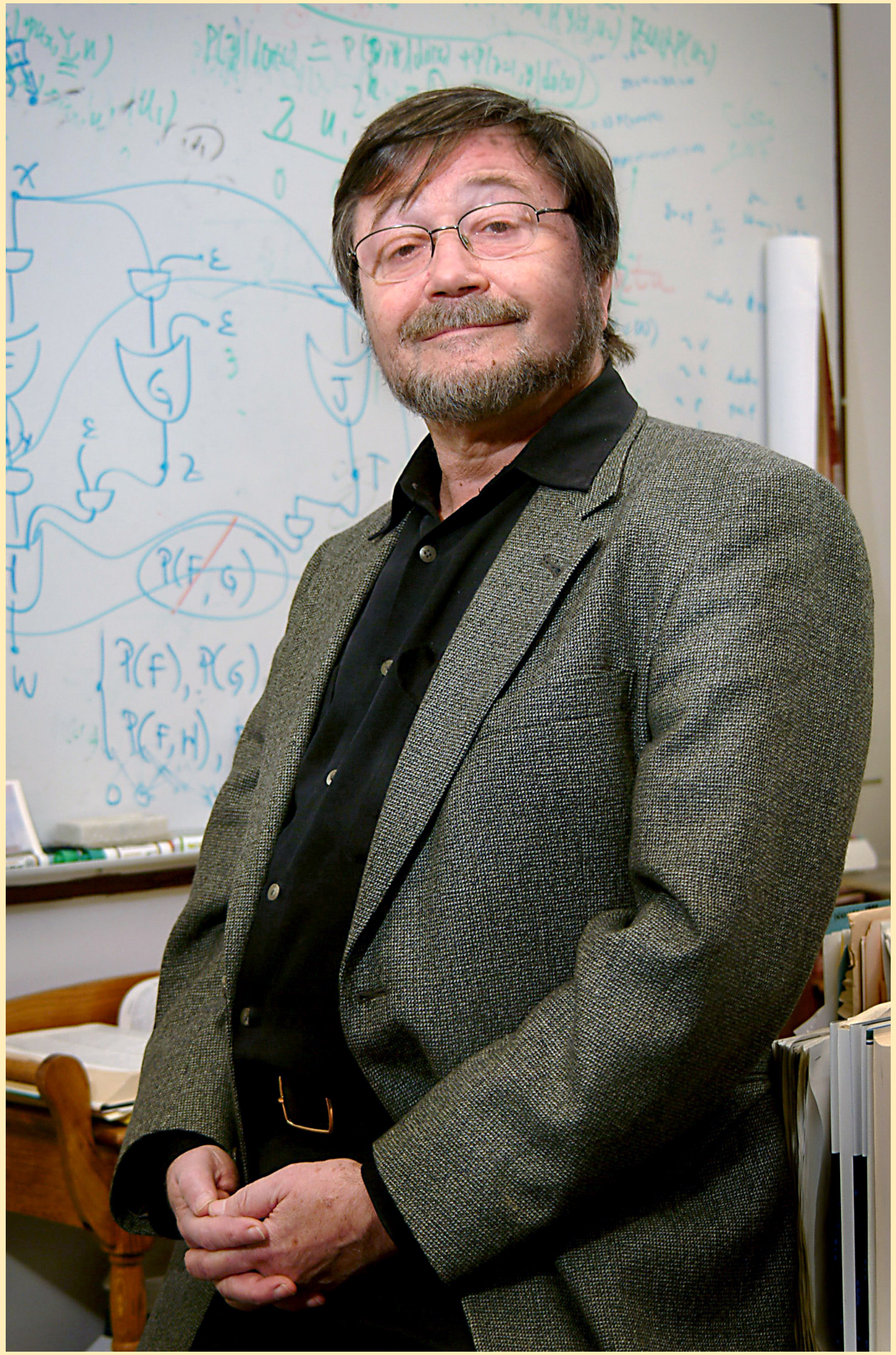
Directed Acyclic Graphs

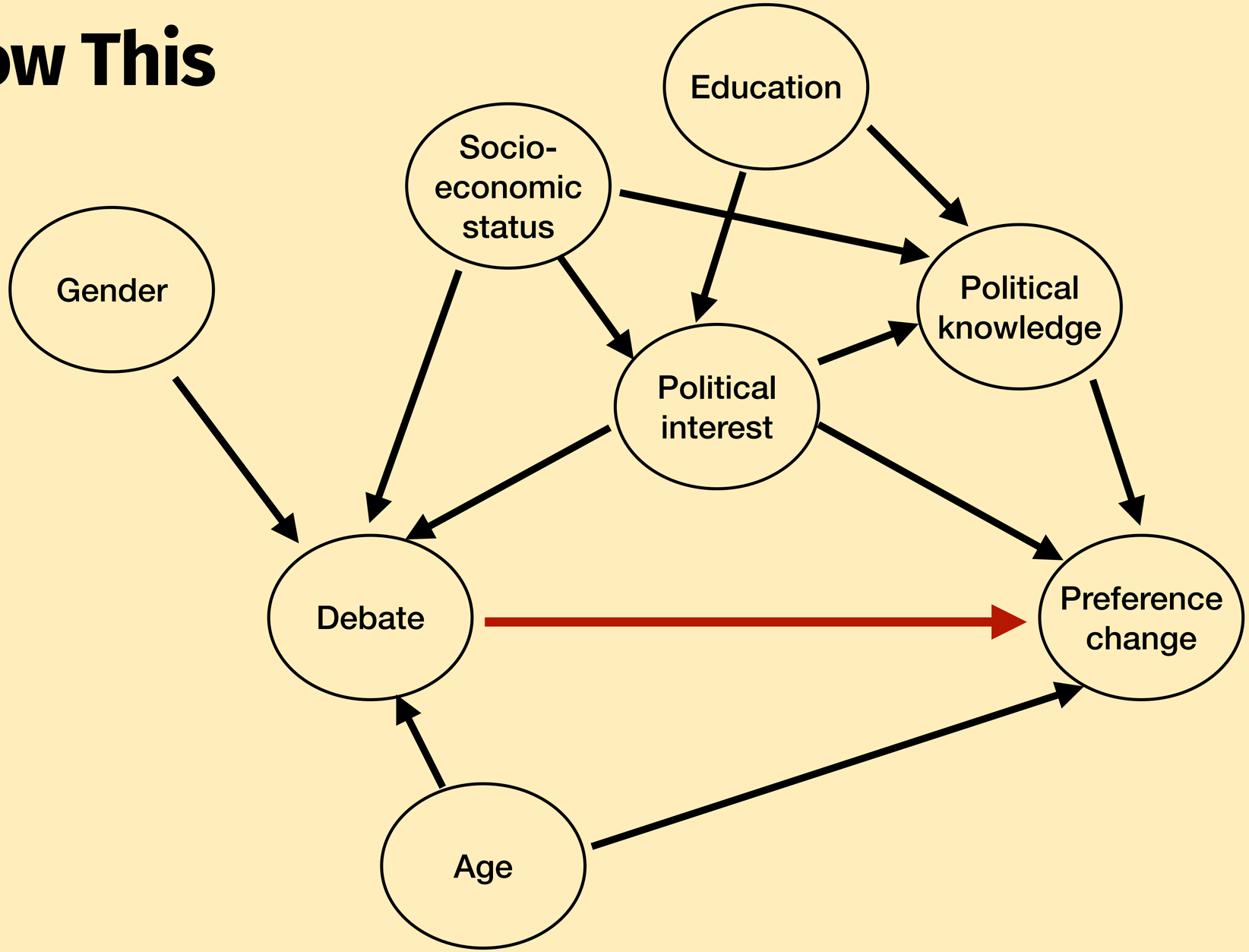
Aleš Vomáčka

Directed Acyclic Graphs

- Developed by Judea Pearl
- Second most popular approach (after Potential outcomes)
- Based on graph theory



We Already Know This



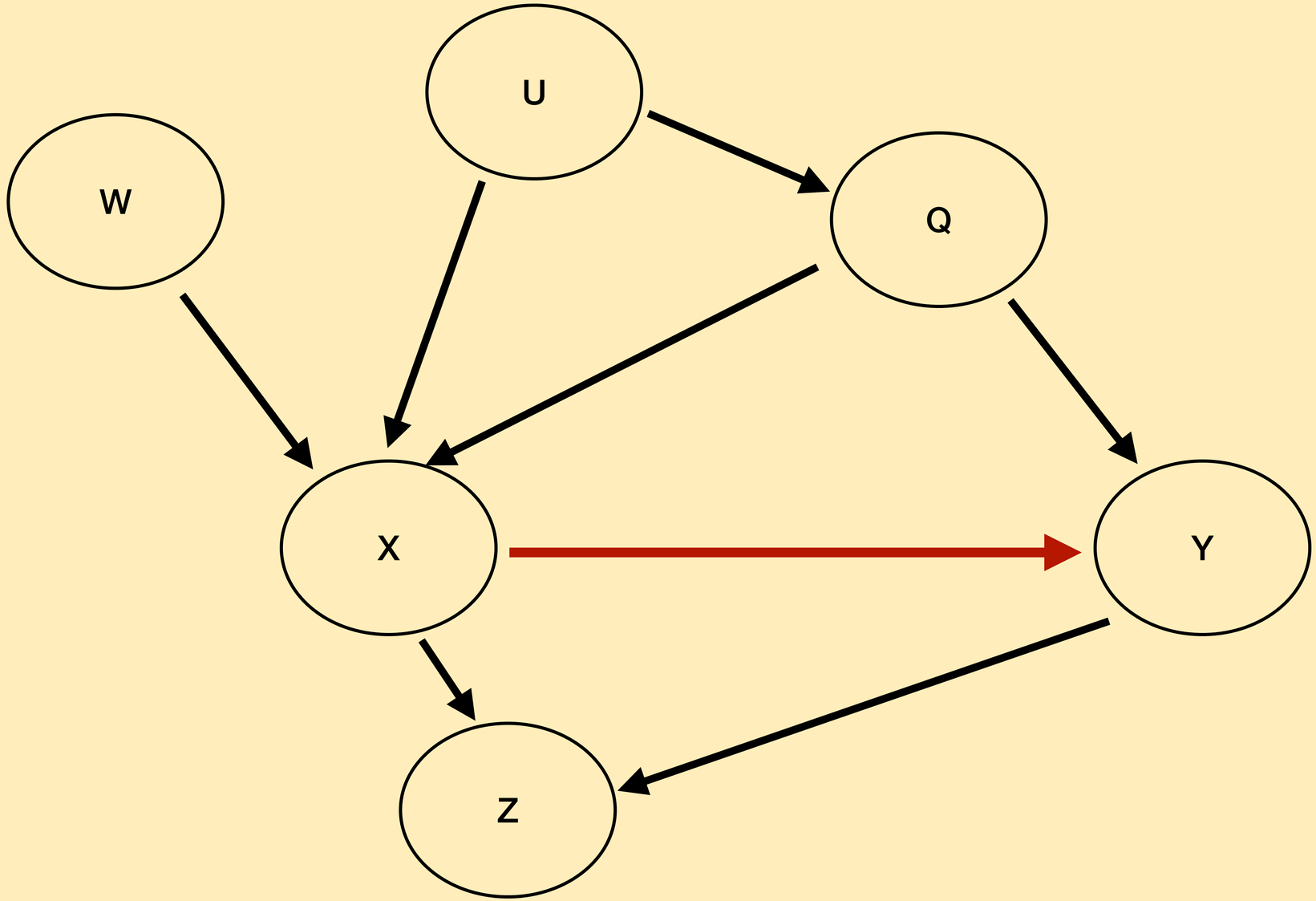
What to Read

- Popular, easy to read:
  - Pearl, J., & Mackenzie, D. (2018). The Book of Why: The New Science of Cause and Effect (1st edition). Basic Books.
- More in-depth, harder:
  - Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal Inference in Statistics—A Primer* (1st edition). Wiley.

- But don't follow him on Twitter

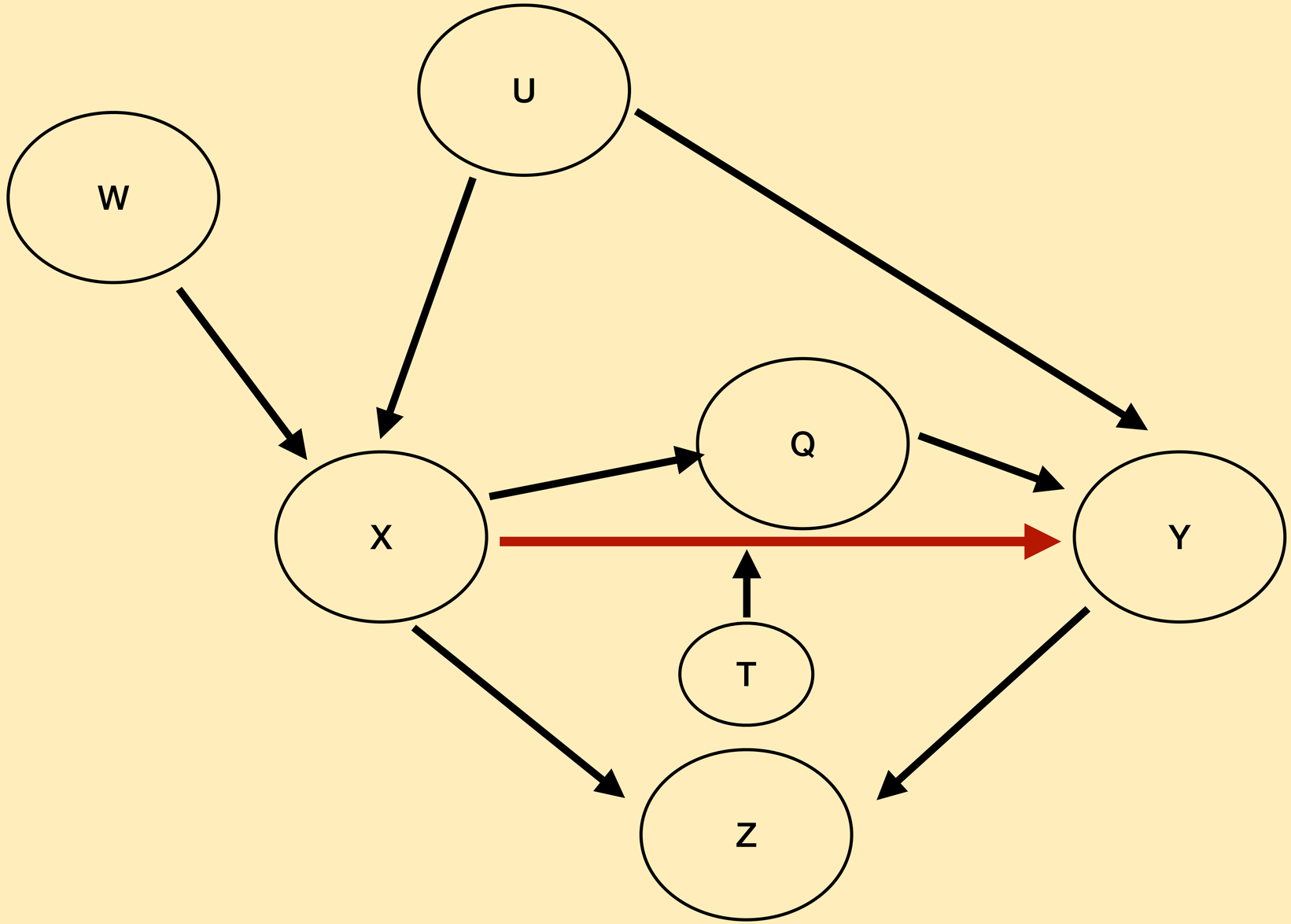
DAGs

- Directed - causal effects have direction
- Acyclic - No variable can cause itself
- Graph - fancy word for a network



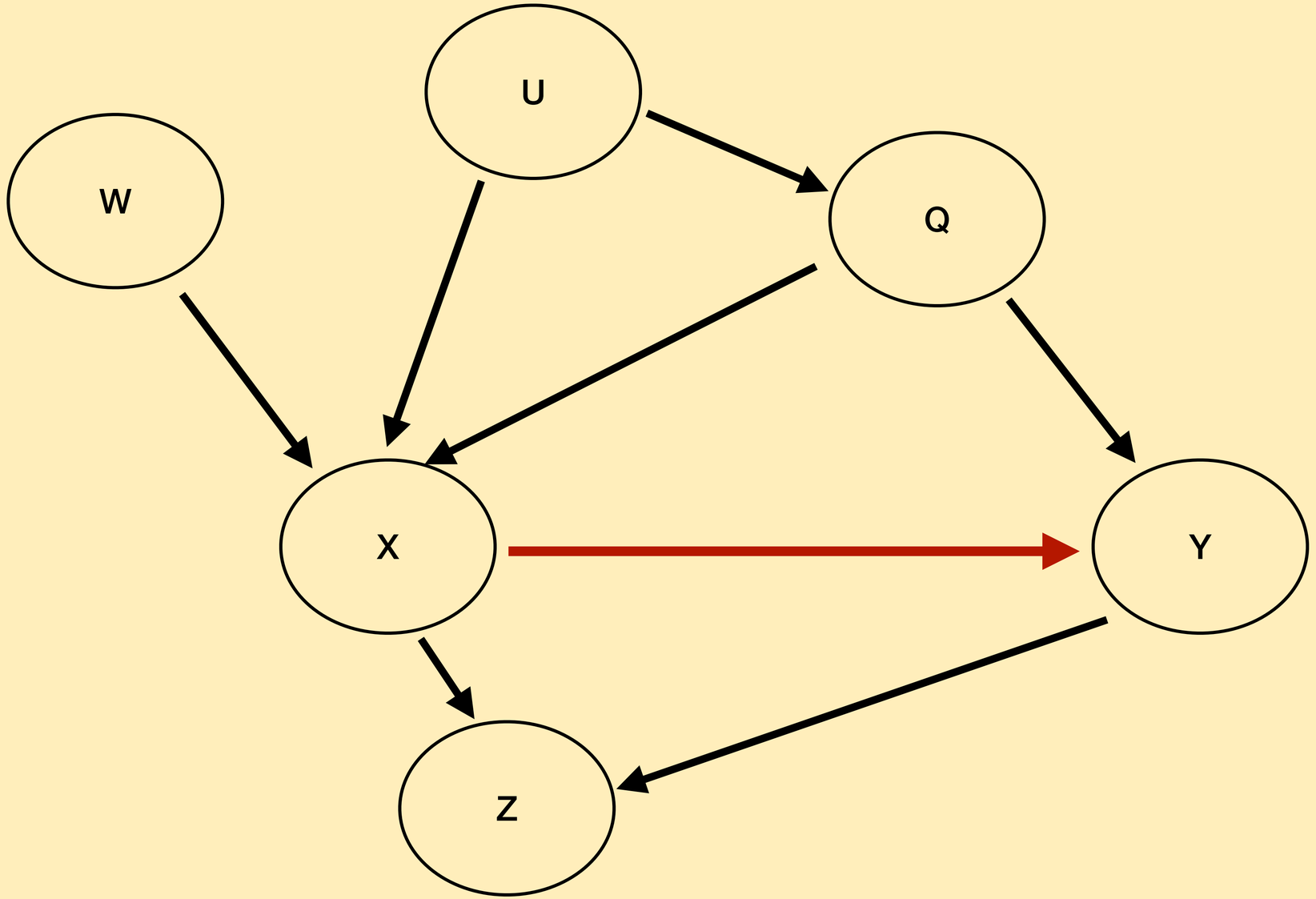
Variables in the Wild

- Confounders - common parents
- Colliders - common children
- Mediators - middle steps
- Moderators - interactions



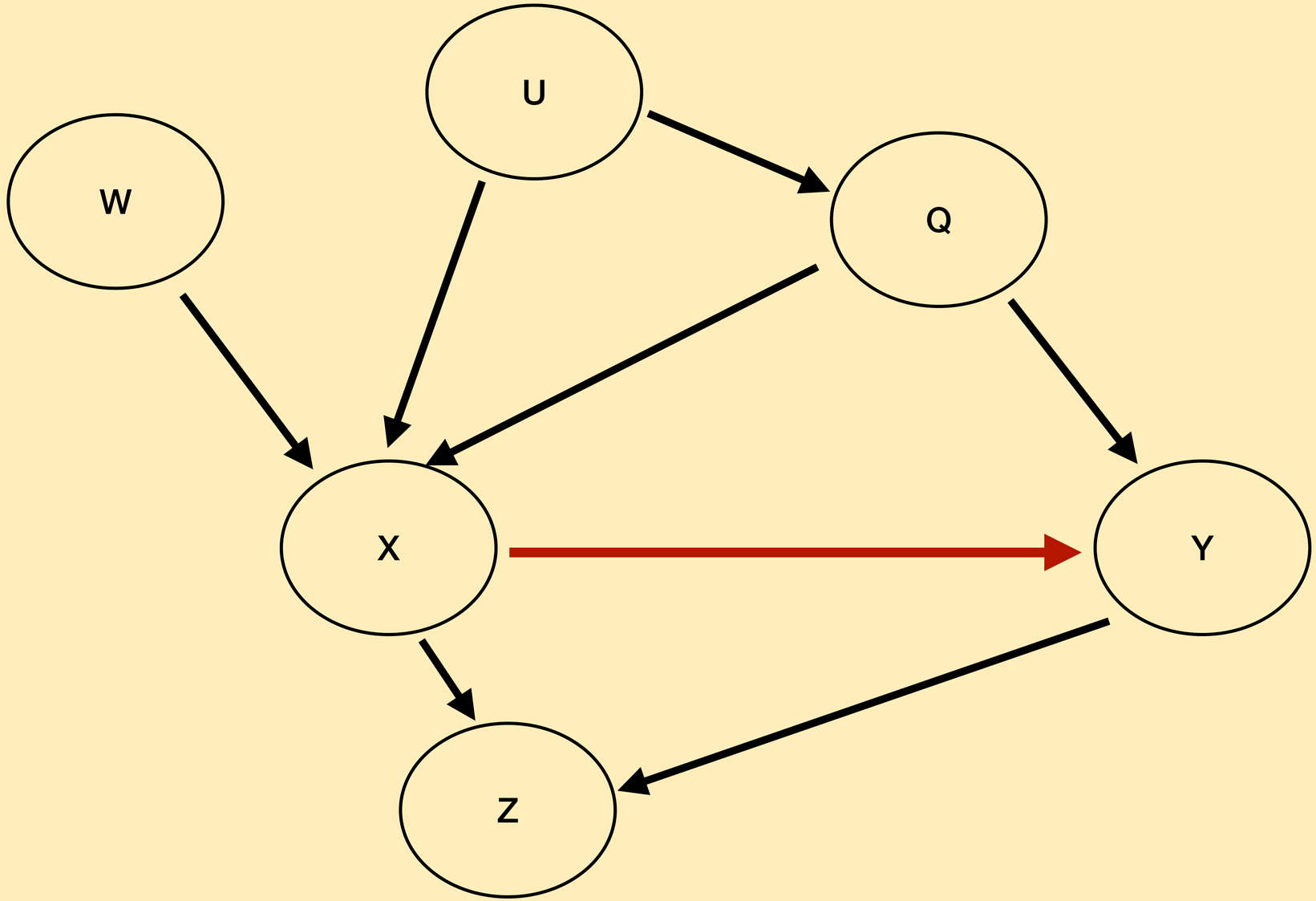
Backdoor Criterion

- Math behind DAGs based on do-calculus.
- Most important is the backdoor criterion.



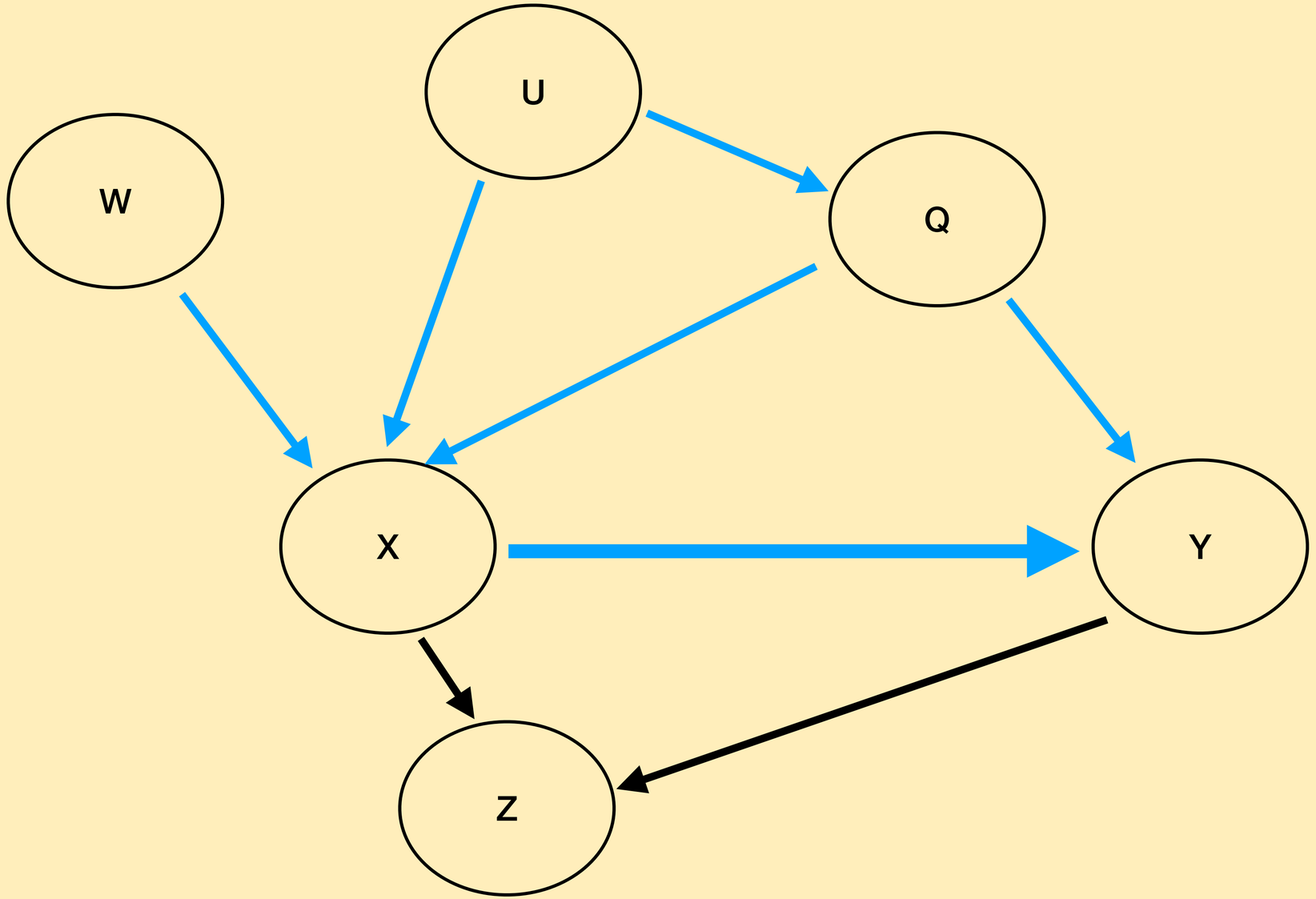
Backdoor Criterion

- Backdoor criterion - To estimate causal effect of X on Y, we need to close all backdoor paths between them.



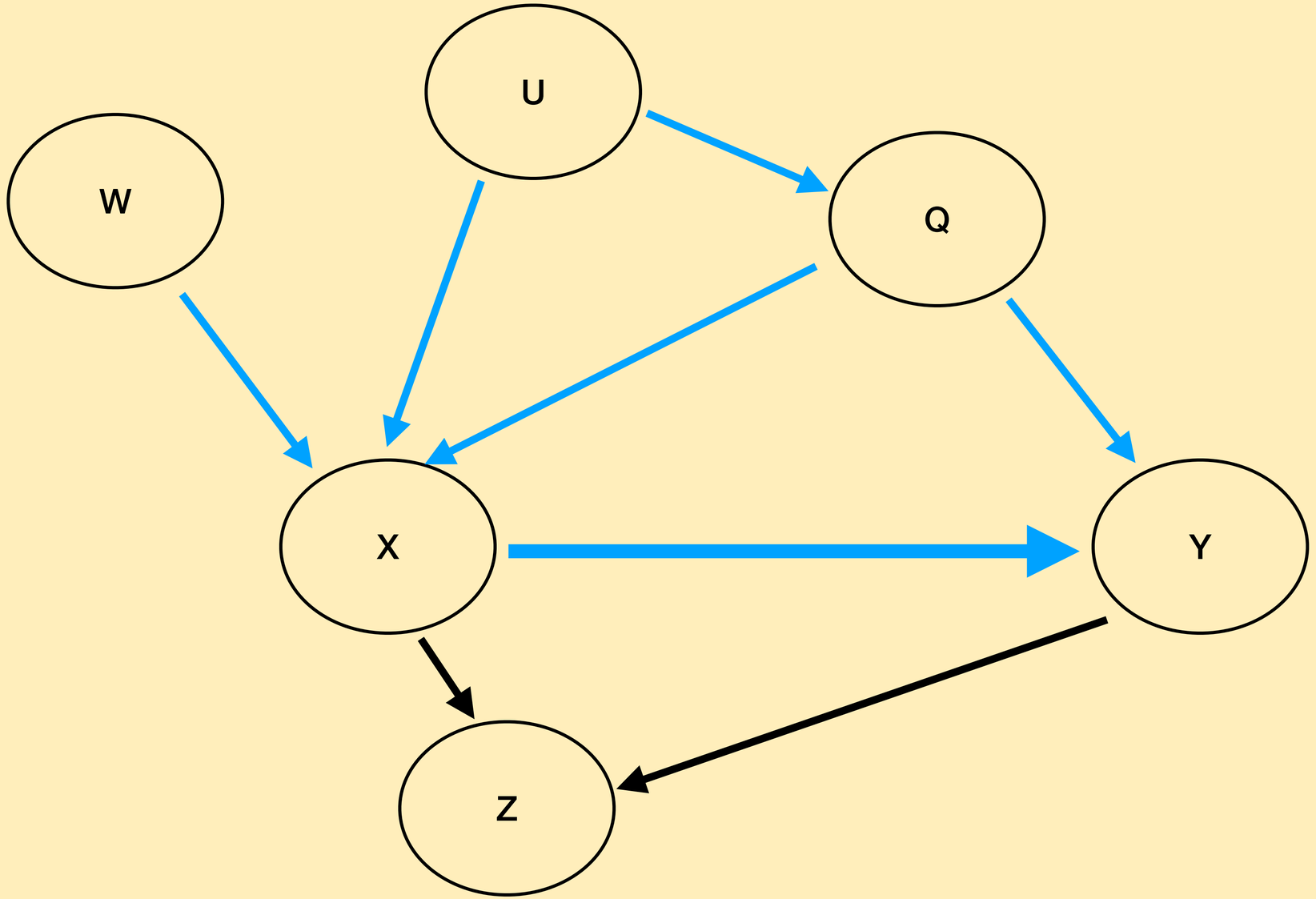
Backdoor Criterion

- Kinda like a system of ponds and rivers.
- We want for water to flow only directly from X to Y, no indirect flows!
- By controlling for variables, we open/ close floodgates.



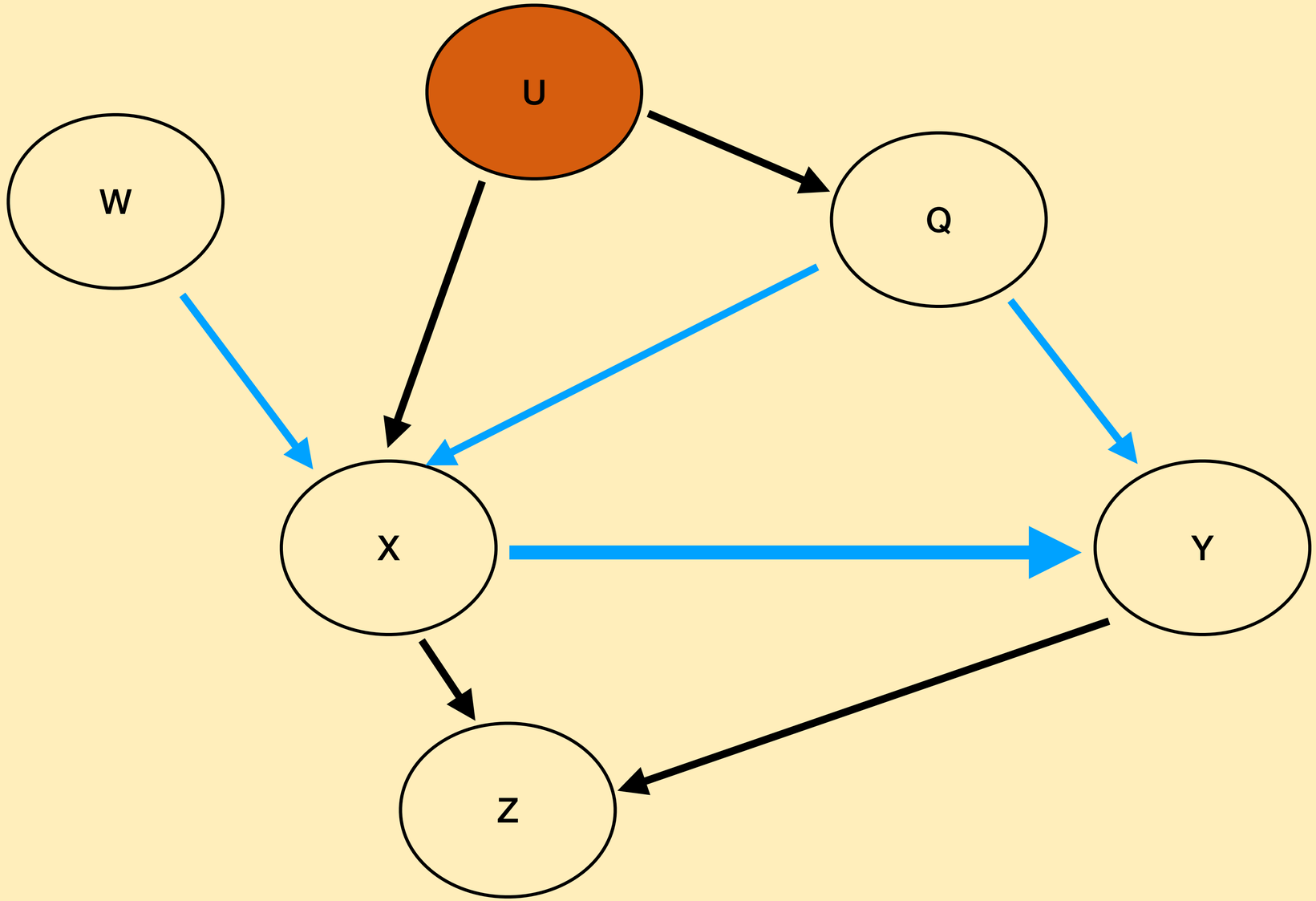
Backdoor Criterion

- Ponds where waters flow outside (confounders) or through (mediators) start open.
- Ponds where the currents collide (colliders) start closed.



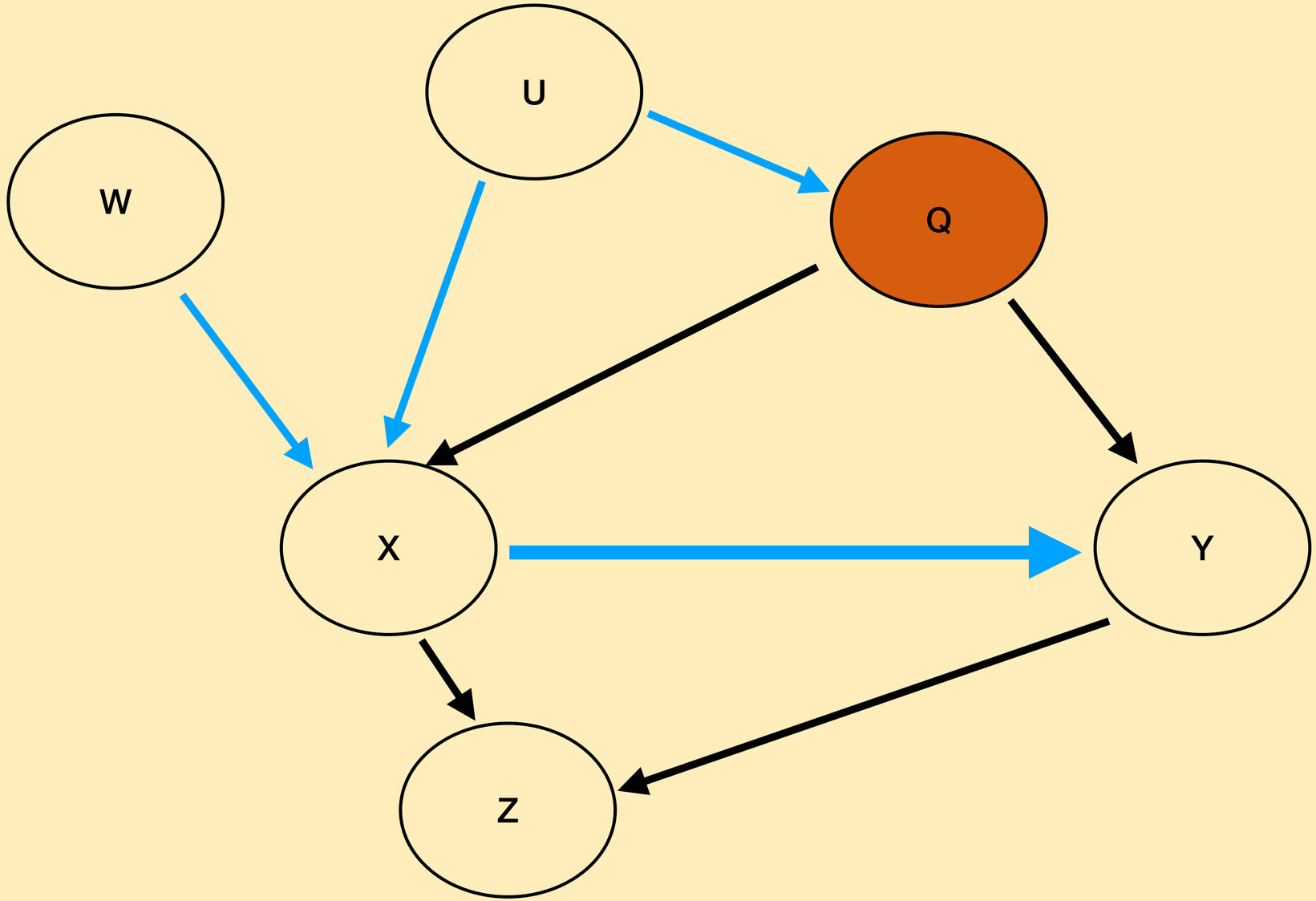
Backdoor Criterion

- Controlling for U closes the flow from X->U->Q
- But X->Q->Y is still open!



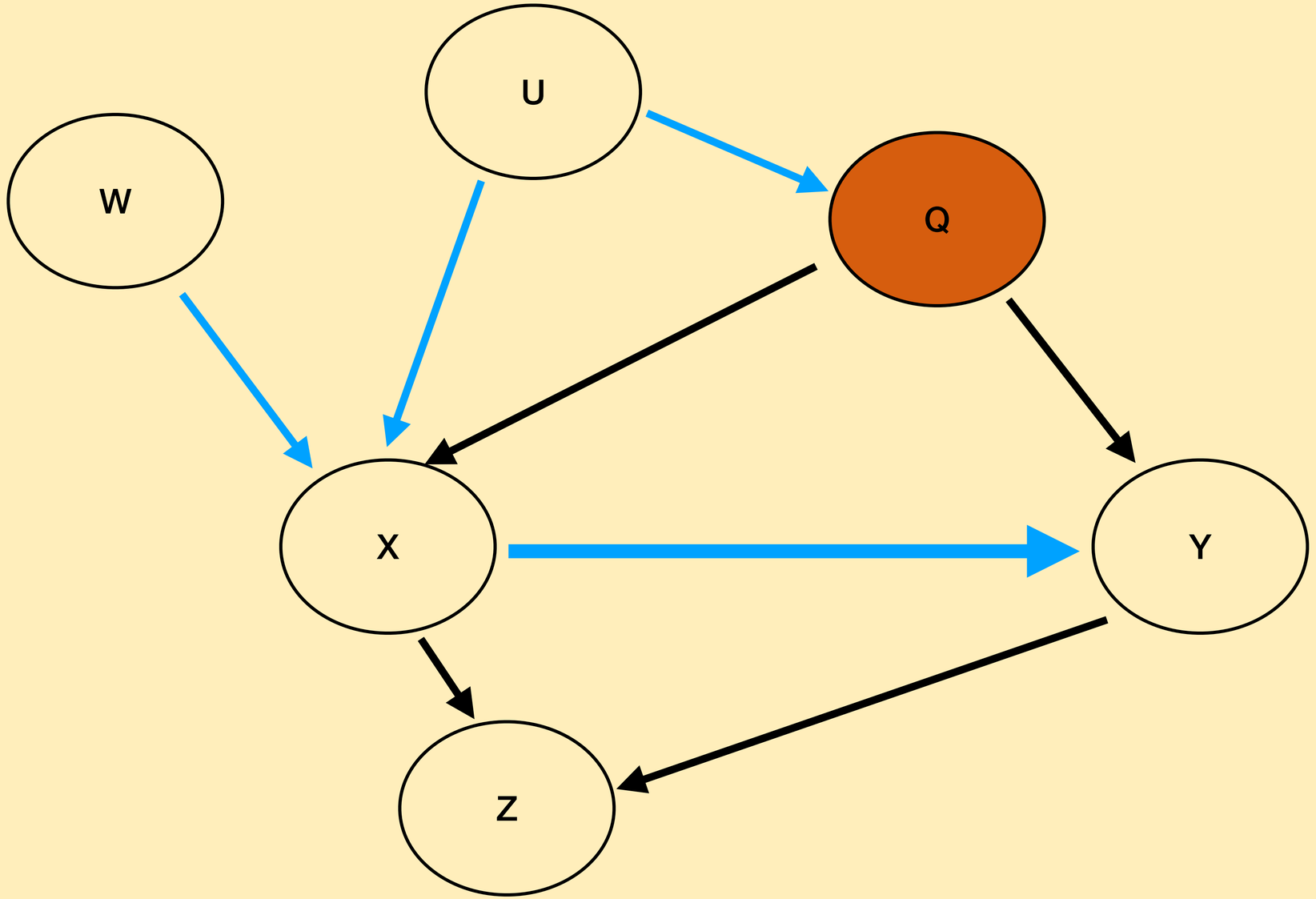
Backdoor Criterion

- Controlling for Q closes flow both through Q an U.
- Q is the only variable we need to control for.



Backdoor Criterion

- We don't care about W (no flow to Y)
- Don't control for Z, you'd opened a backdoor path!



Questions?