Post-doctoral position jointly proposed by project-teams SISTM and FLOWERS
Inria Bordeaux Sud-Ouest, France

Title: Optimization of intervention strategies in epidemic models using deep reinforcement learning techniques.

Co-supervision: Mélanie Prague (SISTM) and Clément Moulin-Frier (Flowers)
(with participation of Rodolphe Thiébaut (SISTM) and Pierre-Yves Oudeyer (Flowers)).

Duration: 1 year

Context:

Infectious diseases, and especially the last COVID-19 pandemics, have important impact on our societies in term of public health, social and economic issues. To mitigate this impact, scientific understanding of the dynamics of spreading of such diseases, associated to methods enabling to optimize and quantify the impact of intervention strategies and their uncertainties, are key to inform policy making. For example, in the COVID-19 context, major decisions to confine populations at large scale were made based on analysis and predictions of mathematical models (Ferguson et al., 2005, 2006, 2020; Cauchemez et al., 2019). In this process, the method used has been to consider a few relatively coarse pre-defined intervention strategies (such as isolation or not) and run predictions of their impact on the epidemic dynamics on mathematical models of epidemic spread (Ferguson et al., 2020). However, given the complexity of the epidemic dynamics (and the associated complexity of models), these pre-defined coarse strategies are bound to be sub-optimal, especially when considering that the problem is multi-objective (e.g. ranging from public health objectives related to number of deaths and ICU saturation to societal and economic objectives) and that strategies may be heterogeneous and multiscale (Halloran et al., 2008).

Project:

The hypothesis we make in this project is that more sophisticated and adaptive strategies could be more efficient, and finding them involves using advanced optimization methods over different kinds of epidemic models. More precisely, we aim to study and adapt the use of state-of-the-art deep reinforcement learning methods which have been proven in other domains to enable find efficient and robust action policies in high-dimensional non-stationary environments with uncertainty and partial observation of the state of the system (Mnih et al., 2015; Haarnoja et al., 2018).

In a first phase of the project, one will compare and select one representative model of epidemic dynamics of the COVID-19 in each of the two following families:
  1. Macroscopic mechanistic models based on ordinary differential equations (or their stochastic counterparts), which are useful when one has only access to aggregated surveillance data at the level of large geographical areas. For this, we will build on and reuse ongoing work on developing such a model in the SISTM Inria team, related to a recent model presented in (Wang et al., 2020), and using maximum likelihood techniques based on Stochastic Approximation Expectation Maximization algorithm to estimate the parameters of the model.

2. Individuo-centered mechanistic models (also called multi-agent models) that take into account more fine-grained information such as the structure of social and spatial networks in the population (Salje et al., 2016), and used recently to inform emergency decision of isolation for the COVID-19 pandemics (Ferguson et al., 2020).

In a second phase, one will focus on studying the optimization of centralized global intervention strategies by using deep reinforcement learning systems (i.e. strategies applied uniformly at the scale of a large geographical area, with few parameters). In particular variants of these algorithms that enable to deal with complex time-dependent dynamics and partial observability (thus we will consider in particular approaches using recurrent neural networks or attention-based architectures). Since models of the epidemic dynamics are available, we will consider model-based deep reinforcement learning approaches (e.g; Chua et al., 2018; Wang et al., 2019), and compare them with model-free approaches (e.g. Mnih et al., 2015; Haarnoja et al., 2018) as well as with more traditional optimization techniques ranging from black-box stochastic optimization to model-based predictive control. The robustness and interpretability of the solutions will be of particular importance in the evaluation process.

In a third phase, one will study the optimization of heterogeneous multi-scale decentralized intervention strategies (e.g. different mitigation actions in different places for different categories of people) using deep multi-agent reinforcement learning algorithms (with centralized learning and decentralized action, Lowe et al., 2017, Rashid et al. 2020) . These will be tested by focusing on multi-agent models or variants of ODE models that are spatialized and incorporate finer-grained compartments.

The scientific outcome of this project will aim to be published in wide audience interdisciplinary journals (e.g. PNAS, Nature Communication/Methods) as well as in specialized venues in epidemiology (introducing the community to deep reinforcement learning tools) and machine learning (raising the interest of this community for this societally important application area).

The project will be co-supervised by Mélanie Prague (SISTM) and Clément Moulin-Frier (FLOWERS), benefitting from the expertise of SISTM in methods for modeling phenomena associated to infectious diseases and their evaluation (e.g. Prague et al. 2012, Vilain et al. 2019, Pasin et al. 2019), and from the expertise of FLOWERS in deep reinforcement learning and multi-agent deep reinforcement learning (e.g. Colas et al., 2018, 2019; 2020).

Skills:

Candidates should have outstanding competences in at least one of these areas, and ideally have experience in several of them:
- Deep reinforcement learning
- Multi-agent reinforcement learning
- Models of epidemic dynamics
- Optimization methods

References

Cauchemez, S., Hoze, N., Cousien, A., & Nikolay, B. (2019). How modelling can enhance the analysis of imperfect epidemic data. *Trends in parasitology*.

Chua, K., Calandra, R., McAllister, R., & Levine, S. (2018). Deep Reinforcement Learning in a Handful of Trials using Probabilistic Dynamics Models. arXiv preprint arXiv:1805.12114.

Colas, C., Sigaud, O., & Oudeyer, P. Y. (2018). Gep-pg: Decoupling exploration and exploitation in deep reinforcement learning algorithms. *ICML 2018.*

Colas, C., Fournier, P., Sigaud, O., Chetouani, M., & Oudeyer, P. Y. (2018). CURIOUS: intrinsically motivated modular multi-goal reinforcement learning. *ICML 2019.*

Colas, C., Karch, T., Lair, N., Dussoux, J. M., Moulin-Frier, C., Dominey, P. F., & Oudeyer, P. Y. (2020). Language as a Cognitive Tool to Imagine Goals in Curiosity-Driven Exploration. *arXiv preprint arXiv:2002.09253.*

Ferguson, N., Laydon, D., Nedjati Gilani, G., Imai, N., Ainslie, K., Baguelin, M., ... & Dighe, A. (2020). Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand.

Ferguson, Neil M., Derek AT Cummings, Christophe Fraser, James C. Cajka, Philip C. Cooley, and Donald S. Burke. "Strategies for mitigating an influenza pandemic." Nature 442, no. 7101 (2006): 448-452.

Ferguson, N. M., Cummings, D. A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., ... & Burke, D. S. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature, 437(7056), 209-214.

Halloran, M. E., Ferguson, N. M., Eubank, S., Longini, I. M., Cummings, D. A., Lewis, B., ... & Wagener, D. (2008). Modeling targeted layered containment of an influenza pandemic in the United States. Proceedings of the National Academy of Sciences, 105(12), 4639-4644.

Haarnoja, T., Zhou, A., Abbeel, P., & Levine, S. (2018). Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. *arXiv preprint arXiv:1801.0129*

Kuhn, E., & Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational statistics & data analysis*, *49*(4), 1020-1038.

Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, O. P., & Mordatch, I. (2017). Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in neural information processing systems* (pp. 6379-6390).

Massonnaud, C., Roux, J., & Crépey, P. (2020). COVID-19: Forecasting short term hospital needs in France. medRxiv.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Petersen, S. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.

Pasin, C., Balelli, I., Van Effelterre, T., Bockstal, V., Solforosi, L., Prague, M., ... & Thiébaut, R. (2019). Dynamics of the humoral immune response to a prime-boost Ebola vaccine: quantification and sources of variation. *Journal of virology*, *93*(18), e00579-19.

Prague, M., Commenges, D., Drylewicz, J., & Thiébaut, R. (2012). Treatment monitoring of HIV-infected patients based on mechanistic models. *Biometrics*, *68*(3), 902-911.

Rashid, T., Samvelyan, M., De Witt, C. S., Farquhar, G., Foerster, J., & Whiteson, S. (2020). Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning. *arXiv preprint arXiv:2003.08839.*

Salje, H., Lessler, J., Paul, K. K., Azman, A. S., Rahman, M. W., Rahman, M., ... & Cauchemez, S. (2016). How social structures, space, and behaviors shape the spread of infectious diseases using chikungunya as a case study. *Proceedings of the National Academy of Sciences*, *113*(47), 13420-13425.

Villain, L., Commenges, D., Pasin, C., Prague, M., & Thiébaut, R. (2019). Adaptive protocols based on predictions from a mechanistic model of the effect of IL7 on CD4 counts. *Statistics in medicine*, *38*(2), 221-235.

Wang, C., Liu, L., Hao, X., Guo, H., Wang, Q., Huang, J., ... & Wei, S. (2020). Evolving Epidemiology and Impact of Non-pharmaceutical Interventions on the Outbreak of Coronavirus Disease 2019 in Wuhan, China. *medRxiv*.

Wang, T., Bao, X., Clavera, I., Hoang, J., Wen, Y., Langlois, E., ... & Ba, J. (2019). Benchmarking model-based reinforcement learning. arXiv preprint arXiv:1907.02057.