Bayesian filtering, parallel hypotheses and uncertainty: A new, combined model for human learning

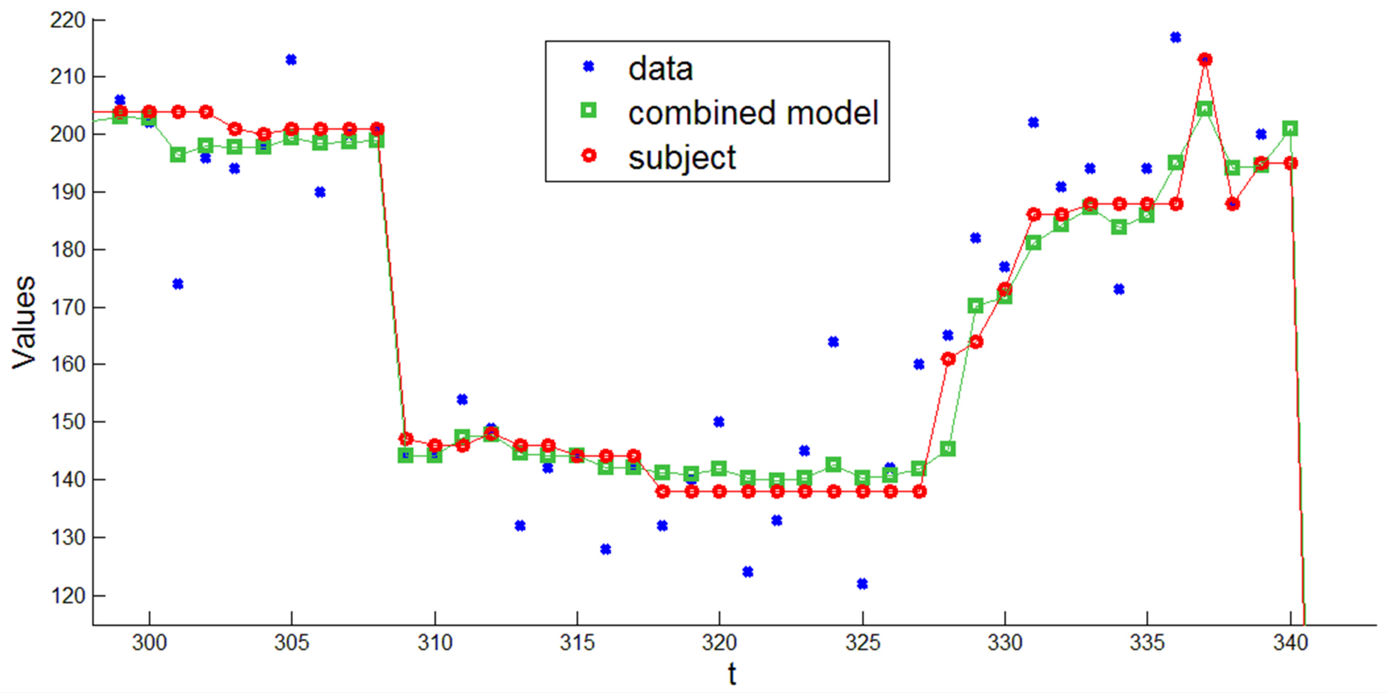
Marco Lehmann[1], Alexander Aivazidis[1], Mohammad Javad Faraji[1], Kerstin Preuschoff[1,2]
[1]Ecole Polytechnique Fédérale de Lausanne, [2]Université de Genève

Abstract

Human learning in unstable environments can be modelled by change-point detection algorithms. We compare different existing algorithms and propose a new one. Our model is simple, solves the change-point problem with remarkable performance, fits well to experimental data, and has an intuitive interpretation. It combines three cognitive processes: Bayesian filtering, parallel memory traces at long and short timescales, and attentional selection by taking into account unexpected uncertainty.

The task

Numbers are drawn from a normal distribution. With hazard rate H, the mean can abruptly change. The goal is to sequentially estimate the underlying mean from noisy observations. (Wilson et al, 2013)



Overview of Algorithms

- The algorithms differ by the information they keep about the past.
- They integrate available (stored) information optimally.

Model	TD estimate	runlength estimates	Estimation Uncertainty
TD	1 fixed	-	-
Wilson	N fixed	-	-
Nassar		1 fixed, 1 adaptive	-
Payzan-LeNestour		1 fixed, 1 adaptive	2
Combined		2 adaptive	2
Adams/MacKay		maximum	maximum

Algorithms

- **Temporal difference learning with constant learning rate (TD α)**
  - Keeps only a single running average.

  $$\theta_t = \theta_{t-1} + \alpha(x_t - \theta_{t-1})$$

- **Full Bayesian model (Adams/MacKay 2007)**
  - Considers *all* past data.
  - Keeps one distribution over θ per run length

  $$P(\theta_t|\vec{X_t}) = \sum_{r=1}^{t} P(\theta|r)P(r|\vec{X_t})$$

- **Wilson et al. 2013**
  - Reduces the full Bayesian model by i) considering only a (fixed) subset of N possible runlengths and ii) computing a running weighted average μ^r per runlength (one constant learning rate α^r per runlength)

  $$\theta_t = \sum_{i=1}^{N} P(r_i|\vec{X_t})\mu^{r_i}$$

- **Nassar et al. 2010**
  - Considers two possible run lengths: r_t=0 (change point) and a longer runlength that either increases or decreases depending on the probability Ω of a change point (cp).

  $$\Omega_t = \frac{P(x_t|cp)^{\lambda}H}{P(x_t|cp)^{\lambda}H + P(x_t|\overline{cp})^{\lambda}(1-H)}$$

  $$\theta_t = \Omega_t x_t + (1-\Omega_t)\frac{x_t + \hat{r}_{t-1} \times \theta_{t-1}}{\hat{r}_t + 1}$$

- **Payzan-LeNestour, Bossaerts 2011**
  - Considers two possible *distributions* over θ, one «prior» and one for a runlength r_t.
  - computes a subjective likelihood λ that *no jump* has occurred.

  $$P(\theta) = (1-\lambda)P(\theta|r=0) + \lambda P(\theta|r_{long})$$

Combined Model (new)

i. for some prior $P(x_t|0)$ and hazard rate H, recursively compute the likelihood of three hypotheses:

$$P(r_t = r_{short}+1|\vec{X_t}) \propto (1-H)P(x_t|r_{t-1} = r_{short})P(r_{t-1} = r_{short})$$
$$P(r_t = r_{long}+1|\vec{X_t}) \propto (1-H)P(x_t|r_{t-1} = r_{long})P(r_{t-1} = r_{long})$$
$$P(r_t = 0|\vec{X_t}) \propto HP(x_t|0)$$

ii. Keep the two most likely runlenths only (pruning), and compute a generalized change point probability: Ω_t = P(changepoint r_{short} timesteps ago):

$$\Omega_t = \frac{P(r_{short})}{P(r_{short}) + P(r_{long})}$$

iii. The distribution over θ_t is a weighted sum of two Bayesian posterior distributions:

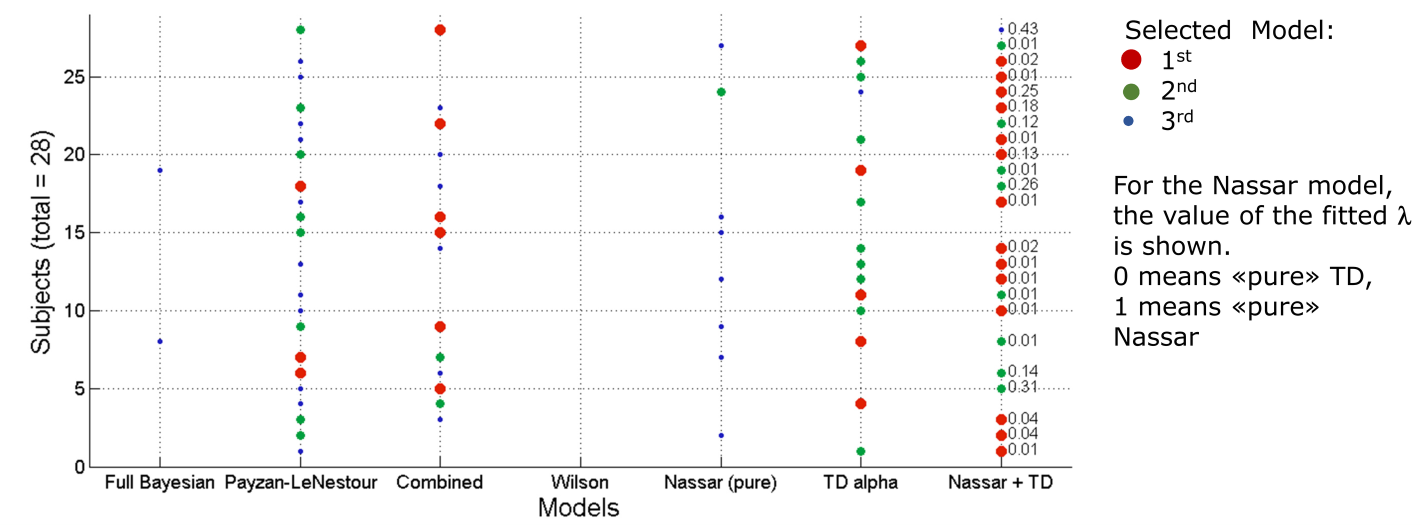
$$P(\theta_t) = \Omega_t P(\theta|r_{short}) + (1-\Omega_t)P(\theta|r_{long})$$

Model Selection

- Model parameters were fit to maximize the log likelihood:

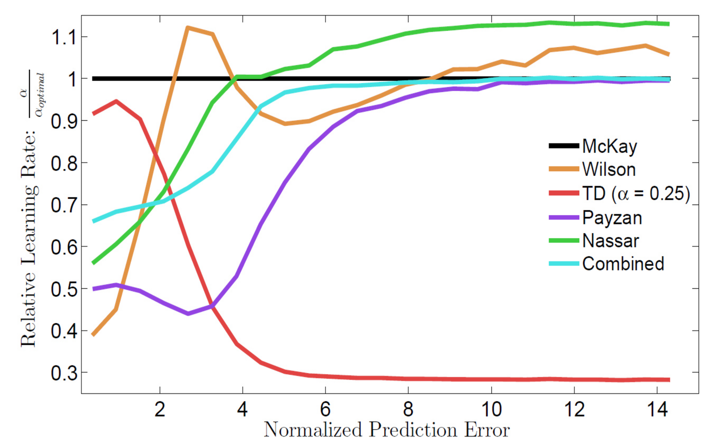
$$logp(\theta^S_{1..n}|Model) = \sum_{n=1}^{N}\frac{(\theta^S_n - \theta^M_n)^2}{2\sigma^2_{decision}} - Nlog(\sigma_{decision}) - \frac{N}{2}log(2\pi)$$

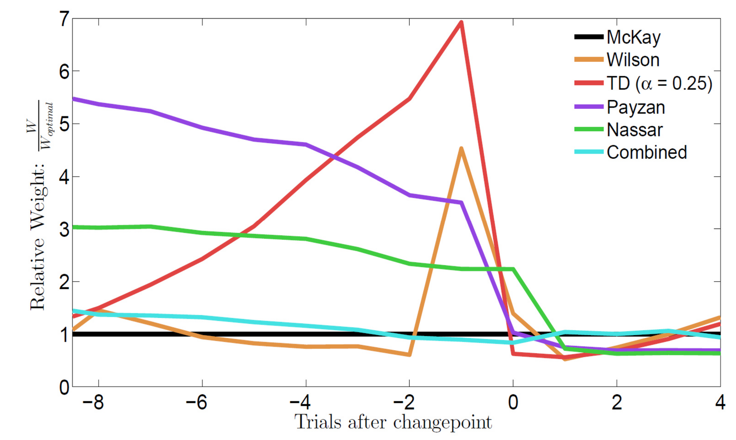
- Bayesian information criterion (BIC) used to «rank» models:



Characteristic Features

- The better an algorithm detects a change-point, the better it can modulate the learning rate. Data before a change point should not influence the current estimate.
- Models that take into account estimation uncertainty detect change points more reliably (e.g. Full, Payzan, Combined)
- Without taking estimation uncertainty into account, an outlier «looks the same» as a change point. (e.g. Nassar Model)





Conclusion

- The combined model closely matches optimal performance and fits well to human behaviour.
- The combined model allows a competing hypothesis to accumulate evidence. This seems to be a strategy also implemented by humans.
- 9 of 28 subjects take estimation uncertainty into account (best fit by either Payzan-LeNestour or Combined model).
- Humans do modulate the learning rate. Yet the models fail to explain those variations for 14 of 28 subjects. TD learning with constant learning rate best explains the data. What are the models missing?

[1] R. P. Adams, D.J.C. MacKay. «*Bayesian Online Changepoint Detection*», arXiv:0710.3742 [stat.ML] 2007
[2] R. C. Wilson et al. «*A mixture of delta-rules approximation to Bayesian inference in change-point problems.* » PLoS Comput Biol. 2013
[3] M. R. Nassar et al. «*An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment*», J Neurosci. 2010
[4] E. Payzan-LeNestour, P. Bossaerts «*Risk, Unexpected Uncertainty, and Estimation Uncertainty: Bayesian Learning in Unstable Settings*», PLoS Comput Biol. 2011
Datasets: Aivazidis 2014 | Wilson, Nassar 2013
This research was supported by Swiss National Science Foundation Sinergia Grant No. CRSII2_147636 (Learning from delayed and sparse feedback)