## E: Applying imperfect information analysis, part 2

Example: using AI in sentencing, continued

Tree from last time:



Probabilities of reports: Prob rR = 0.375 Prob rN = 0.625

Conditional probabilities: Received rR: Prob R given rR = 0.6 Prob N given rR = 0.4 Received rN: Prob R given rN = 0.04 Prob N given rN = 0.96

Tree if AI says defendant is dangerous (rR):



EV of I = 0.6\*(-5) + 0.4\*(-505) = -205 EV of P = 0.6\*(-505) + 0.4\*(-5) = -305



Conclusion: Incarcerate if rR

Mistakes here per 100 initial defendants:

• 15 type N people are classified as R (risk overestimated)

• Given incarceration (too severe)





Mistakes here per 100 initial defendants:

- 2.5 R people are classified as N (risk underestimated)
- Given probation (too lenient)

Finally, stepping back to decision about the test:



Calculating the EV of the algorithm:



Conclusion: Would use the system

Errors:

- RP: 2.5 per 100 are too lenient (much lower, was 25)
- NI: 15 per 100 are too severe (substantially higher, was 0)

Could incorporate additional considerations:

- What if error rates depend on race of defendant?
- Also, AI algorithms often lack transparency: Difficult to say why a defendant was classified in a particular way