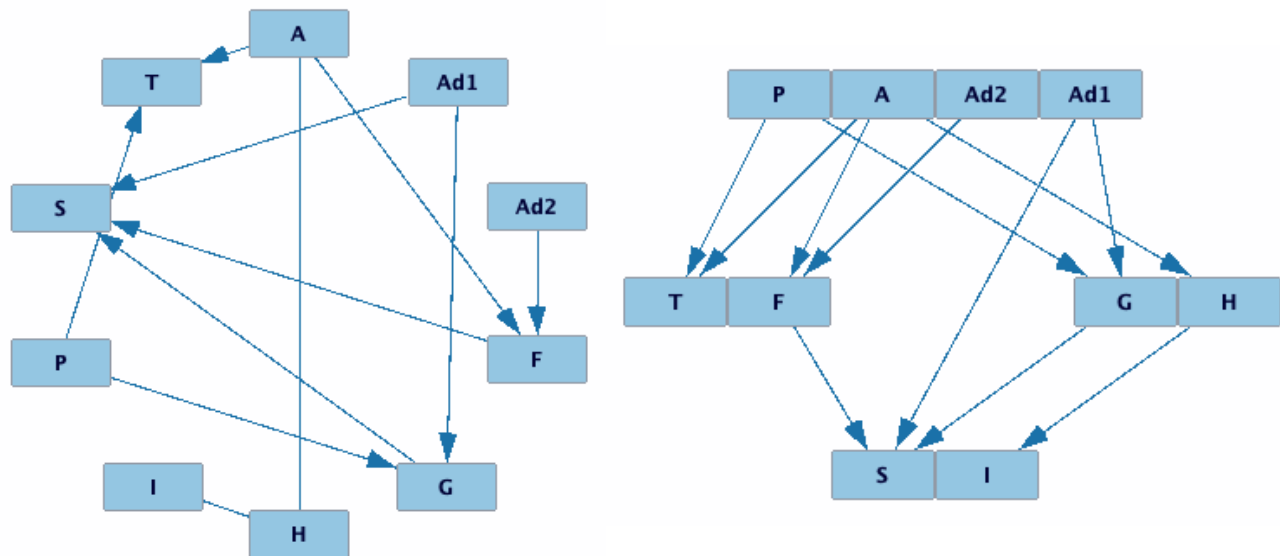


HW2 Report

Problem 1



The initial graph, including undirected edges AH and IH, and the final graph format which directs the two aforementioned edges AH and HI.

H and I both have to do with home ownership, so they have an obvious connection: one cannot own home insurance if they don't own a home. However, someone who owns a home may not necessarily have home insurance. The fact that homeowners can, but do not necessarily, have home insurance, might confuse the algorithm about which of the two variables "causes" the other. It's only through our contextual knowledge (the meaning of the variables) that we can presume that home ownership "causes" possession of home insurance.

A and H have a similar problem. Older people, because they typically have more financial stability, usually own their home (as opposed to renting it), but the data alone wasn't definitive enough to determine that link. At the same time, younger people are perfectly capable of home ownership, but may not be at the point where their financial stability allows them to own it.

Problem 2

$P(W) =$

W	$\phi(W)$
W(0)	0.3013
W(1)	0.6987

$P(X | W = 1) =$

X	$\phi(X)$
X(0)	0.3963
X(1)	0.6037

$P(Z | W = 0, X = 0) =$

Z	$\phi(Z)$
Z(0)	0.2615
Z(1)	0.7385

Problem 7

answer_7_1 = 20.779393798843557

answer_7_2 = 0.0

answer_7_3 = {'I': 1, 'P': 1, 'G': 1, 'F': 0, 'T': 1, 'S': 0, 'A': 0, 'H': 1}

Problem 8.1

- Learning a person's stance on gun control from Facebook is absolutely worth the squeeze. For \$2 per person, we would get \$20.78 worth of information, a profit of \$18.78.
- Learning the person's political affiliation, given their stance on gun control, is not worth the squeeze. The VPI is 0, which means that we would actually lose \$0.25 to Google.

Problem 8.2

These practices, to us, are considered unethical. Social media is a form of expression, and harvesting data through that, especially without the direct consent of its users, is an unethical practice. Furthermore, the data harvested from social media may not be indicative of the bigger picture, so the data itself may be inherently flawed, barring its unethical premise. This practice, as opposed to political advertising, is much more invasive due to the much larger pool of information that social media has access to.