

BioE 2696/ECE 2695: Control Theory in Neuroscience
(3 Credits, Spring 2009)

Adaptive Control and Reinforcement Learning

April 1, 2009

Instructor: Zhi-Hong Mao
Assistant Professor of ECE and Bioengineering
University of Pittsburgh, Pittsburgh, PA

1

Outline

- Why adaptive control?
- What is adaptive control?
- Approaches to adaptive control
 - Gain scheduling
 - Model reference adaptive control
 - Self tuning regulation
 - **Reinforcement learning**

2

Why adaptive control?

- Systems to be controlled have parameter uncertainty
- System dynamics experience unpredictable parameter variations as the control operation goes on

3

Why adaptive control?

- Systems to be controlled have parameter uncertainty
- System dynamics experience unpredictable parameter variations as the control operation goes on

- **Examples**

- Aircraft control



4

Why adaptive control?

- Systems to be controlled have parameter uncertainty
- System dynamics experience unpredictable parameter variations as the control operation goes on

- **Examples**

- Aircraft control
- Robot manipulation

[Video from Atkeson Lab](#)



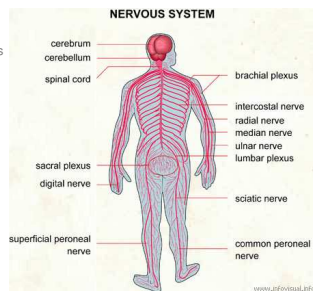
5

Why adaptive control?

- Systems to be controlled have parameter uncertainty
- System dynamics experience unpredictable parameter variations as the control operation goes on

- **Examples**

- Aircraft control
- Robot manipulation
- Neural control of movement



6

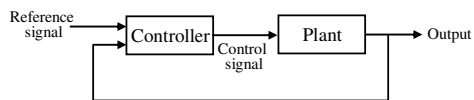
What is adaptive control?

- Definition
 - A form of control in which the control parameters are automatically adjusted as conditions change so as to optimize performance [cited from Oxford English Dictionary]

7

What is adaptive control?

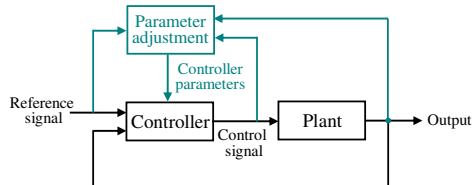
- Definition
- Difference between adaptive control and ordinary feedback control



8

What is adaptive control?

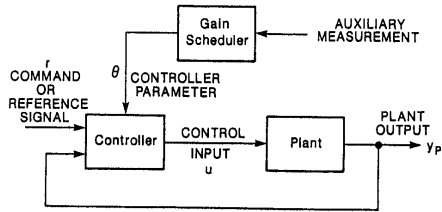
- Definition
- Difference between adaptive control and ordinary feedback control



9

Approaches to adaptive control

- Gain scheduling



10

Approaches to adaptive control

- Gain scheduling
 - Advantages of gain scheduling
 - Parameters can be changed quickly in response to changes in the plant dynamics. It is convenient especially if the plant dynamics depend in a well-known fashion on a relatively few easily measurable variables

11

Approaches to adaptive control

- Gain scheduling
 - Advantages of gain scheduling
 - Disadvantages of gain scheduling
 - It is an open-loop adaptation scheme, with no real "learning" or intelligence
 - The extent of design required for its implementation can be enormous (example: flight control system of CH-47 helicopter)

12

Approaches to adaptive control

- Gain scheduling

CH-47
(Chinook)

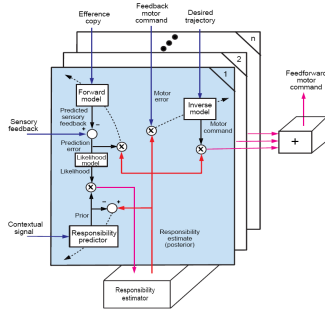


13

Approaches to adaptive control

- Gain scheduling

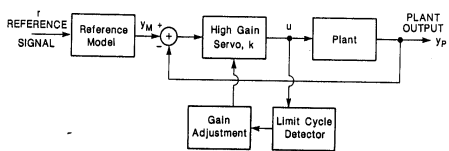
Multiple paired
internal models
in the
cerebellum



14

Approaches to adaptive control

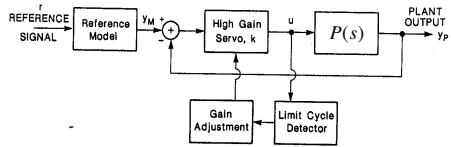
- Gain scheduling
- Model reference adaptive control
 - Series high-gain scheme



15

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
 - Series high-gain scheme

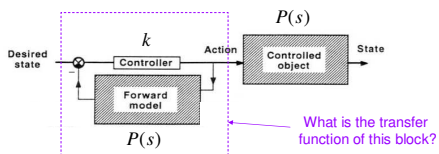


Questions: What is the transfer function between y_m and y_p ? Why use high gain k ?

16

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
 - Series high-gain scheme



Using high-gain and a plant model to implement an inverse model of the plant

17

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
 - Series high-gain scheme
 - Advantages: intuitive and simple—only one parameter is updated

18

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
 - Series high-gain scheme
 - Advantages
 - Disadvantages:
 - (1) Oscillations are constantly present in the system
 - (2) Noise in the frequency band of the limit cycle detector causes the gain to decrease well below the critical value
 - (3) Reference inputs may cause saturation due to the high-gain
 - (4) Saturation may mask limit cycle oscillations, allowing the gain to increase above the critical value, and leading to instability

19

Approaches to adaptive control

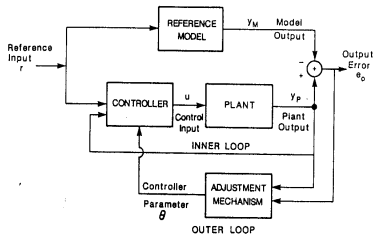
- Gain scheduling
- Model reference adaptive control
 - Series high-gain scheme



20

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
 - Series high-gain scheme
 - Parallel scheme



21

Approaches to adaptive control

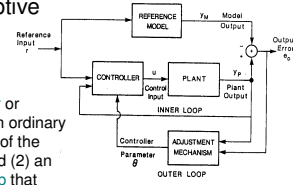
- Gain scheduling

- **Model reference adaptive control**

- Series high-gain scheme

- Parallel scheme

- Two loops: (1) an **inner or regulator loop** that is an ordinary control loop consisting of the plant and regulator, and (2) an **outer or adaptation loop** that adjusts the parameters of the regulator in such a way as to drive the error between the model output and plant output to zero



22

Approaches to adaptive control

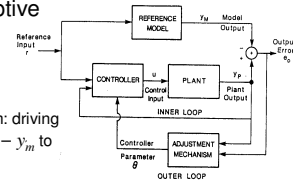
- Gain scheduling

- **Model reference adaptive control**

- Series high-gain scheme

- Parallel scheme

- Two loops
- Adjustment mechanism: driving the output error $e_0 = y_p - y_m$ to zero



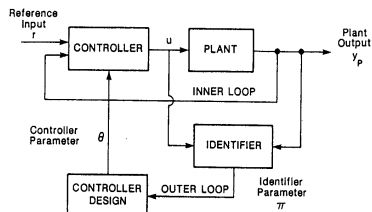
23

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control

- **Self tuning regulation**

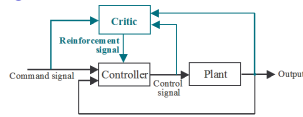
- Certainty equivalence principle



24

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
- Self tuning regulation



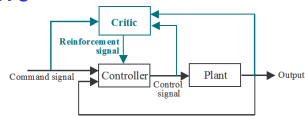
• Reinforcement learning

- What are critic and reinforcement signal?
 - Critic generates evaluative learning feedback on the basis of observing the control signals and their consequences on behavior of the plant
 - Critic produces reinforcement signal (“reward” or “punishment”) that provides a measure of the “goodness” of the behavior or simply an indication of success or failure

28

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
- Self tuning regulation



• Reinforcement learning

- What are critic and reinforcement signal?
- Differences between reinforcement learning and other adaptive control approaches
 - Critic does *not* directly tell the controller what it should have done (critic is not a “teacher”)
 - Learning from trial-and-error
 - Reinforcement learning may create surprises and real intelligence!

29



Massachusetts Institute of Technology

Learning to walk

30

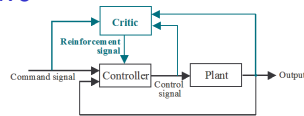


Learning to walk up slope
<http://www.nld.ds.mpg.de/~poramate/Funbot.html>

31

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
- Self tuning regulation



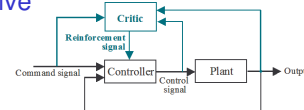
• Reinforcement learning

- What are critic and reinforcement signal?
- Differences between reinforcement learning and other adaptive control approaches
- Credit assignment problems
 - Spacial credit assignment problem
 - How is credit assigned to the internal working of a complex structure?
 - Hypothesis: Dopamine mediates synaptic enhancement in the corticostriatal pathway in the manner of a broadcast reinforcement signal

32

Approaches to adaptive control

- Gain scheduling
- Model reference adaptive control
- Self tuning regulation



• Reinforcement learning

- What are critic and reinforcement signal?
- Differences between reinforcement learning and other adaptive control approaches
- Credit assignment problems
 - Spacial credit assignment problem
 - Temporal credit assignment problem
 - How can reinforcement learning work when the system behavior is temporally extended and evaluations occur at varying and unpredictable times?
 - It is especially relevant in motor control because movements extend over time and evaluative feedback may become available, for example, only after the end of a movement

33

References

- K. J. Astrom, Theory and applications of adaptive control—a survey, *Automatica*, vol. 19, no. 5, pp. 471-486, 1983.
- K. J. Astrom, Adaptive control around 1960, *IEEE Control Systems*, pp. 44-49, June 1996.
- K. J. Astrom and B. Wittenmark, *Adaptive Control*, 2nd Edition, Addison-Wesley, 1995.
- S. H. Collins, A. Ruina, R. Tedrake, and M. Wisse, Efficient bipedal robots based on passive-dynamic walkers. *Science*, vol. 307, pp. 1082-1085, 2005.
- L. Itti, Lecture 11: reinforcement learning, CS564-Brain Theory and Artificial Intelligence, University of South California.
- R. C. Miall, D. J. Weir, D. M. Wolpert, and J. F. Stein, Is the cerebellum a Smith predictor? *Journal of Motor Behavior*, vol. 25, no. 3, pp. 303-216, 1993.
- S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence, and Robustness*, Prentice-Hall, 1989.
- J.-J. E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice Hall, 1991.
- D. M. Wolpert, R. C. Miall, and M. Kawato, Internal models in the cerebellum, *Trends in Cognitive Sciences*, vol. 2, no. 9, pp. 338-347, 1998.
- http://www.geocities.com/Hollywood/2307/ch-47d_chinook.jpg
- http://www.infovisual.info/03/038_en.html
- <http://www.nld.ds.mpg.de/~poramate/Runbot.html>
- <http://www.photosfan.com/images/stealth-aircraft-flying-in-formation1.jpg>
- http://www.robotokyo.com/wp-content/uploads/2007/04/domo_robot.jpg
- <http://www.surclaro.com/photo6937.html>
- <http://www.youtube.com/watch?v=JLmmEr7Y7ic>

34

