explicit solution which can be found based upon the well-known bang-bang principle. For simultaneous movements of actuators (like an XY-stage), the problem quickly becomes much more difficult.

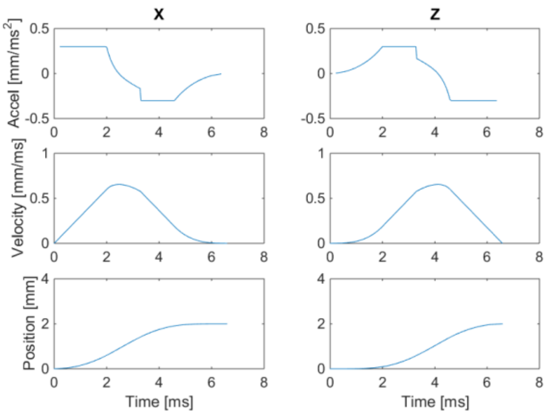


Figure 1: Example of a time-optimal circular trajectory

In this internship, we tackle these types of problems by developing NLP solvers tuned to this problem structure from scratch. In prior work, the problem has been solved using generic NLP solvers like IPOPT in frameworks like CasADi and JuMP, but these have proven to be too slow for the application, as IPOPT is designed for large sparse problems rather than small dense problems with real-time constraints. Development will be a combination of Python and C++ (if speedup is necessary).

Time-optimal trajectories through reinforcement learning

Keywords: Reinforcement learning, TensorFlow

The Eindhoven region hosts many manufacturers of high-tech machines. To have a high throughput, components of such machines need to make very precise movements as fast as possible. The limiting factors are often the dynamic constraints on the actuators, such as maximum velocity, acceleration and higher order derivatives. Completing a movement subject to such constraints as fast as possible is a challenging mathematical optimization problem. For the simple 1D case, there is a simple explicit solution which can be found based upon the well-known bang-bang principle. For simultaneous movements of actuators (like an XY-stage), the problem quickly becomes much more difficult.

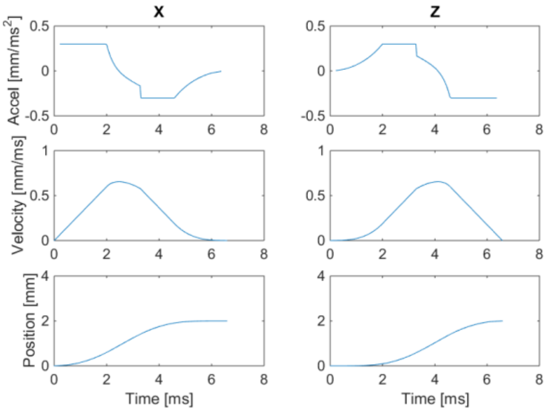


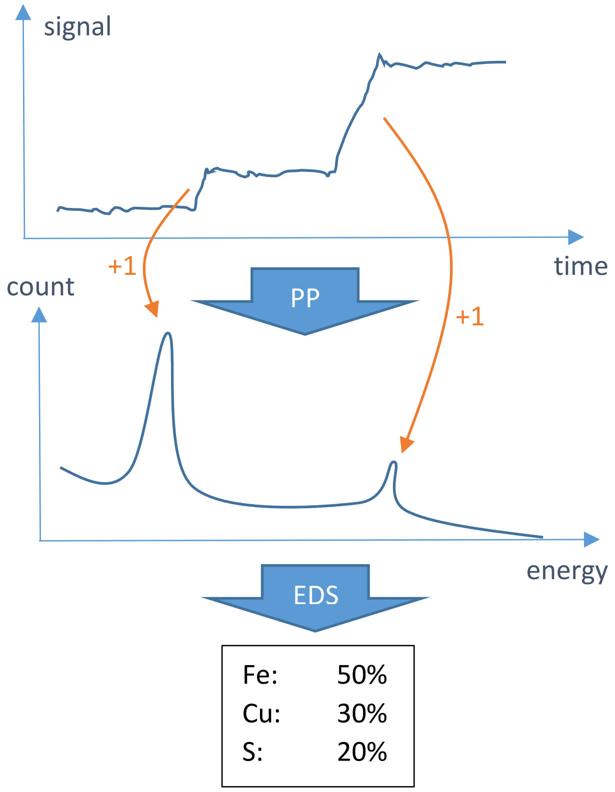
Figure 2: Example of a time-optimal circular trajectory

Solving these problems using non-linear optimization techniques easily results in computation times that are too high for real-time application.

In this internship you will develop a Reinforcement Learning approach to this kind of time-optimal control problems. The primary goal is calculating fast trajectories in a limited amount of computation time (excluding the learning phase). Reinforcement Learning is an area of machine learning that has been used as the core of the AlphaGo and AlphaGo Zero programs. One or multiple agents take actions in an environment with the goal of maximizing the total cumulative reward. These agents in practice are often neural networks that are being trained by learning. The hypothesis behind this internship is that if we can learn an agent to design time-optimal trajectories for a set of training benchmark cases, that it can also solve the validation benchmarks well, and moreover very fast. To develop the methodology, the TensorFlow library and Python programming language will be used.

Spectrum-free spectroscopy

Keywords: Statistical processes, physics, inverse problems

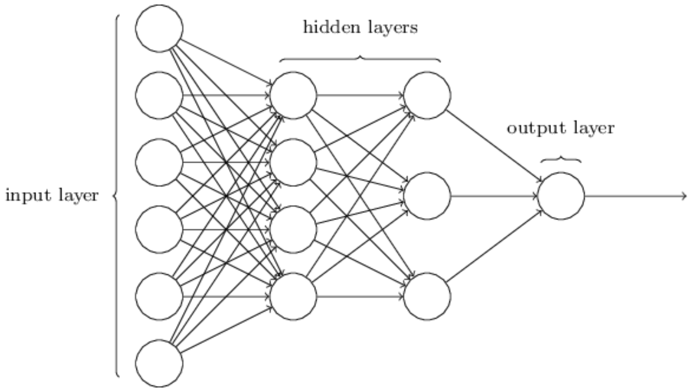


Energy-dispersive X-ray spectroscopy (EDS) is a class of techniques that allow extraction of elemental information from the 'byproduct' X-rays that are generated by electron microscopes. As a sample is being bombarded with electrons, X-rays are produced which are registered by a detector which builds up a spectrum (counts as a function of energy). Such spectra contain characteristic peaks which indicate presence of iron, for example. In the building of this spectrum, a low-level signal processing step, called the 'pulse processor' (PP) detects steps in a signal and adds counts to a bin with respects to a step height. Any information which could have been obtained about the accuracy of the step detection is thrown away.

In this internship, we propose to devise a solution in which we avoid the intermediate step of spectrum-building and modify standard EDS methods to work with event-type data and try to have benefits from doing so in terms of spectral resolution and accuracy. Though the application is quite physical, the techniques used will be related to simulation of Poisson processes and maximum-likelihood-estimation-type parameter fitting.

INTERNSHIP: NEURAL NETWORKS FOR ENERGY DISPERSIVE X-RAY SPECTROSCOPY

Keywords: Neural networks, Inverse problems, Tensorflow



Sioux LIME has developed for one of its clients a methodology to extract elemental compositions from x-ray spectra. The used approach uses a forward model and an inverse solver. The forward model predicts spectra when given an elemental composition. The inverse solvers solve the corresponding optimization problem. As the forward model is made progressively more complex (including higher order physical effects and constraints), the requirements for the inverse solvers become more difficult to satisfy.

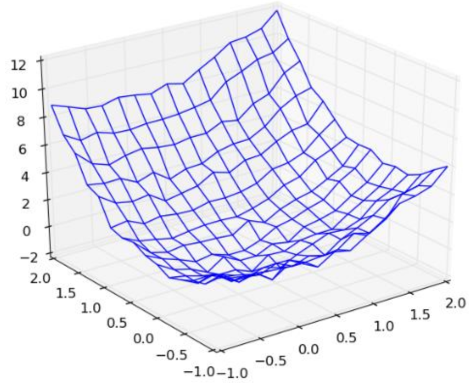
In this internship, we investigate and tune the performance of a solution in which the inverse solver is replaced with a trained neural network. The neural network will be trained using artificially generated ground truths using the forward model. The Tensorflow toolkit is recommended for this exercise.

Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

INTERNSHIP: OPTIMIZATION OF NOISY COST FUNCTIONS

Keywords: Numerical optimization, Python



A common optimization problem is the parameter fitting of a physical model to measurements. We consider the case where the physical model has complex/involved computations which are iterative in nature (differential equation solving, some optimization, …) and which are terminated well before numerical precision, more typically at 3-4 digits. This makes use of gradient-based optimizers troublesome. On the other hand, global optimizers don't quite exploit the specific structure of the problem, which is often 'Rosenbrock-like' (e.g. having a single global optimum). As the forward model is already computationally heavy, putting a slow optimizer around it is undesireable.

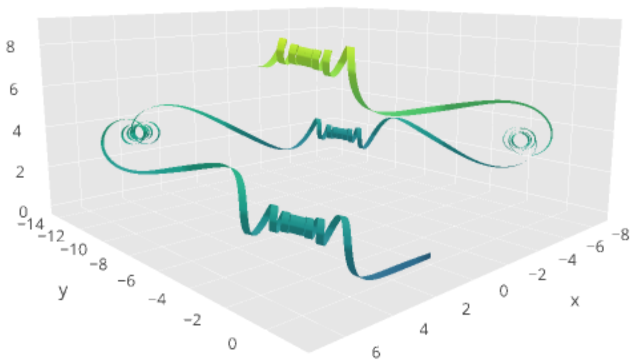
We consider a practical application coming from the bending manufacturing industry. The Nelder-Mead method has been found to perform reasonably well. We expect a more advanced method (e.g. using quadratic interpolation), may still give significant improvements. Investigation and development of such a method is the goal of the internship.

Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

INTERNSHIP: EXTENDING THE FRENET-SERRET EQUATIONS TO RIBBONS

Keywords: Differential geometry, Integration of differential equations, Computer graphics



Sioux LIME is developing algorithms and tooling for a bending machine manufacturing company. Currently, the use case under consideration is the bending of round (a priori straight) metal pipes into arbitrary shapes. Because of the circular cross-section, representation of a pipe by a curve (0D cross-section) is sufficient. A square tube (not necessarily free-form-bendable by the clients current machinery) would get an extra degree of freedom, such that a curve representation should be replaced by a 'ribbon' representation. No theoretical framework (extending the Frenet-Serret framework) is known for such objects.

In this internship, we try to develop a theoretical framework for such geometrical objects. Additionally, we develop tooling to visualize and manipulate such shapes.

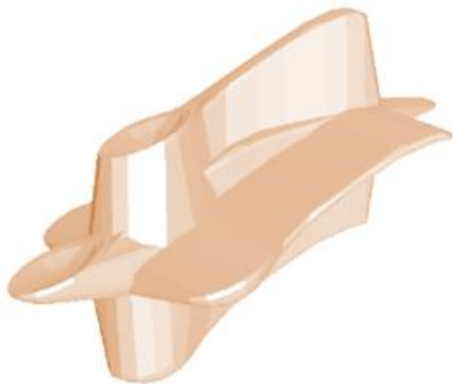
Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

INTERNSHIP: CIRCULAR EXTRUSION OF 3D MESHES

Keywords: Computational geometry, collision detection

Sioux LIME has develop collision detection methodologies for safe navigation of stages in a vacuum chamber. A particular feature involving 'blind-folded' return to a reference position can be solved with extrusion of meshes. For linear/translating axes, a linear extrusion operation has already been developed. For rotating axes, a circular extrusion would be required, which is a much more challenging problem. Furthermore, it may be desirable for the extrusion to be able to run on-line, and thus be fast/dynamic. In this internship, we develop the methodology which performs this circular extrusion



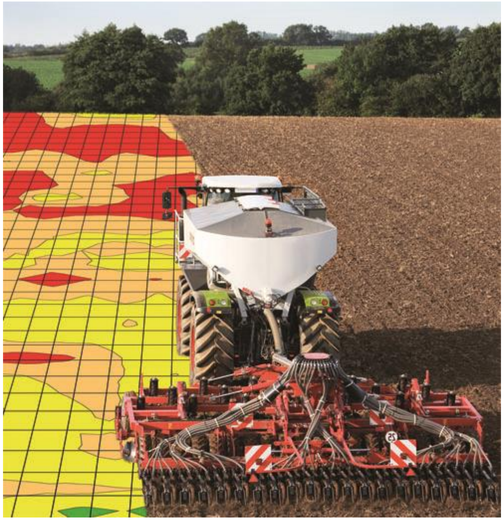
Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

SOURCE OF YOUR TECHNOLOGY

INTERNSHIP: PRECISION FARMING

Keywords: data analysis, visualization, data processing, R, multi-sensor data, indirect measurements



Precision farming is a farming management concept based on observing, measuring and responding to inter and intra-field variability in crops. By using data from various sources, a farmer tries to maximize the yield of a field using as few resources as possible. This can be achieved by using high precision positioning systems (up to 2cm accuracy) on agricultural vehicle: the work place of a farmer gradually shifts from field to office. Example of data sources vary from pictures taken from satellites/drones (remote sensing), information on soil (water content, soil map, …), last year’s harvest quantities and measured amount of biomass. As of this moment, these data sources are unconnected and, due to a lack of standards, difficult to process.

In this internship you will extract as much value as you can from the data sources available for one specific field. As an experienced farmer is more than able to translate the insights from the data to actions (e.g. add fertilizer), an important part of the data’s value is a convenient way to import and visualize the data over time. Eventually you might even be able to predict this year’s harvest based on the intermediate measurements.

Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

INTERNSHIP: OPTIMIZATION FOR AUTOMATED PACKAGING

Keywords: Optimization, Scheduling

Sioux LIME has developed for one of its clients a solution that concerns real-time optimization algorithms for the automated packaging of products by robots. Both the packaging boxes and the different types of products flow on conveyor belts, leaving a limited amount of time for the robots to perform packaging each product. The algorithm controls the actions of the robots, determining in which order the products should be packed and which product should go in which box.

Typically, given the machine setup, the objective function is to minimize the number of products that leave the conveyor belt unpacked. There are also several constraints that should be taken into account, for instance, each box should contain the correct number of each type of product. An important other type of constraint is that everything should perform real-time. This makes that the algorithm should be fast and even faster with an increased speed of the conveyor belts. There is obviously a commercial incentive to use higher belt speeds.

Our developed solutions perform a last-minute optimization, taking into account (only) all the products that are currently on the conveyor belt. However, the flow of products is typically quite predictable. How can we use this to better optimize the packaging? This requires rethinking the current optimization approach. A second goal is to design a(n) (meta) algorithm that solves a more generic case. For instance, up to now LIME has developed two algorithms for two different machine types and these two algorithms partially differ. As an illustration of a more generic algorithm, imagine that the algorithm can handle any specified number of robots. As a nice-to-have, we would like to obtain a bound on the performance of the algorithm. This would enable us to tell how far the algorithm is maximally off the (in hindsight) optimal solution.

Overall, the goal of this intern project is to design and implement an appropriate method to optimize a(n) (ideally generic) packaging machine aiming to also make use of product inflow predictability.

Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

SOURCE OF YOUR TECHNOLOGY

INTERNSHIP: SCHEDULING USING REINFORCEMENT LEARNING

Keywords: Reinforcement Learning, Scheduling, Optimization, Markov Decision Processes, Approximate dynamic programming, Deep Learning

LIME has carried out two packaging machine optimization projects for a specific client. This concerned the optimization of two slightly different machine types. The setup of such a packaging machine is that both the packaging boxes and the products flow on conveyor belts. Alongside the belts, robots are placed that can pick and place products. The algorithm controls the actions of the robots, determining in which order the products should be packed and which product should go in which box.

After having developed two distinct solutions, the natural question comes up: can we develop a more generic and potentially even better solution to this kind of scheduling problems? A single solution framework that can easily handle a wide range of different machine instances. Using Reinforcement Learning to tackle these scheduling problems could be part of such a solution. Another alternative could be to use an integer programming approach.

In this assignment you will develop a Reinforcement Learning approach to this kind of scheduling problems. Reinforcement Learning is an area of machine learning. The context is one of agents taking actions in an environment with the goal to maximize the total cumulative reward. The reinforcement learning framework is quite general and can be applied to many problems in different domains. There is also a connection with deep learning. Quite often, the problem size is too large, so that one has to move to approximation methods. Deep learning can be used e.g. to approximate the value function in a high-dimensional state space.

Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

INTERNSHIP: ANALYZING SENSOR DATA USING DYNAMIC BAYESIAN NETWORKS

Keywords: Sensor fusion, Bayesian Networks

Dynamic Bayesian Networks form a subclass of Probabilistic Graphical Models (PGMs). PGMs use a graph-based representation as the foundation for encoding a complete distribution over a multi-dimensional space and a graph that is a compact or factorized representation of a set of independences that hold in the specific distribution.

PGMs form a very general framework for building probabilistic models with applications in many different directions. PGMs form a key element of so-called model-based machine learning. This is a paradigm that uses a Bayesian viewpoint together with PGMs and generic inference algorithms to build all kind of specific machine learning algorithms. The clean split between model (specification/building) and inference is one of the key benefits.

Dynamic Bayesian Networks (DBNs) can be used to model systems that have a temporal aspect and vary over time. As such they are a very natural candidate for modelling and analyzing systems that are measured by (multiple) sensors (of different modalities) over time.

Typically, the complexity is in doing inference for a specific (type of) DBN. There exist a few rather generic inference engines that can be used, but they might not give fast and accurate results. As such, it might be required to build/develop new inference algorithms or adopt existing ones tailored to the network at hand.

We have several datasets available, but we can also generate (artificial) data and see if and how we can get good inference results for these datasets (where we now the truth). In this assignment you will be building and analyzing specific DBNs.

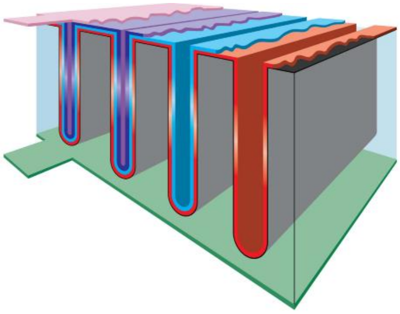
Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

GRADUATION INTERNSHIP: FREE SURFACE DEPOSITION SIMULATION

Keywords: Finite element method, Free boundary problem, Time integration, Topological changes, Python.

Chemical Vapour Deposition (CVD) is a common manufacturing technique for adding a thin layer of material to a product. For example, to achieve a high energy density in micro-batteries, anodes and cathodes are deposited in thin layers on deep trenches in a silicon substrate. Also , coating the skeleton of an aluminium foam with a nickel layer, creates a very stiff but light-weight structure, with applications in mechanical and biomedical engineering. Our primary concern is for graphite wafer carriers, on which a SiC layer is deposited to enhance stiffness, temperature conduction and to seal the porous core.



3D micro-battery schematic

The above mentioned processes all involve the diffusion of ions in a highly complex geometry. This geometry accrues material in amounts significant enough to change the topology of the flow domain. In order to investigate CVD on the microscale, you will develop a numerical model satisfying the following requirements:

1. The diffusion of the ions should be captured accurately on “arbitrarily complex'” geometries. Geometries will be derived from a μCT scan, so the topology is arbitrary, but the scales are limited by resolution.
2. The evolution of the deposition boundary allows for topological changes. In practice, pores in the flow domain can clog up and eventually become sealed from the ion source.
3. The flow domain can be constrained to decrease monotonically. When material is deposited in an area of the void, it will remain there.

Interested? Please contact us via stage@siouxlime.nl.

Sioux LIME BV
Esp 405
5633 AJ Eindhoven
The Netherlands
Tel: +31 (0)40 751 61 16
ww.siouxlime.nl
stage@siouxlime.nl

Internship proposals 2018

Lense Swaenen (lswaenen@limebv.nl)

Real-time non-linear optimization of time-optimal trajectories

Keywords: Optimization, Automatic differentiation

The Eindhoven region hosts many manufacturers of high-tech machines. To have a high throughput, components of such machines need to make very precise movements as fast as possible. The limiting factors are often the dynamic constraints on the actuators, such as maximum velocity, acceleration and higher order derivatives. Completing a movement subject to such constraints as fast as possible is a challenging mathematical optimization problem. For the simple 1D case, there is a simple explicit solution which can be found based upon the well-known bang-bang principle. For simultaneous movements of actuators (like an XY-stage), the problem quickly becomes much more difficult.

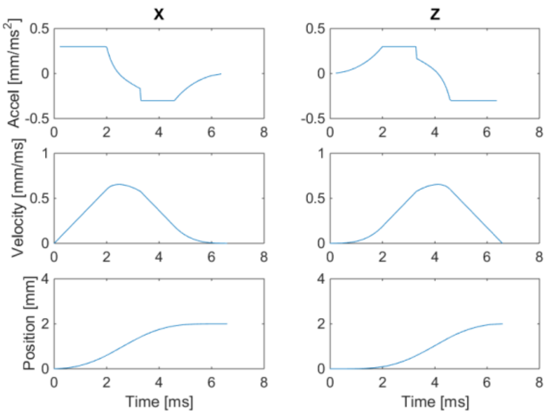


Figure 1: Example of a time-optimal circular trajectory

In this internship, we tackle these types of problems by developing NLP solvers tuned to this problem structure from scratch. In prior work, the problem has been solved using generic NLP solvers like IPOPT in frameworks like CasADi and JuMP, but these have proven to be too slow for the application, as IPOPT is designed for large sparse problems rather than small dense problems with real-time constraints. Development will be a combination of Python and C++ (if speedup is necessary).

Time-optimal trajectories through reinforcement learning

Keywords: Reinforcement learning, TensorFlow

The Eindhoven region hosts many manufacturers of high-tech machines. To have a high throughput, components of such machines need to make very precise movements as fast as possible. The limiting factors are often the dynamic constraints on the actuators, such as maximum velocity, acceleration and higher order derivatives. Completing a movement subject to such constraints as fast as possible is a challenging mathematical optimization problem. For the simple 1D case, there is a simple explicit solution which can be found based upon the well-known bang-bang principle. For simultaneous movements of actuators (like an XY-stage), the problem quickly becomes much more difficult.

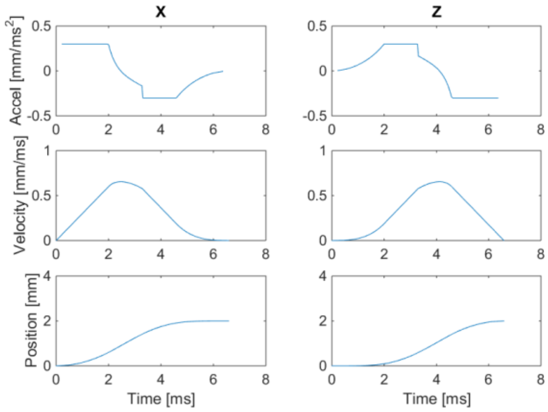


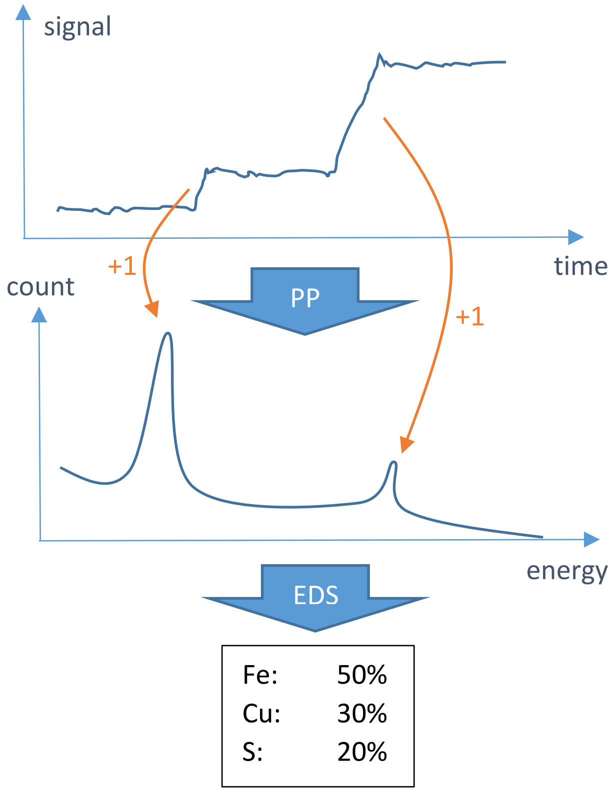
Figure 2: Example of a time-optimal circular trajectory

Solving these problems using non-linear optimization techniques easily results in computation times that are too high for real-time application.

In this internship you will develop a Reinforcement Learning approach to this kind of time-optimal control problems. The primary goal is calculating fast trajectories in a limited amount of computation time (excluding the learning phase). Reinforcement Learning is an area of machine learning that has been used as the core of the AlphaGo and AlphaGo Zero programs. One or multiple agents take actions in an environment with the goal of maximizing the total cumulative reward. These agents in practice are often neural networks that are being trained by learning. The hypothesis behind this internship is that if we can learn an agent to design time-optimal trajectories for a set of training benchmark cases, that it can also solve the validation benchmarks well, and moreover very fast. To develop the methodology, the TensorFlow library and Python programming language will be used.

Spectrum-free spectroscopy

Keywords: Statistical processes, physics, inverse problems



Energy-dispersive X-ray spectroscopy (EDS) is a class of techniques that allow extraction of elemental information from the ‘byproduct’ X-rays that are generated by electron microscopes. As a sample is being bombarded with electrons, X-rays are produced which are registered by a detector which builds up a spectrum (counts as a function of energy). Such spectra contain characteristic peaks which indicate presence of iron, for example. In the building of this spectrum, a low-level signal processing step, called the ‘pulse processor’ (PP) detects steps in a signal and adds counts to a bin with respects to a step height. Any information which could have been obtained about the accuracy of the step detection is thrown away.

In this internship, we propose to devise a solution in which we avoid the intermediate step of spectrum-building and modify standard EDS methods to work with event-type data and try to have benefits from doing so in terms of spectral resolution and accuracy. Though the application is quite physical, the techniques used will be related to simulation of Poisson processes and maximum-likelihood-estimation-type parameter fitting.