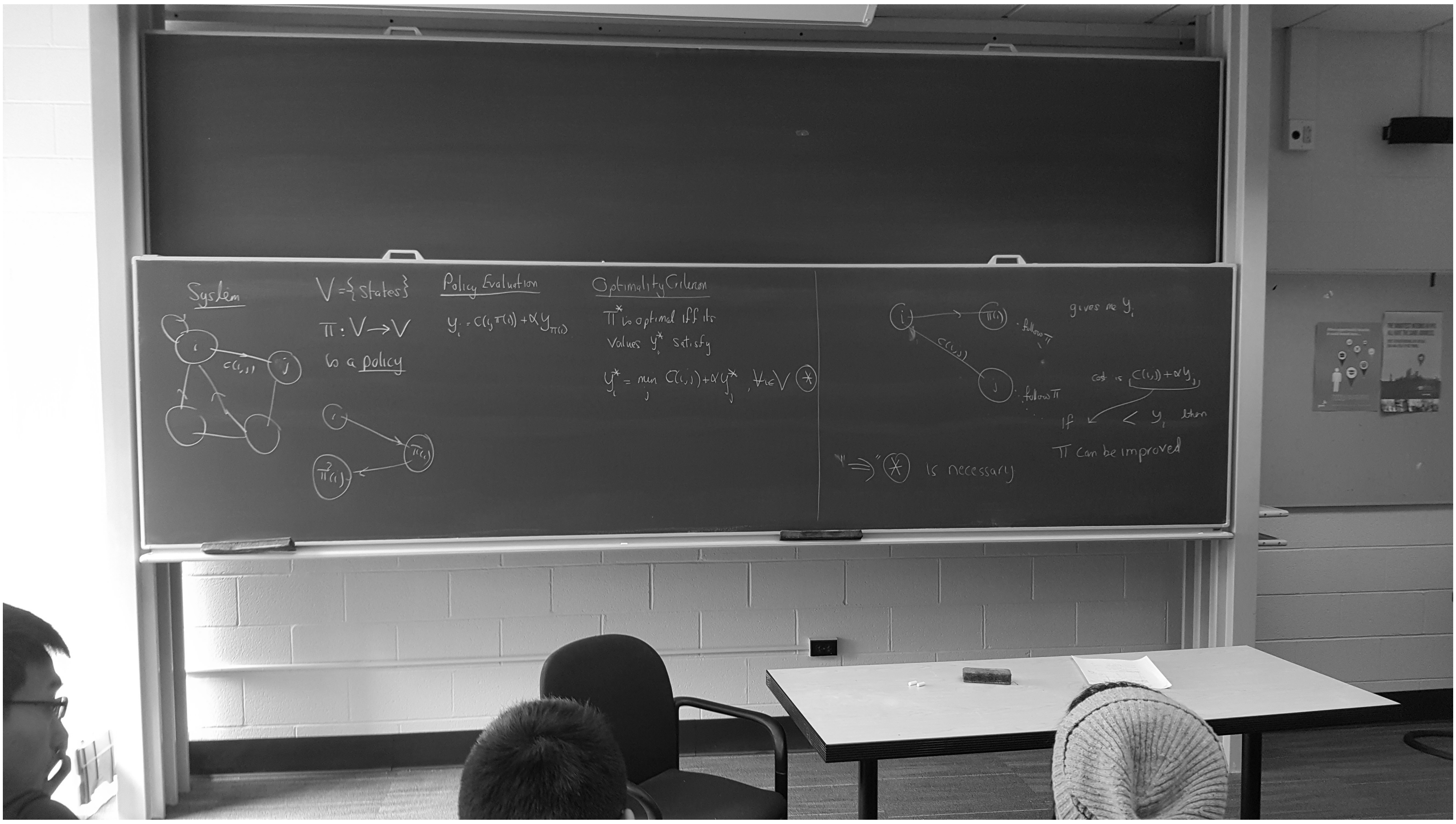
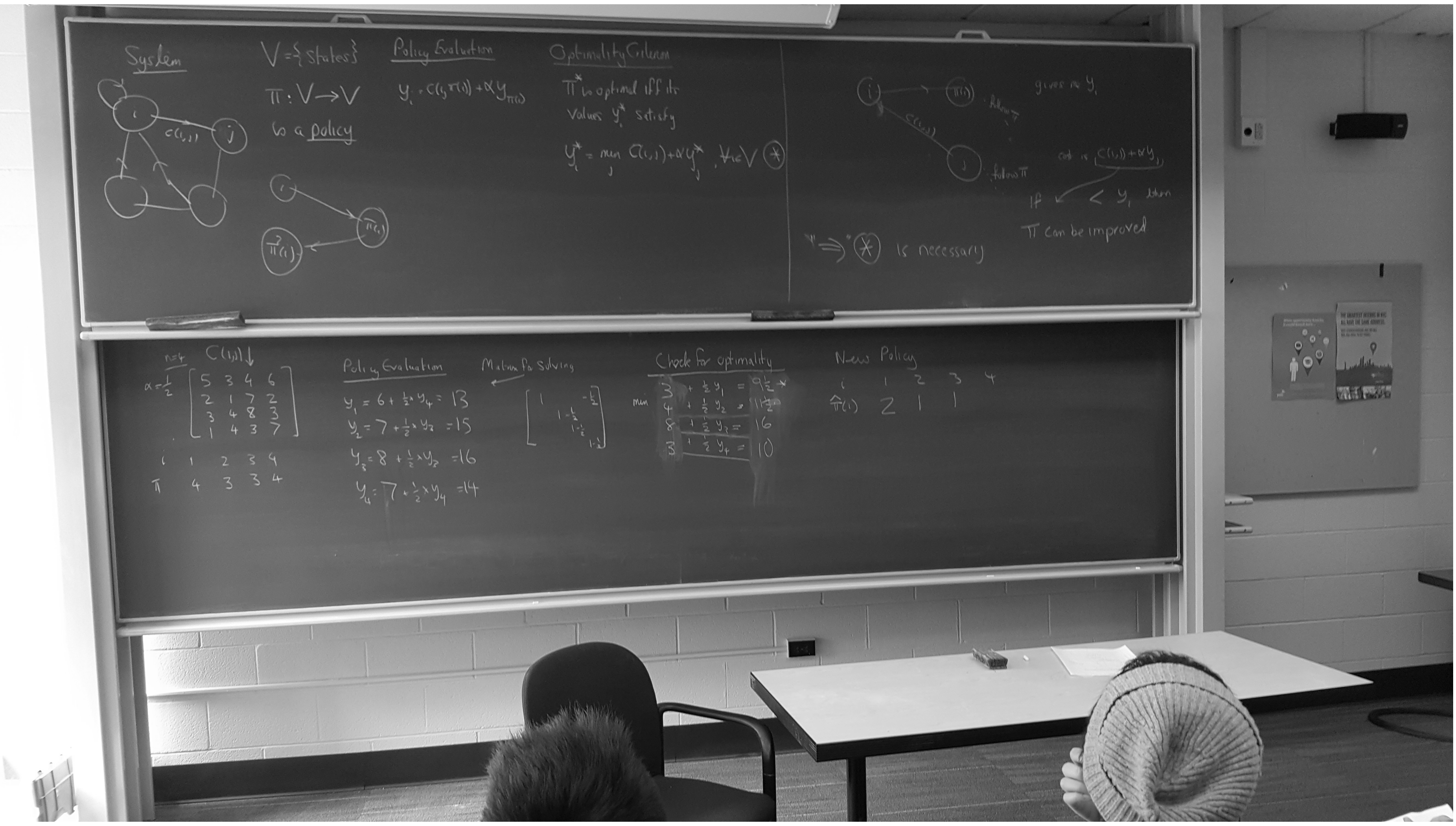


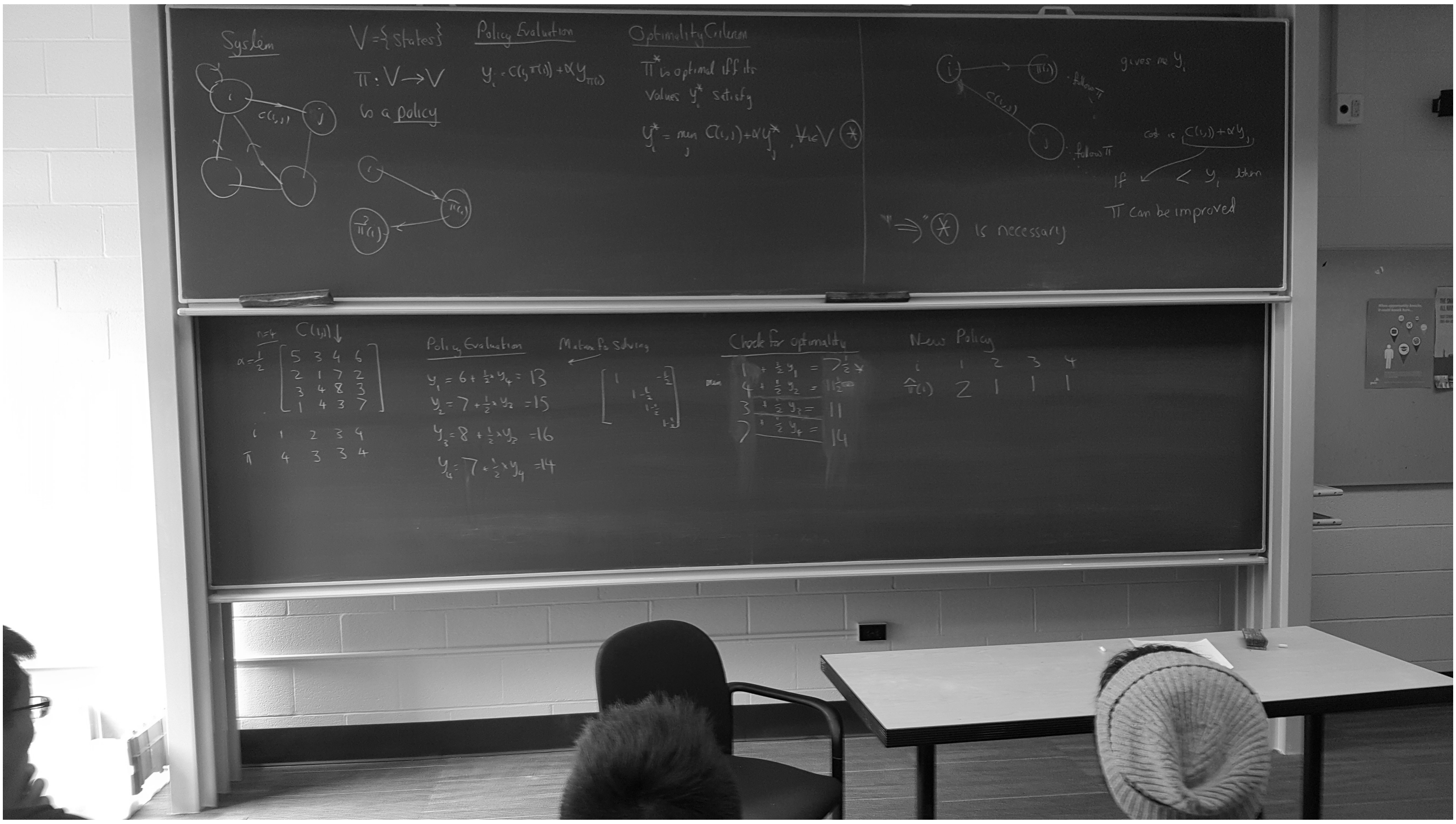
System
c(i,j)
V = {states}
π : V → V
is a policy
π²(i)
Policy Evaluation
y_i = c(i, π(i)) + α y_π(i)
Optimality Criterion
π* is optimal iff its
values y*_i satisfy
y*_i = min_j c(i,j) + α y*_j , ∀i∈V (*)
gives me y_i
follow π
c(i,j)
follow π
cost is c(i,j) + α y_j
If < y_i then
π can be improved
"⇒" (*) is necessary


System
V = { states }
π : V → V
is a policy
Policy Evaluation
Optimality Criterion
π* is optimal iff its
values y* satisfy
is necessary
π can be improved
Policy Evaluation
Matrix for Solving
Check for optimality
New Policy
y1 = 6 + ½ × y4 = 13
y2 = 7 + ½ × y2 = 15
y3 = 8 + ½ × y3 = 16
y4 = 7 + ½ × y4 = 14

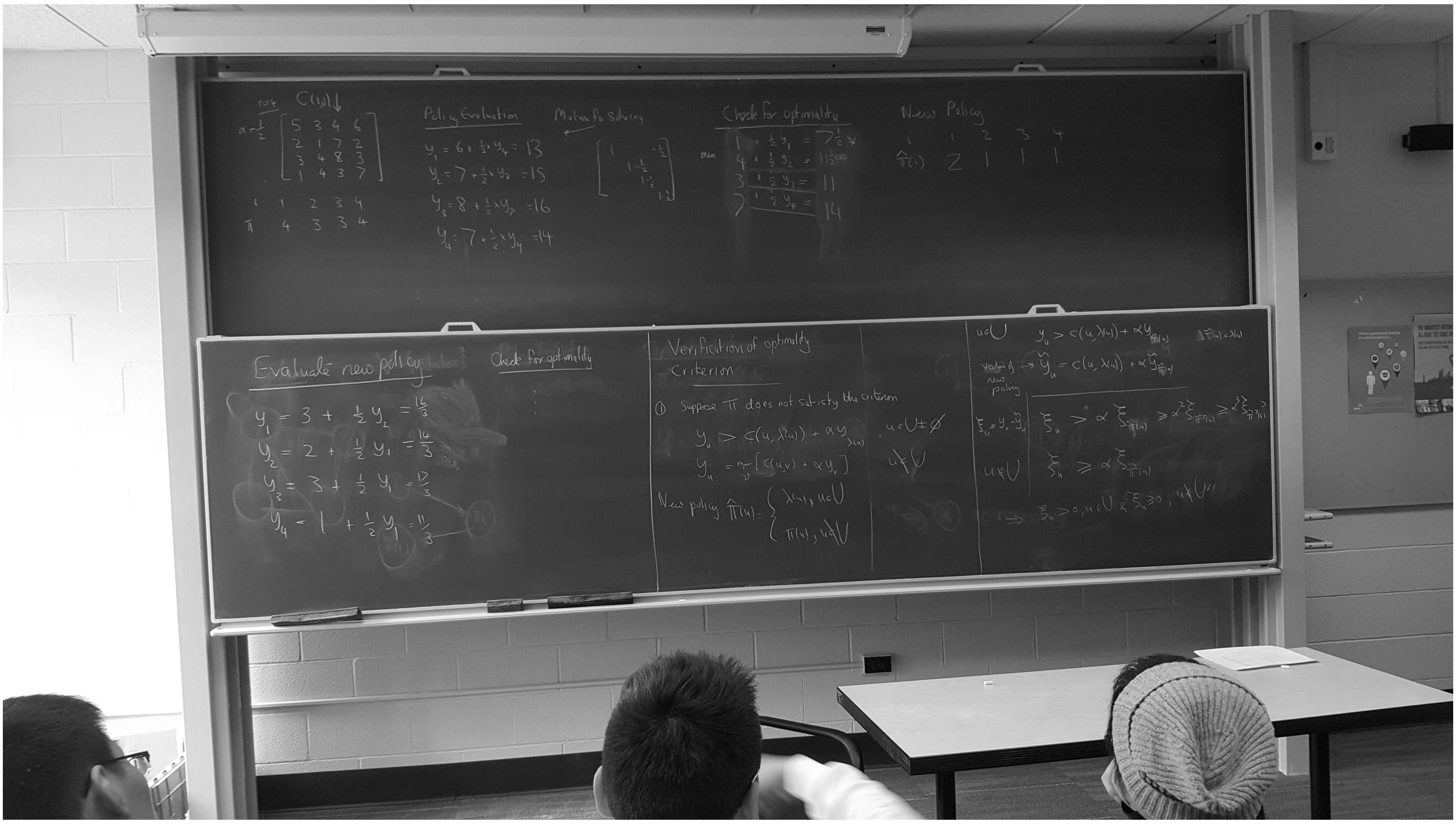

Π : V → V
is a policy
Π* is optimal iff its
Values y* satisfy
gives me y
Π can be improved
is necessary
Policy Evaluation
Matrix for solving
Check for optimality
New Policy

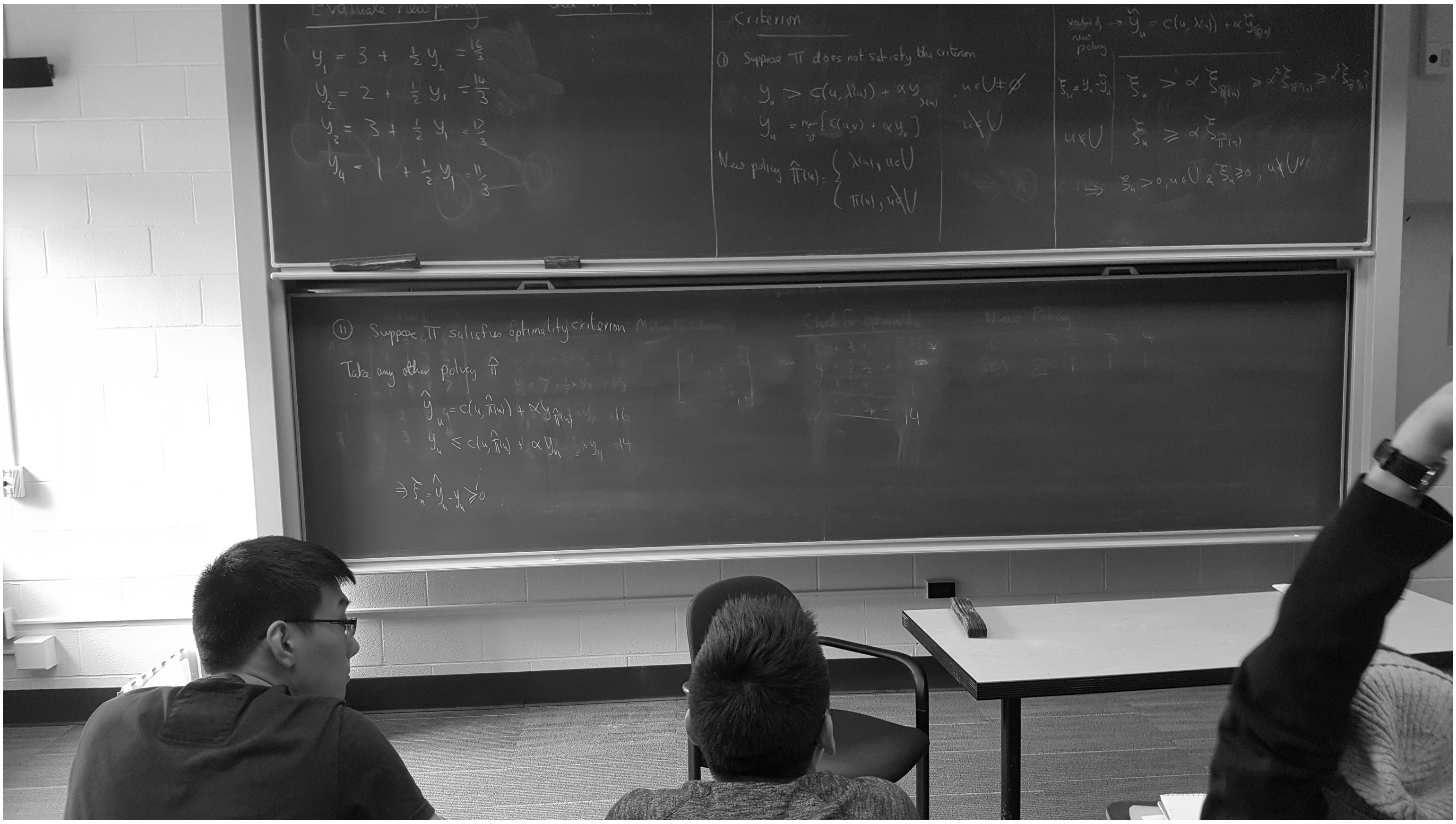


System
V = { states }
π : V → V
is a policy
Policy Evaluation
y_i = c(i, π(i)) + α y_π(i)
Optimality Criterion
π* is optimal iff its
Values y*_i satisfy
y*_i = min_j c(i,j) + α y*_j , ∀i∈V
is necessary
π can be improved
n=4
C(i,j)
α = 1/2
Policy Evaluation
Matrix for solving
y1 = 6 + 1/2 × y4 = 13
y2 = 7 + 1/2 × y2 = 15
y3 = 8 + 1/2 × y2 = 16
y4 = 7 + 1/2 × y4 = 14
Check for optimality
New Policy



System
c(i,j)
V = {states}
π : V → V
is a policy
Policy Evaluation
y_i = c(i, π(i)) + α y_π(i)
Optimality Criterion
π* is optimal iff its
values y_i* satisfy
y_i* = min_j c(i,j) + α y_j*, ∀i ∈ V (*)
gives me y_i
follow π
c(i,j)
follow π
cost is c(i,j) + α y_j
if < y_i then
π can be improved
"⇒" (*) is necessary
n=4
C(i,j)
α = 1/2
5 3 4 6
2 1 7 2
3 4 8 3
1 4 3 7
i 1 2 3 4
π 4 3 3 4
Policy Evaluation
y_1 = 6 + 1/2 × y_4 = 13
y_2 = 7 + 1/2 × y_2 = 15
y_3 = 8 + 1/2 × y_3 = 16
y_4 = 7 + 1/2 × y_4 = 14
Matrix for solving
Check for optimality
New Policy
i 1 2 3 4
π̂(i) 2 1 1 1


Policy Evaluation
Method for solving
Check for optimality
New Policy
Evaluate new policy
Check for optimality
Verification of optimality Criterion
Suppose π does not satisfy the criterion
New policy



Criterion
Suppose π does not satisfy the criterion
New policy
Suppose π satisfies optimality criterion
Take any other policy