Natural Actor Critic

Do you have a subtitle?
If so, write it here

First Author · Second Author

F. Author
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

S. Author
second address

1 Paper

Natural Actor Critic:

- Main Version [6].
- 2nd Version: Natural Actor-Critic in Neurocomputing [4].
- 3rd version: RL of motor skills with policy gradients in NN [5].

Recommended by Jan:

- Policy Evaluation with TD [3].
- Incremental NAC algorithms [1].
- Jan said that a paper form C. Dann is very important. Did he mean Policy
  Evaluation with TD by Dann or did he mean a second paper?

Research:

- Comparison of four natural gradient algorithms (co-author Sutton) [2].

2 Meetings & Notes

Meetings:

- 12.12.18: Notes from Jan can be found in ".\Notes Jan 12.12.18"

Table 1 Please write your table caption here

first	second	third
number	number	number
number	number	number

Abstract Insert your abstract here. Include keywords, PACS and mathematical subject classification numbers as needed.

Keywords First keyword · Second keyword · More

3 Introduction

- Steepest ascent direction of performance object with respect to any metric $M(\theta)$: $M(\theta)^{-1}\nabla_\theta J(\mu_\theta)$
- The natural gradient ist the steepest ascent direction with respect to the Fisher information metric $M_\pi(\theta) = E_{s\sim\rho^\pi, a\sim\pi_\theta}[\nabla_\theta log\pi_\theta(a|s)^T \nabla_\theta log\pi_\theta(a|s)]$
- For deterministic policies: $M_\mu(\theta) = E_{s\sim\rho^\mu}[\nabla_\theta\mu_\theta(s)\nabla_\theta\mu_\theta(s)^T w]$
  - Limiting case of the Fisher information metric: policy variance reduced to zero
- Combining DPG theorem with compatible function approximation gives $\nabla_\theta J(\mu_\theta) = E_{s\sim\rho^\mu}[\nabla_\theta\mu_\theta(s)\nabla_\theta\mu_\theta(s)^T w]$ so steepest ascent direction reduces to $M_\mu(\theta)^{-1}\nabla_\theta J_\beta(\mu_\theta) = w$

Your text comes here. Separate text sections with

4 Section title

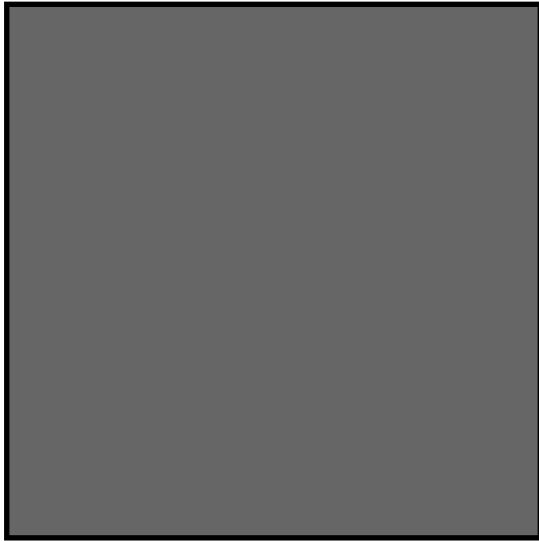Text with citations [**?**] and [**?**].

4.1 Subsection title

as required. Don't forget to give each section and subsection a unique label (see Sect. 4).

Paragraph headings Use paragraph headings as needed.

$$a^2 + b^2 = c^2 \tag{1}$$

Fig. 1 Please write your figure caption here

Fig. 2 Please write your figure caption here

References

1. Bhatnagar S, Ghavamzadeh M, Lee M, Sutton RS (2008) Incremental natural actor-critic algorithms. In: Advances in neural information processing systems, pp 105–112
2. Bhatnagar S, Sutton RS, Ghavamzadeh M, Lee M (2009) Natural actor–critic algorithms. Automatica 45(11):2471–2482
3. Dann C, Neumann G, Peters J (2014) Policy evaluation with temporal differences: A survey and comparison. The Journal of Machine Learning Research 15(1):809–883
4. Peters J, Schaal S (2008) Natural actor-critic. Neurocomputing 71(7-9):1180–1190
5. Peters J, Schaal S (2008) Reinforcement learning of motor skills with policy gradients. Neural networks 21(4):682–697
6. Peters J, Vijayakumar S, Schaal S (2005) Natural actor-critic. In: European Conference on Machine Learning, Springer, pp 280–291