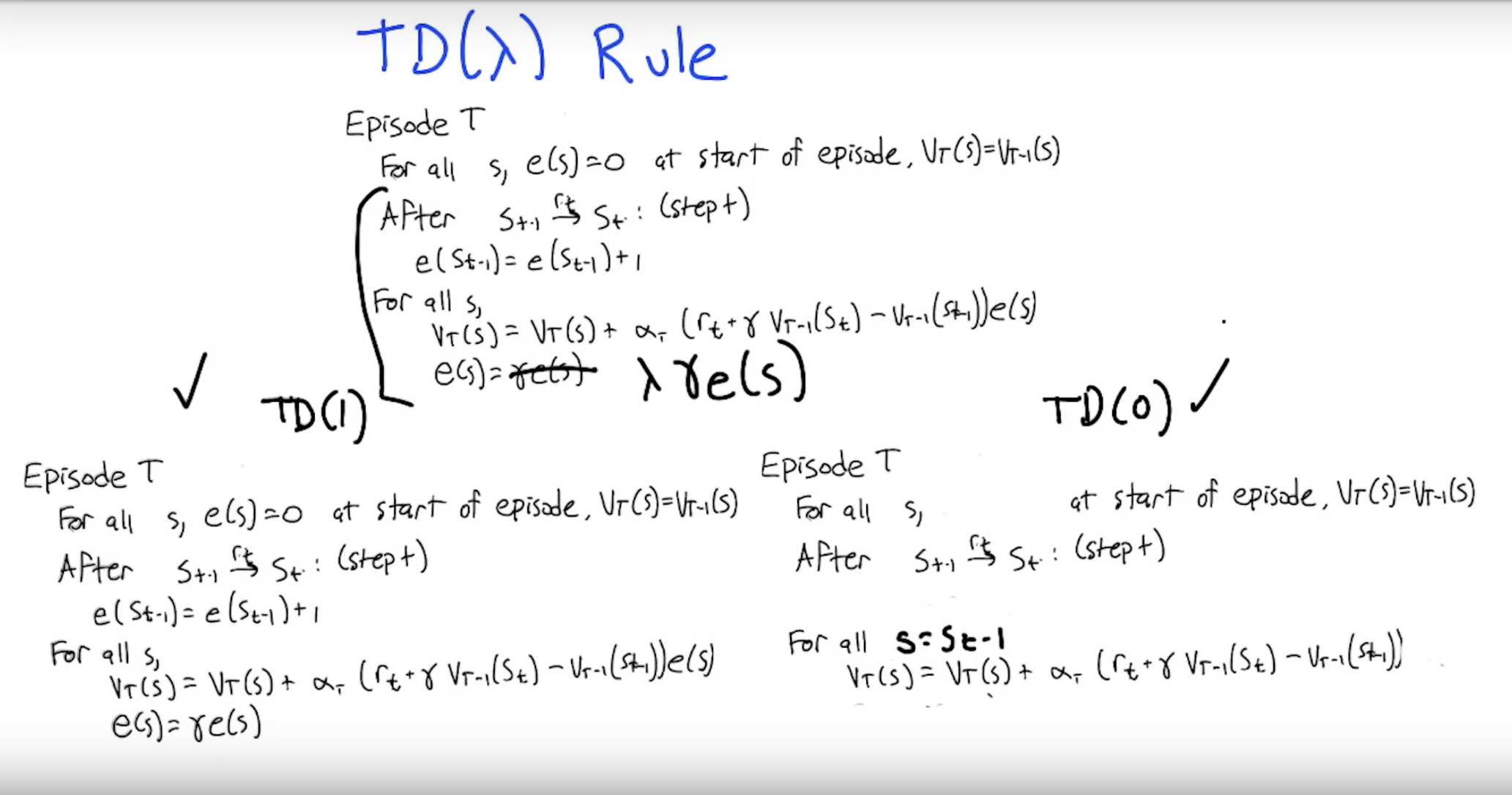
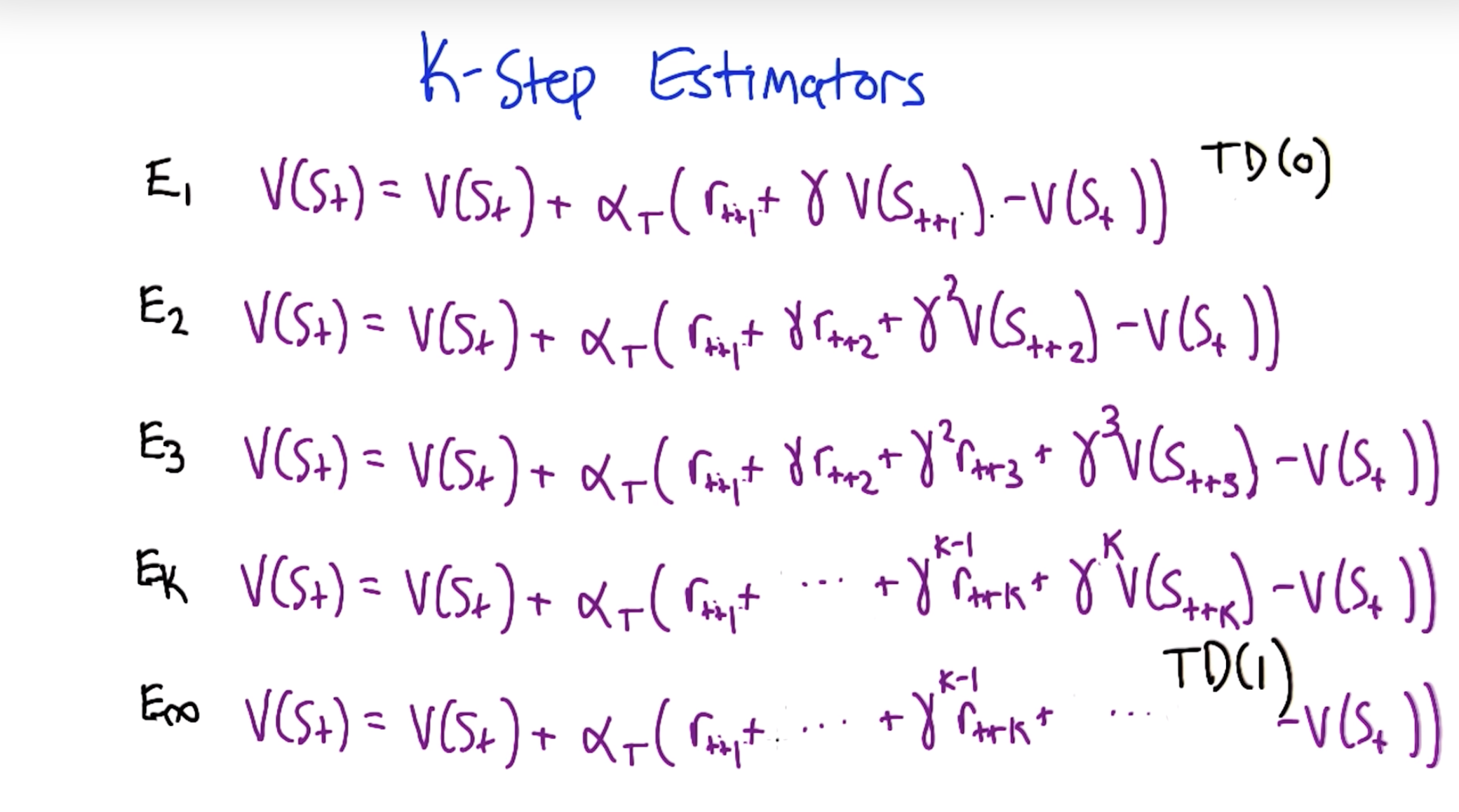
1. Get Souza, explain what do we want to do, how we move forward.
2. Which one is good for us to buy, on T-cells? On CD4+, CD8+

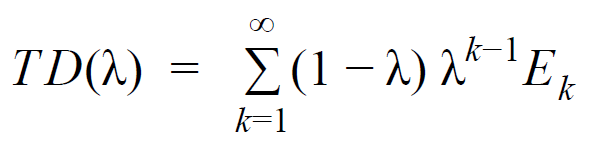
Section 1 - Exceeded expectations. Section 2 - Exceeded expectations. Section 3 - Exceeded expectations.





# Won't TD(1) Always be Zero?

Given this formula from HW 2:



Won't TD(1) always be zero because the first factor is zero and everything else is multiplied by it?

[hw2](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)

[**edit**](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)·[good question](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)0

Updated 6 days ago by

Bruce Nielson

**the students' answer,**

*where students collectively construct a single answer*

Click to start off the wiki answer

**the instructors' answer,**

*where instructors collectively construct a single answer*

If this were a finite sum then that would be true.  However, since it's not, you need to be careful of 0⋅∞, since you're summing an infinite number of k-step returns.  Check out Section 12.1 in S&B for more information and some math.

**~ An instructor (**

**Farrukh Rahman**

**) endorsed this answer  ~**

[thanks!](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)2

Updated 6 days ago by

Vahe Hagopian

**followup discussions**

*for lingering questions and comments*

Resolved Unresolved



**[Kirill Luzin](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)**

[1 day ago](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)

According to Section 12.1, formula (12.3) in S&B: TD(1) = Gt1 = Gt which is conventional return. If my understanding of the lectures is correct then Gt (conventional return) might be calculated as outcome-based estimate V(s0)... I calculated it for the 1st example and got 16.678. When I try to put different values λ⊂[0,1) to the formula (12.3) result TD(λ) = Gtλ increases from 0.6 to 16.687... there is no way it touches 16.687 at 0.6227 point... I assume my Ek=Gt:t+k (formula 12.1) calculations are wrong. Does anyone else get the value of TD(1) = 16.687?

[helpful!](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1) 0

/var/folders/m2/02s5_6cx2gnfg7__3r24l_ph0000gp/T/com.microsoft.Word/WebArchiveCopyPasteTempFiles/1546887719_35.png

**[Bruce Nielson](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)**

[1 day ago](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)

I believe that is correct.

[helpful!](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1) 3



**[Kirill Luzin](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)**

[1 day ago](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)

Thanks, Bruce. If we assume it is correct, how do we calculate the TD(λ) for the 1st example? Equation (12.3) sort of clear. My K-step estimators might be wrong. I calculated them as  
Gt:t+1 = 0.81\*(R0+Vs1)+0.19\*(R1+Vs2)

Gt:t+2 = 0.81\*(R0+R2)+0.19\*(R1+R3) + Vs3  
Gt:t+2 = 0.81\*(R0+R2)+0.19\*(R1+R3) + R4 + Vs4  
...  
Gt:t+5 = 0.81\*(R0+R2)+0.19\*(R1+R3) + R4 + R5 + R6 + Vs6  
  
do you see any flaws here?

[helpful!](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1) 0



**[Vahe Hagopian](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)**

[23 hours ago](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)

Did you also compute E6,E7,E12,E237, etc.?

[helpful!](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1) 2

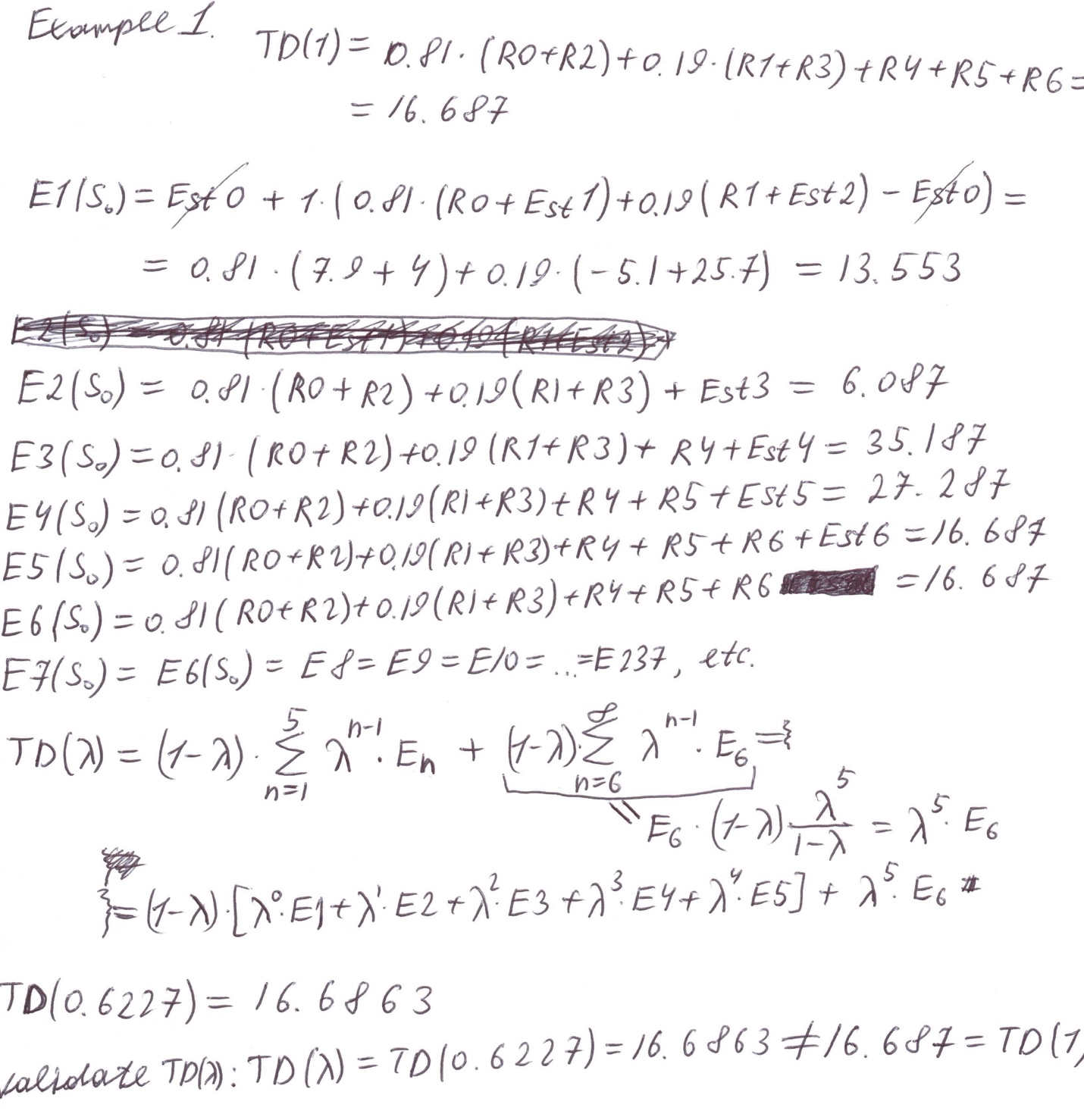


**[Kirill Luzin](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)**

[4 hours ago](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)

From S&B "After a terminal state has been reached, all subsequent n-step returns are equal to the conventional return, Gt.", so

E6,E7,E12,E237 = 0.81\*(R0+R2)+0.19\*(R1+R3) + R4 + R5 + R6. Since they are all the same we can simplify the infinite sum. Below is my pseudocode, which leads to a very close answer by not the correct one. May you please point me in the direction to see where my error is.  
  
In the last line, I try to confirm (and fail) that TD(0.6227) = TD(1), what means my formula is missing something.



[helpful!](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1) 1



**[Evelyn](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)**

[3 hours ago](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)

In hw2 shouldn't T =5 and n goes from 1 to 4?

[helpful!](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1) 0



**[Kirill Luzin](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)**

[2 hours ago](https://piazza.com/class/k4ws6a8i2uu71u?cid=130_f1)

Nope... we should calculate Ek till the terminal state is reached. So, with E1 we reach S1 and S2; E2 reaches S3; E3 reaches S4; E4 reaches S5; E5 reaches E6.  
S&B: **After a terminal state** has been reached, all **subsequent** n-step returns are equal to the conventional return, Gt and post-termination terms **can be separated from the main sum**.