Optimization of a Sequential Decision Making Problem for a Rare Disease Diagnostic Application.

Rémi Besson
with Erwan Le Pennec and Stéphanie Allassonnière

École Polytechnique - 09/07/2018

Prenatal Ultrasound Diagnostic

Prenatal Diagnosis

- France: three compulsory ultrasound tests during pregnancy.
- Some classical measures (e.g. trisomy 21).
- But no strict examination protocol.

Outline

Data at Hands

Dataset

- An expert database build from litterature (Emmanuel Spaggiari, M.D. in gynecology obstetrics).
- 81 diseases, 307 symptoms (signs visible with ultrasound):
  - Disease prevalences: $P[D = d_j]$
  - Symptom prevalences given each disease: $P[S_i = k \mid D = d_j]$.

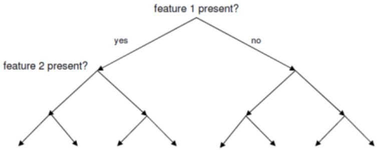
- Database will be enriched from the future exams.

id disease	id symptom	probability of symptom knowing the disease
16	29	0.39
16	136	0.67
16	149	0.50
16	176	0.16
16	181	0.50
16	231	0.75

Our Goals

Medical Goals

- Help obstetricians by improving ultrasonic diagnostic.
- Guide a (non rare disease expert) sonographer to assess as fast as possible potential diseases.



Technical Goals

- Learn a "good" policy $\rightarrow \pi : \mathcal{S} \in \mathbb{S} \mapsto a \in \mathbb{A}$
- State: $\mathbb{S} = \{0, 1, 2\}^{307}$ (presence,absence,not yet looked at) for each symptom.
- Action: $\mathbb{A} = \{1, \ldots, 307\}$ next symptom.
- Develop the final (rather intermediate) product: Shiny app.

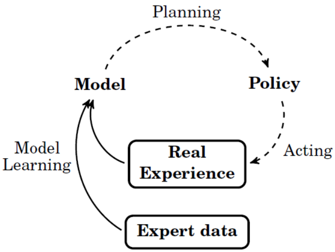
Problems to be Solved

Environment Learning

- We have $P[S_i \mid D]$ but we need to know $P[S_{i_1}, ..., S_{i_K} \mid D]$.
- Idea: add some expert knowledge and maximize uncertainty.

Diagnostic Strategy Optimization

- Find a policy that allows to detect the disease while minimizing the average duration.
- Idea: recast the problem as a non adversial game and find the optimal strategy.



Outline

Diagnostic Strategy Optimization

Diagnostic Strategy Optimization.

- Find a policy that allows to detect the disease while minimizing the average duration.

Measure of Performance

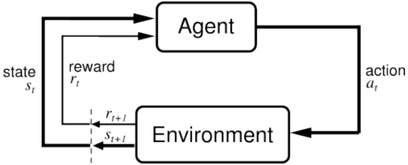
- Number of questions before being able to diagnose a disease.

Alternative Formulations

- Trade-off: cost of misdiagnosis/cost of medical tests to perform.
- Reach the lowest uncertainty under fixed budget constraint (time, money).

Reinforcement Learning

- Sutton (98): An agent takes actions in a sequential way, receives rewards from the environment and tries to maximize his long-term (cumulative) reward.
- Has been applied a lot in robotics or game (Go...).



Reward Signal Design

- We give a -1 reward for each question posed before the diagnostic can be made and 0 afterwards.
- When should we stop and consider that we can make a diagnostic? $\rightarrow$ when the disease entropy falls below ϵ.

Value Functions and Optimization

Reward and Optimization Task

- T the stopping time, r_t the reward obtained at time t:

$$\mathcal{R} = \sum_{t=1}^{T} r_t$$

- $s_0 = (2, \ldots, 2)$:

$$\pi^\star = \operatorname*{argmax}_{\pi} \mathbb{E}_\pi[\mathcal{R} \mid s_0]$$

Value Functions

- Given a strategy π: two values functions measuring
  - the performance of the strategy starting from s

$$V_\pi(s) = \mathbb{E}_\pi[\mathcal{R}|s]$$

  - the performance if one forces the first action to be a

$$Q_\pi(s, a) = \mathbb{E}_\pi[\mathcal{R}|s, a]$$

Dynamic Programming

Two observations

- We can evaluate a policy using Bellman expectation backup:

$$v_\pi(s) = \sum_{s'} p(s' \mid s, \pi(s))(v_\pi(s') + r(s, a, s'))$$

- We enhance π if we replace it by π' (policy improvement):

$$\begin{aligned} \pi'(s) &= \operatorname*{argmax}_a Q_\pi(s, a) \\ &= \operatorname*{argmax}_a \sum_{s'} p(s' \mid s, a)(v_\pi(s') + r(s, a, s')) \end{aligned}$$

Dynamic Programming

- $\pi_0 \xrightarrow{\text{E}} v_{\pi_0} \xrightarrow{\text{I}} \pi_1 \xrightarrow{\text{E}} v_{\pi_1} \xrightarrow{\text{I}} \dots \xrightarrow{\text{I}} \pi^\star \xrightarrow{\text{E}} v_{\pi^\star}$.
- Issue: impossible to use in high dimension!
  $\rightarrow$ State dimension: 3^{307}

Policy Parametrisation

Parametrization

- $\pi_\theta(s) = f_\theta(\Phi(s))$ with $\theta \in \mathbb{R}^d$.
- Example: f_θ depends only on $P(A|s)$ and $H(D|s)$ through a logit model $\rightarrow$ inspired from Breiman algorithm for decision tree optimization.
- Values functions:

$$V_\theta(s) = \mathbb{E}_{\pi_\theta}[T|s]$$
$$Q_\theta(s,a) = \mathbb{E}_{\pi_\theta}[T|s,a]$$

Parametric Optimization

- Optimization in θ:
  - *gradient* descent, order 0 optimization...
- Issue: neither V or Q are known...
- Monte Carlo technique to estimate those quantities from game/simulation.

Action-Value Function Parametrisation

Reminder

- We "only" need to learn $Q_{\pi^\star}$ since:

$$\pi^\star(s) = \operatorname*{argmax}_a Q_{\pi^\star}(s, a)$$

Action-Value Function Look-up table

- Action-Value function satisfies the following Bellman equation:

$$Q^\star(s, a) = E_{s'}[\max_{a'} Q^\star(s', a') + r(s, a, s')]$$

- Value iteration algorithm solve the Bellman equation:

$$Q_{i+1}(s, a) \leftarrow E_{s'}[\max_{a'} Q_i(s', a') + r(s, a, s') \mid s, a]$$

- $Q_i \rightarrow Q^\star$ as $i \rightarrow \infty$.

Action-Value Function Parametrisation

Deep Q-network algorithm

- Represent action-value function by a deep Q-network with weights w:

$$Q(s, a) \approx Q_w(s, a)$$

- Loss function:

$$L_i(w_i) = E_{s,a}[(y_i - Q_{w_i}(s, a))^2]$$

  where $y_i = E_{s'}[\min_{a'} Q_{w_i}(s', a') + r(s, a, s') \mid s, a]$ is the target.
- Some tricks to overcome instability: experience replay, freeze target Q-network:

$$L_i(w_i) = E_{s,a,s' \sim D}[(r(s, a, s') + \min_{a'} Q_{w_{i-1}}(s', a') - Q_{w_i}(s, a))^2]$$

.

High-dimensional issues.

Issue

- DQN algorithm is not tractable for the main task: to find the best path starting from $s_0 = (2, ..., 2)$.

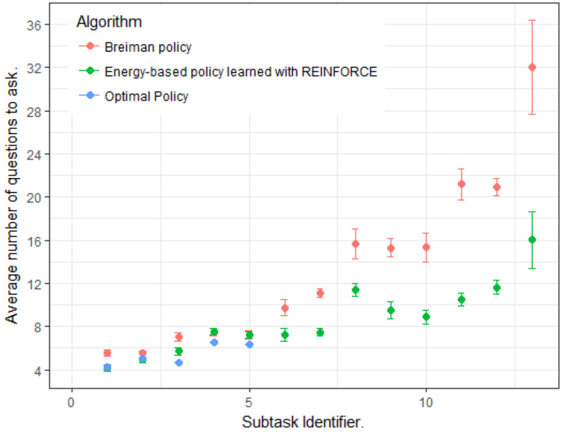
Dimension reduction

- Idea: Create subproblems of lower dimension.
- Learn a strategy starting from each $s_0^{(i)} = (2, ..., 2, 1, 2, ..., 2)$.
- Assumption: this first observed symptom is relevant (we can focus on the diseases for which this initial symptom is typical $\rightarrow$ reduce dimension).

Transfer Learning

- Learn the strategy for the global task from what have been learned on subtasks: transfer learning.
- Ongoing research: promising results.

Some Results



Some Results

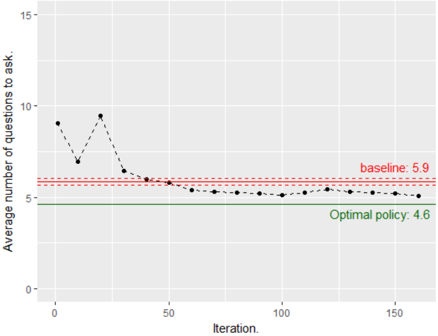


Figure: Evolution of the performance of the neural network during the training phase with DQN-MC. Task dimension: 10.

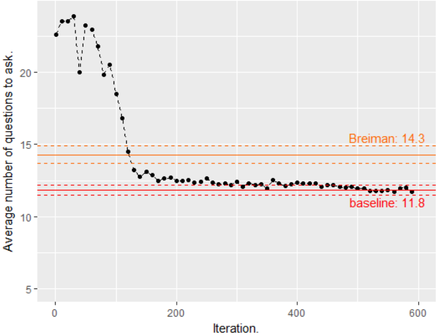


Figure: Evolution of the performance of the neural network during the training phase with DQN-MC. Task dimension: 26.

Outline

Uncertainty and Entropy

Environment Learning

- We have $P[S_i \mid D]$ but we need to know $P[S_{i_1}, ..., S_{i_K} \mid D]$.
- Idea: add some expert knowledge and maximize uncertainty.

Expert knowledge

- Some symptoms can not occur simultaneously...
- You need at least a certain number of symptoms to talk about a syndrome.

Uncertainty

- General idea: choose a solution that maximize the uncertainty while respecting the constraints (probability/impossibility).
- Uncertainty measured by entropy.

MaxEnt Principle

Environment Learning

- We have $P[S_i \mid D]$ but we need to know $P[S_{i_1}, ..., S_{i_K} \mid D]$.

- Naive idea: $P[S_{i_1}, ..., S_{i_K} \mid D] = P[S_{i_1} \mid D] \times ... \times P[S_{i_K} \mid D]$ (Conditional independence)

Data and Expert Knowledge

- Conditional probabilities: $P[S_i \mid D]$
- Medical constraints: $P[S_{i_k}, S_{i_{k'}} \mid D] = 0$...
- Mathematical constraints: P should be a probability...

MaxEnt Principle

- Maximize the entropy of the distribution $P[S_{i_1}, ..., S_{i_k} \mid D]$ under mathematical and medical constraints.

- Numerical scheme available.

Take Away Message

Medical Goals

- Help obstetricians by improving/systematizing ultrasonic diagnostic (MDP modeling)
- Guide a (non rare disease expert) sonographer to assess as fast as possible potential diseases (first prototype at Necker)

Technical Goals

- Build an optimized decision tree:
  - Need to learn the environment (MaxEnt and data assim.)
  - Reinforcement learning (parametrized policy and MC)
- Not yet (theoretical) guarantees.

Take Away Message

- Formalization requires a true dialog between the mathematicans and the practicians.
- Defining the performance measure (the metric goal) not always easy.
- First prototype already interesting.
- Promising direction.

Acknowledgment

Thesis Team (CMAP/Paris Descartes)

- Rémi Besson, Erwan Le Pennec, Stéphanie Allassonnière

Yves Ville Team (Hôpital Necker Enfants Malades)

- Julien Stirneman, Emmanuel Spaggiari

Anita Burgun Team (Paris Descartes/INSERM/Institut Imagine)

- Nicolas Garcelon, Anne-Sophie Jannot, Antoine Neuraz, Anita Burgun

References

- Bayer-Zubeck V., Dietterich T.G (2005), Integrating Learning from Examples into the Search for Diagnostic Policies Journal of Artificial Intelligence Research, Oregon.
- Breiman L., Friedman J., Stone C.J, Olshen R.A. (1984), Classification and regression trees Wadsworth Statistics/Probability, Berkeley.
- Jaynes, E.T. (1963), Information Theory and Statistical Mechanics, The Physical Review, Standford.
- V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M.A. Riedmiller (2013). Playing atari with deep reinforcement learning. CoRR.
- Sutton R.S., Barto A.G. (1998), Introduction to reinforcement learning Bradford Book, MIT Cambridge.
- K.F Tang, H.C. Kao, C.N. Chou, E.Y. Chang (2016), Inquire and Diagnose: Neural Symptom Checking Ensemble using Deep Reinforcement Learning *NIPS*, Barcelona.