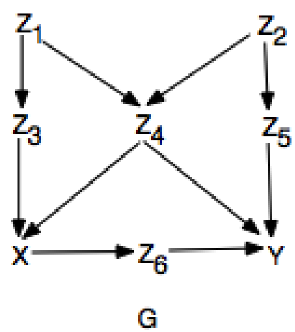
Ralph's covariate selection

Ralph is struggling with appropriate covariate selection to address potential confounding in assessing causal effects. Ralph is considering the following DAG



DAG G

The covariate adjustment algorithm (derived from the rules of do-calculus) is instructive.

Let Z refer to a potential set of sufficient covariates for assessing the effect of X on Y.

1. Z should not be a descendant of X.
2. Delete all non-ancestors of $\{X, Y, Z\}$.
3. Delete all arcs emanating from X.
4. Connect any two parents sharing a common child.
5. Strip arrow-heads from all edges.
6. Delete Z.

Test: If X and Y are disconnected in the remaining graph, then Z is a sufficient set of covariates.

Suggested:

1. Utilize the covarite adjustment algorithm and compare with the rules of do-calculus (see Ralph's back-door adjustment) to determine if $\{Z_3, Z_4\}$, $\{Z_4, Z_5\}$, $\{Z_3, Z_4, Z_5\}$, or $\{Z_4, Z_6\}$ is a sufficient set of covariates to enable identification of the causal effect of X on Y?

2. For any sufficient sets identified in 1, indicate if they involve back-door adjustments, front-door adjustments, neither, or both.

3. Is it necessary to include Z_1 or Z_2 in the set of covariates? Explain.

4. What role, if any, does Z_6 play in the identification of the causal effect of X on Y? Hint: consider a two-step front-door adjustment: $X \rightarrow Z_6$ and $Z_6 \rightarrow Y$; see Ralph's front-door adjustment.