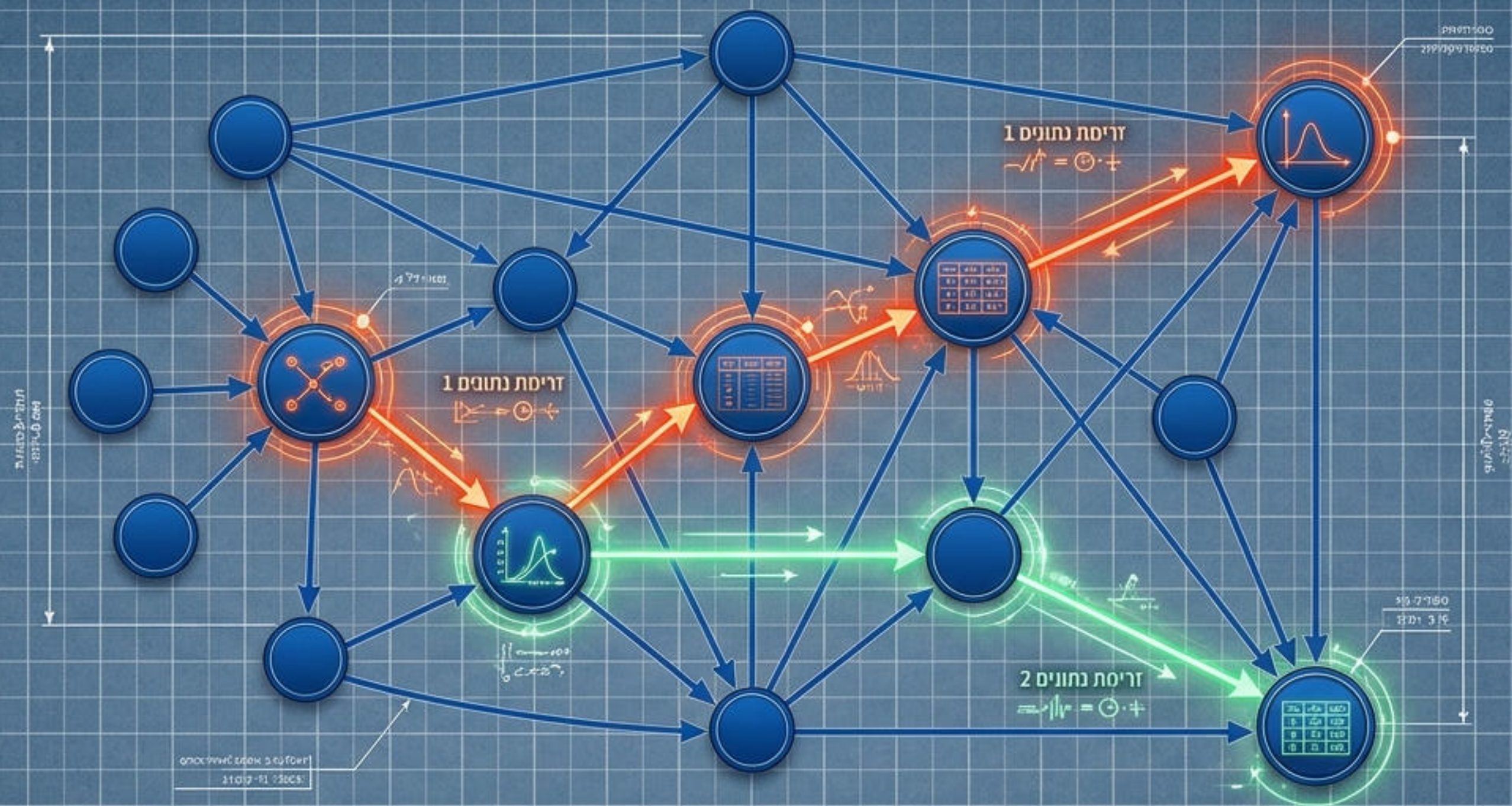
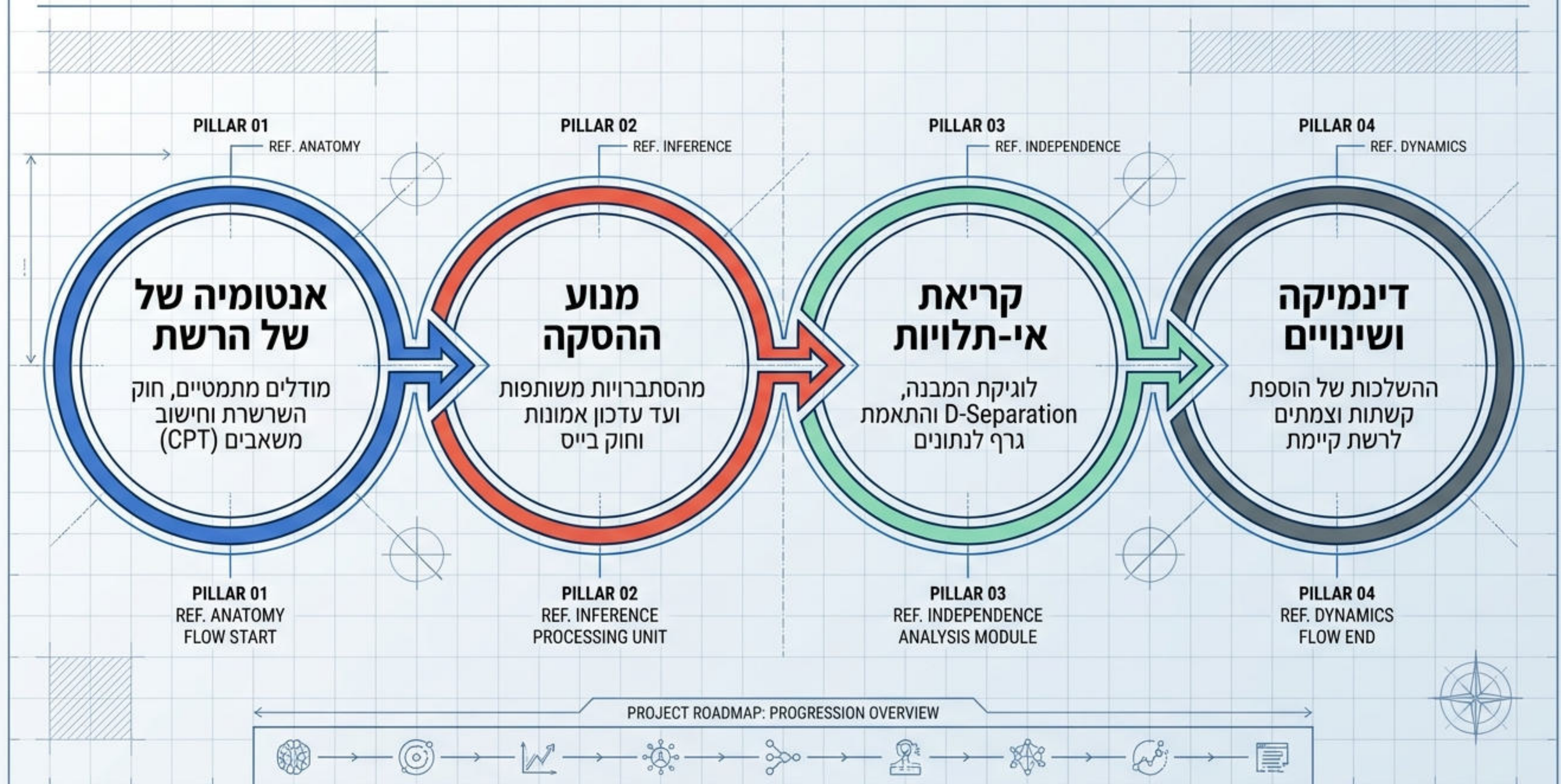




רשתות בייסיאניות: מתיאוריה להבנה עמוקה

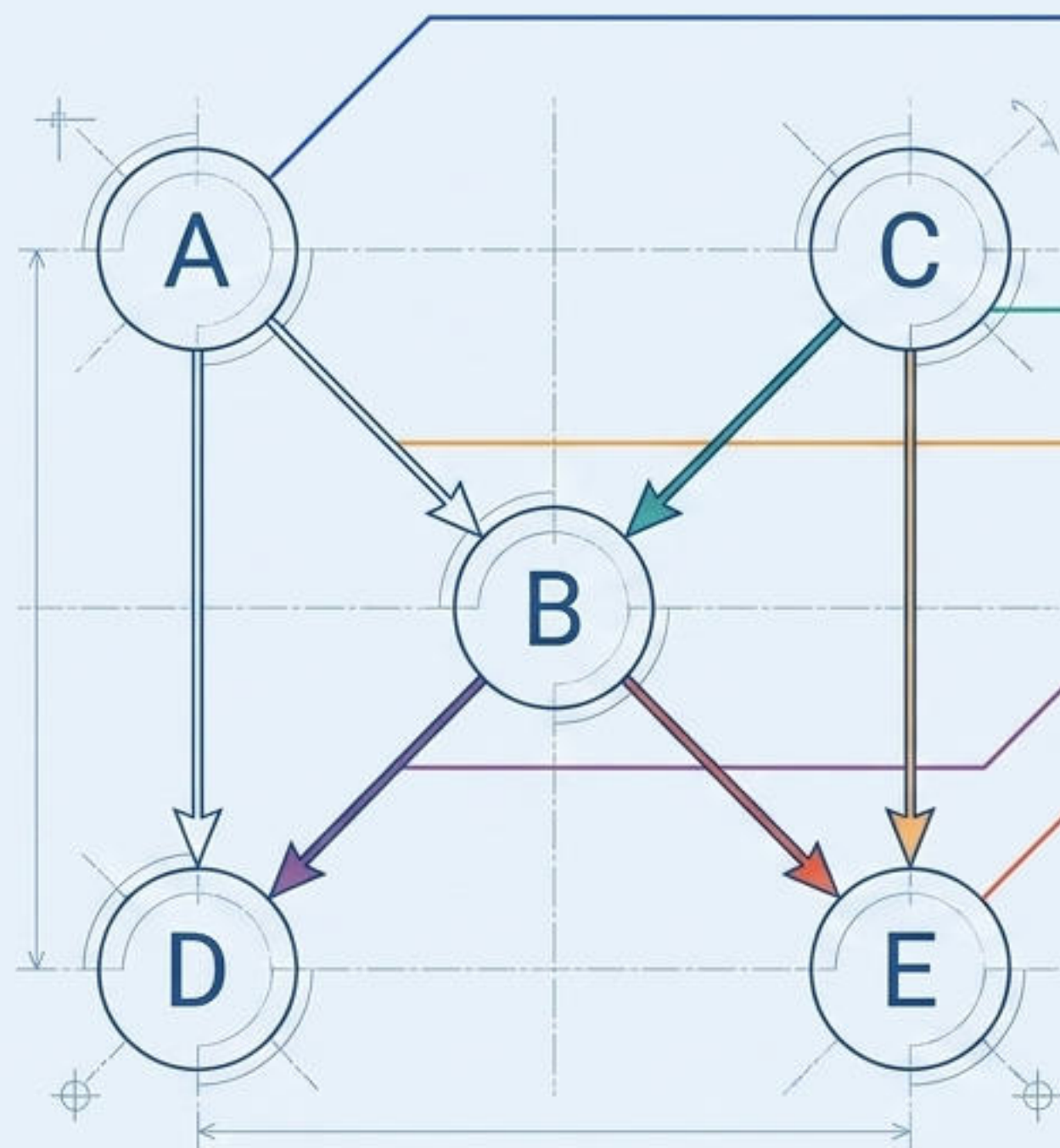
מדריך ויזואלי לסטודנט: מפענוח מבנה הרשת, דרך חישובי טבלאות, ועד להסקה הסתברותית מורכבת.

MASTERCLASS BLUEPRINT: THE FOUR PILLARS OF NETWORK STRUCTURE



Structural Blueprint

16:9



$$P(A,B,C,D,E) = P(A) * P(C) * P(B|A) * P(D|B) * P(E|B,C)$$

$$P(A, \neg S, H, E, \neg C) = P(A) * P(\neg S) * P(H) * P(E|A, \neg S, H) * P(\neg C|E, H)$$

חוק השרשרת: כל צומת מותנה אך ורק בהורים הישירים שלו.
הגרף הוא למעשה 'מפת הדרכים' של הנוסחה.

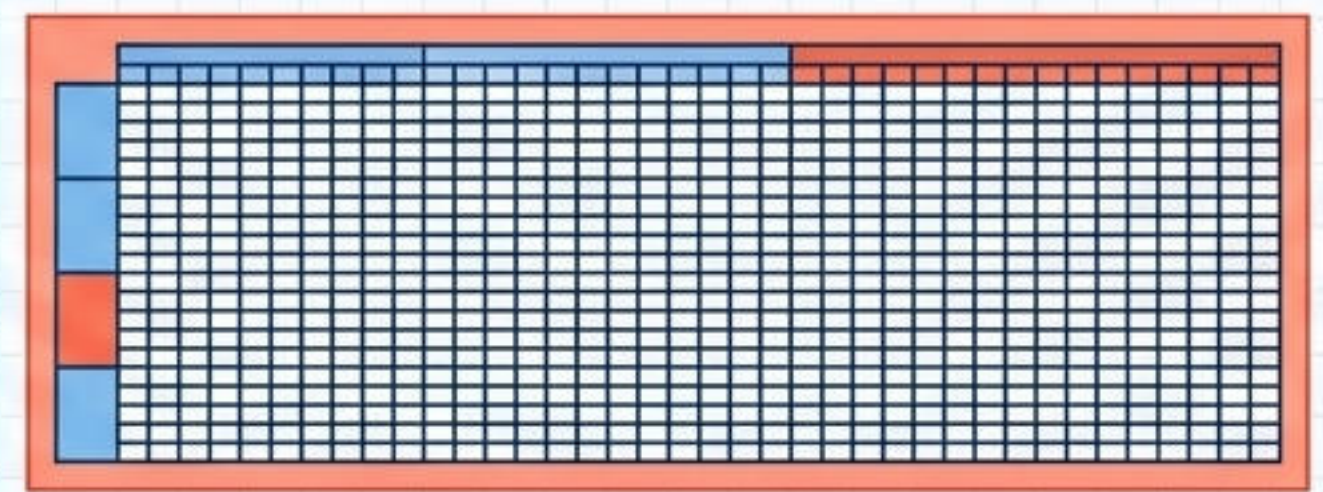
(מספר הערכים האפשריים - 1) × (מכפלת מספר הערכים של כל צומתי ההורה)

משתנים בינאריים (2 ערכים)

CPT	

$$1 + 1 + 1 + (2 \times 2 \times 2) \times 1 + (2 \times 2) \times 1 = 15$$

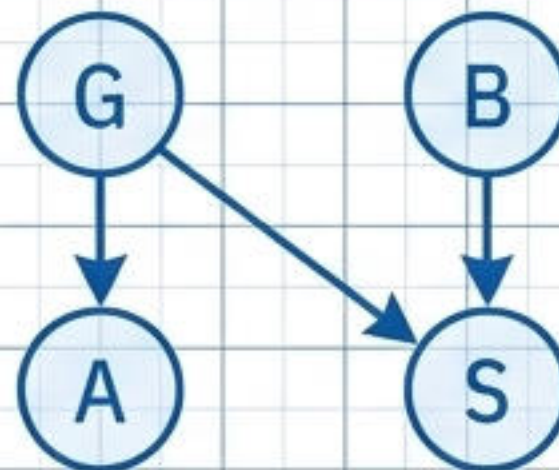
5 ערכים אפשריים לכל משתנה



$$4 + 1 + 1 + (5 \times 2 \times 2) \times 4 + 4 = 90$$

שימו לב: שינוי קל במספר המצבים (מ-2 ל-5 בצמתים A ו-E)
מקפיץ את גודל הרשת מ-15 ל-90 ערכים שיש לאחסן!

Genes (G) -> Disease (A) Network Structure



מהי ההסתברות של מטופל יש את מחלה A?
 $P(+a)$

Path 1 (Green)

$$P(+a|+g)P(+g) = (1.0)(0.1)$$

Path 2 (Orange)

$$P(+a|-g)P(-g) = (0.1)(0.9)$$

$$0.1 + 0.09 = 0.19$$

הסתברות שולית (Marginalization): כדי למצוא את A, אנו 'רצים' על כל המצבים האפשריים של ההורה שלו, G, ומסכמים את התוצאות.





מהי ההסתברות למחלה A בהינתן סימפטום S ומחלה B?

$$P(+a \mid +s, +b)$$



$$P(+a, +b, +s)$$

$$(0.19)(0.4)(1.0) = 0.076$$

$$\Rightarrow = 0.2267$$

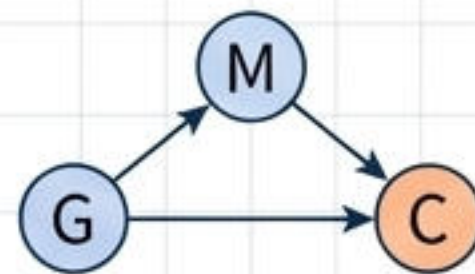
$$P(+a, +b, +s) + P(-a, +b, +s)$$

$$0.076 + [(0.81)(0.4)(0.8)] = 0.076 + 0.2592 = 0.3352$$



אזהרת חישוב מותנה: חובה לאזן! אסור לשכוח לחלק בהסתברות הכוללת של העדות (Evidence) המצויה במכנה.





Explaining Away (הסבר חלופי)

Scenario 1: (M) ידוע שיש לגליזציה
מה הסיכוי שהירוקים נבחרו (G)?

$$P(G|M) = 0.2286$$



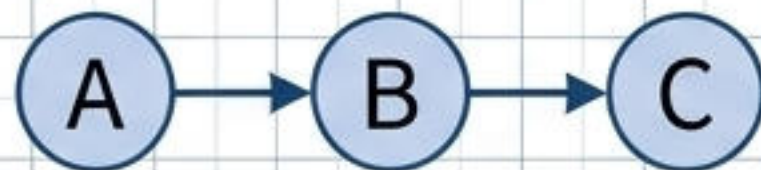
Scenario 2 - New Evidence!:
התגלה נתון נוסף: הנוכחות בשיעורים ירדה ($\sim C$).

$$P(G|\sim C) = 0.1163$$

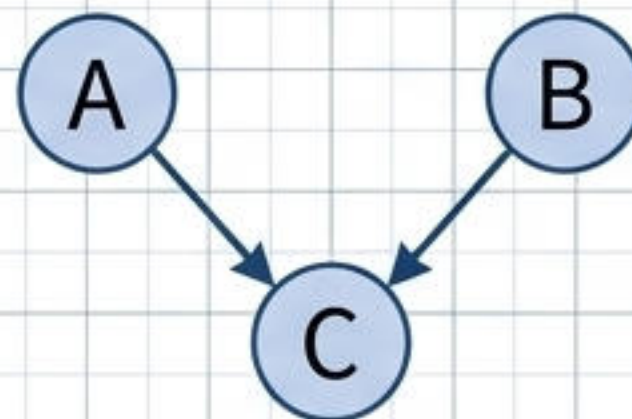


Explaining Away: ראייה חדשה (ירידה בנוכחות) מספקת הסבר חלופי לתופעה (לגליזציה), ולכן מורידה את הסבירות לגורם המקורי (ניצחון הירוקים).

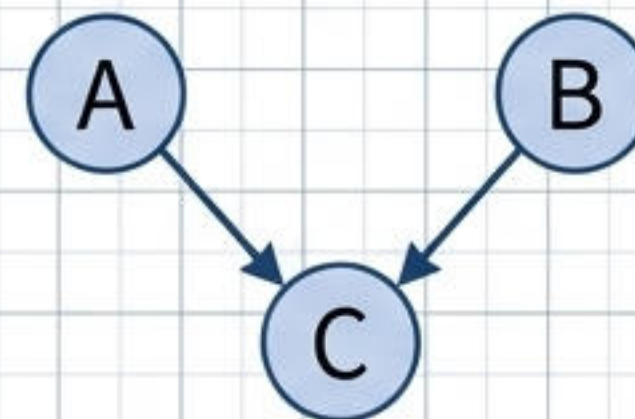
שרשרת



גורם משותף



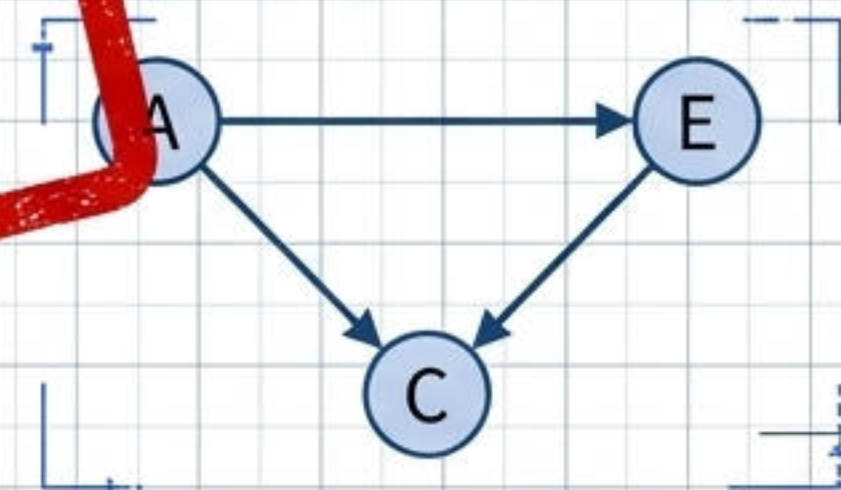
השפעה משותפת



האם מתקיים? $P(A, \neg C, E) = P(A)P(\neg C)P(E) - P(E)$

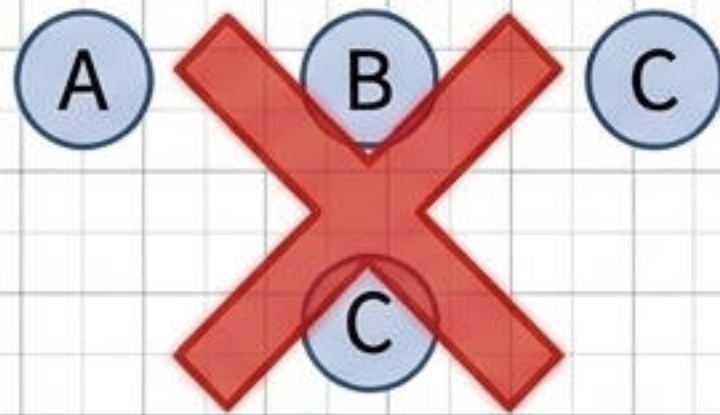
הטענה נכונה רק אם A, C, E בלתי תלויים לחלוטין זה בזה. המבנה הגרפי מראה מסלולים פתוחים, ולכן אי-התלות לא מתקיימת כאן.

לא נכון

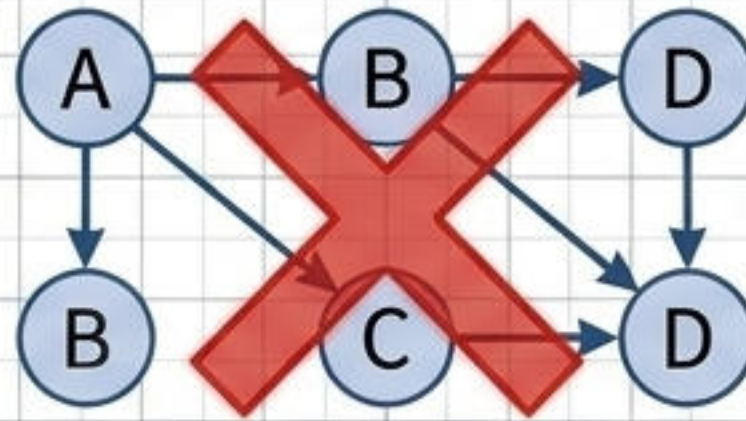


כלל ברזל: חפשו מסלולים פתוחים או חסומים (Evidence) כדי לקבוע תלויות, לפני שאתם צוללים לחישובים.

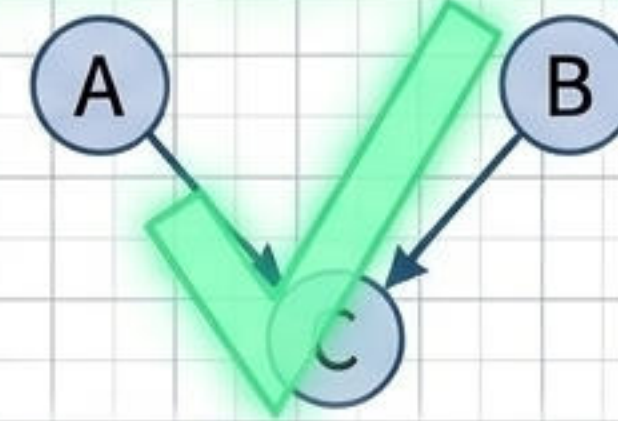
נתון מהטבלה: $P(A,B,C) = 0.25$
 אך מכפלת השוליים $P(A)P(B)P(C) = 0.175$. מסקנה: אין אי-תלות מוחלטת!



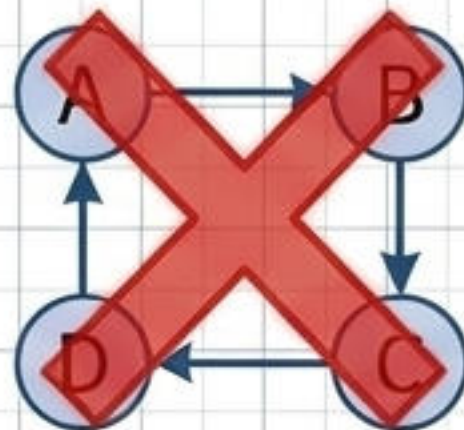
נפסל: מניח אי-תלות מוחלטת
 $P(A,B,C) = P(A)P(B)P(C)$



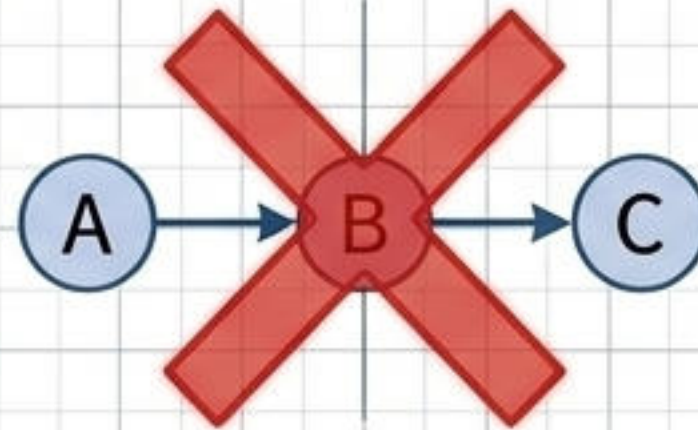
נפסל: מבנה מעגלי בעייתי



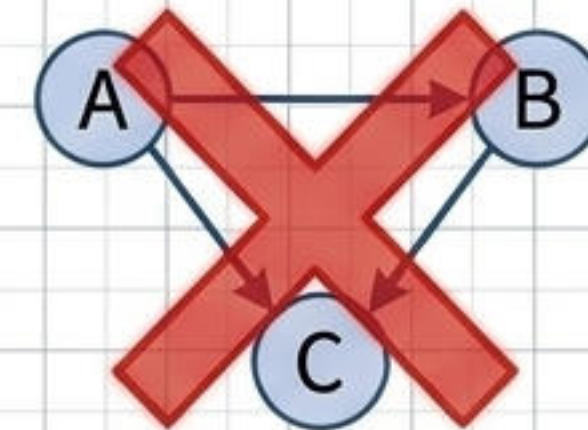
תואם: מבנה V מבטיח תלות
 החישוב תואם את הנתונים
 $P(A,B,C) = P(B|C)P(A)P(C)$



נפסל: מבנה מעגלי בעייתי



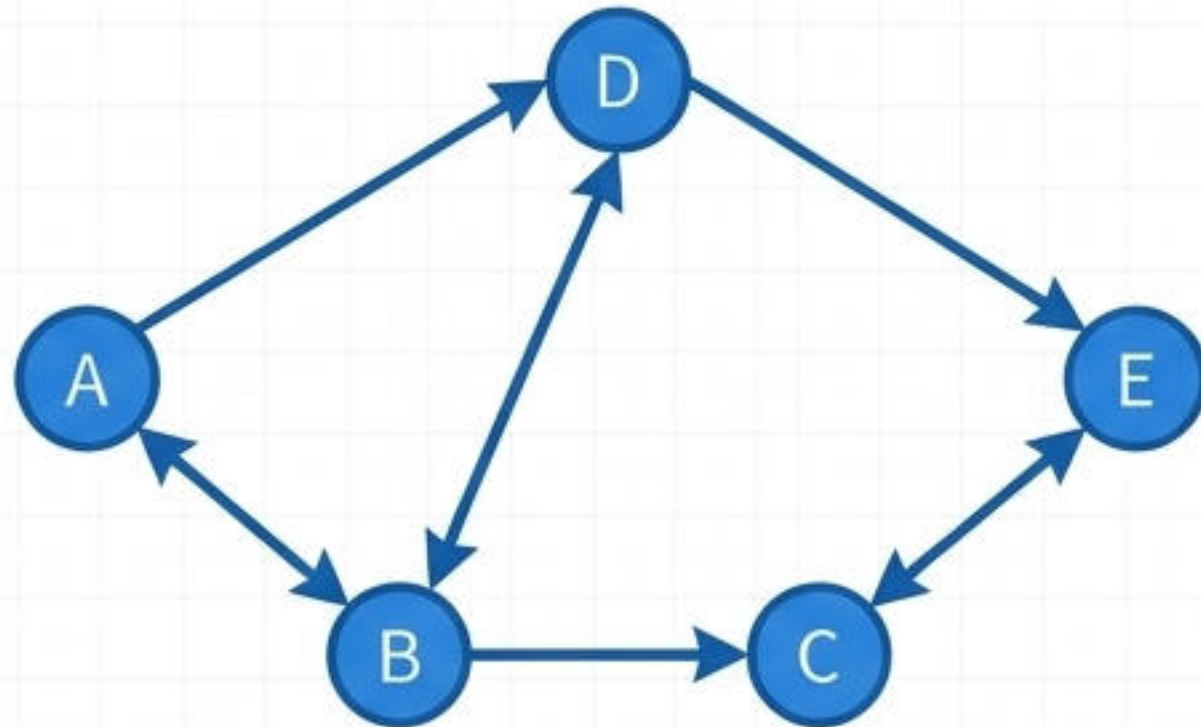
נפסל: מניח תלות מותנית
 A ו-C בהינתן B



נפסל: מניח תלות מותנית
 A ו-B בהינתן C

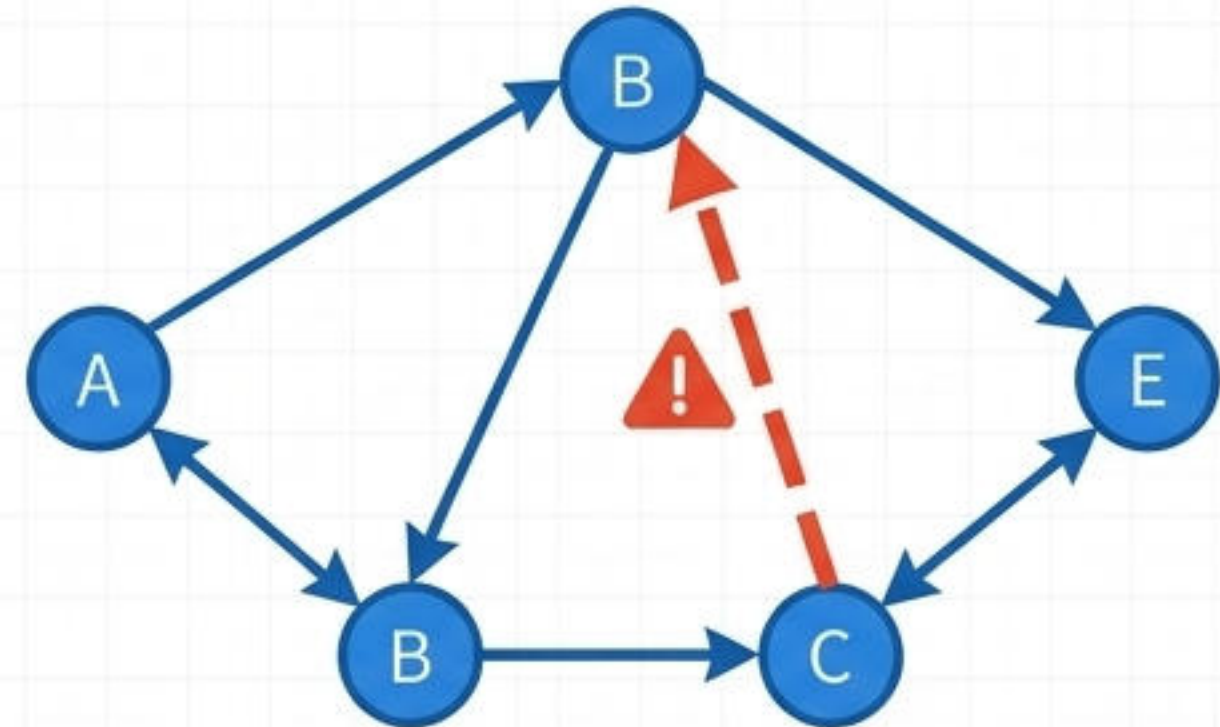
Structural Blueprint

סדר אופטימלי מקורי



סך גודל ה-CPTs: 10

סדר חדש: 4:E, 5:C, 3:B, 2:A, 1:D

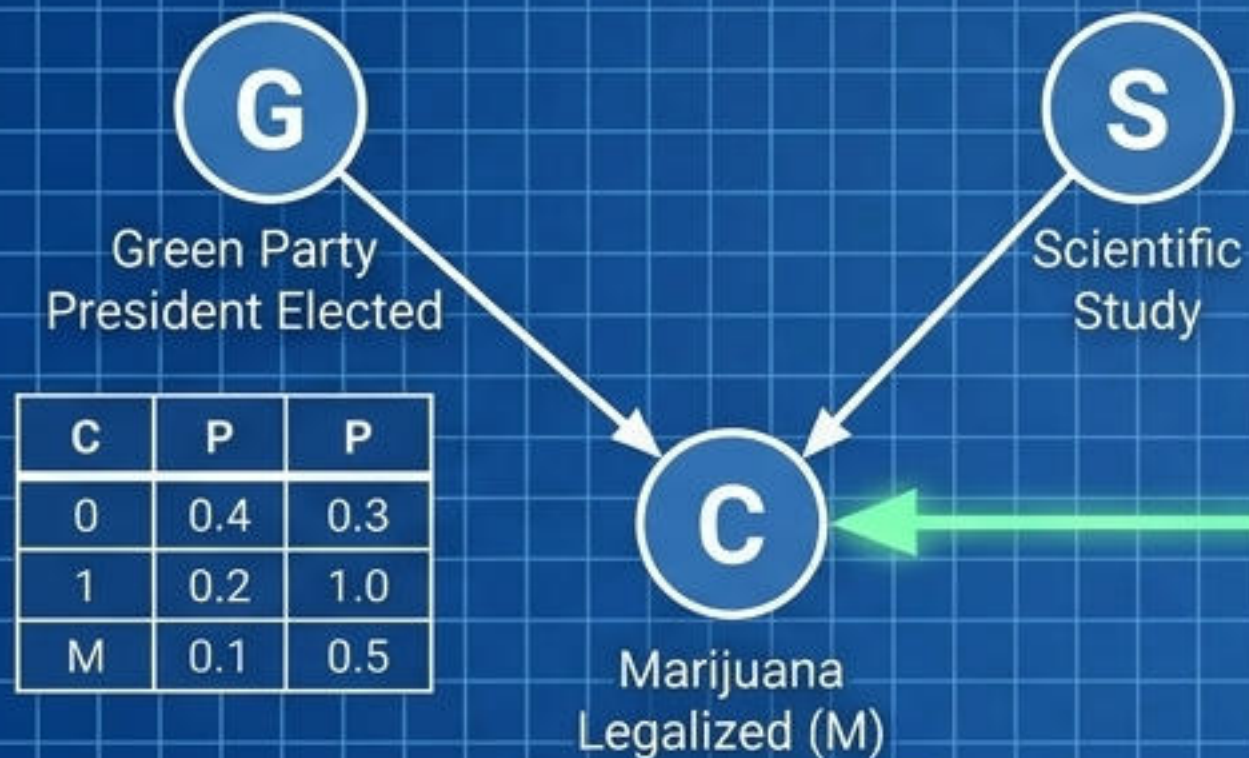


סך גודל ה-CPTs: 13

מדוע קפצנו ל-13? ברשת החדשה נאלצנו להוסיף קשת מיותרת בין C ל-B. מחיר ה הסדר הלא-טבעי הוא אובדן הנחות אי-תלות וגידול אקספוננציאלי בפרמטרים לאחסון.

Structural Blueprint

הוספת משתנה S (מחקר מדעי)
המשפיע על החלטת הלגיזציה (C)



C	P	P
0	0.4	0.3
1	0.2	1.0
M	0.1	0.5

M	P	P
0	0.3	0.4
0	0.2	0.1
1	0.2	0.5
M	0.1	0.3

☒ ~~$P(B,C) = P(B)*P(C)$~~

☒ ~~$P(C|G,M) = P(C|G)$~~

☒ **iii. $P(G) = P(G|S)$ - נכון**

מסלול חסום (D-Separated): מכיוון שאין לנו עדות על צומת האמצע (הלגיזציה), המסלול בין S-G חסום. לכן, מידע על המחקר לא משנה את סבירות ניצחון הירוקים.

Structural Blueprint



לוגיקת מבנה

תמיד חפשו D-Separation
לפני שמחשבים.
מסלול חסום = אי תלות מובטחת.
חוסך חישובים ארוכים!



כלכלת הרשת

סדר טופולוגי נכון שומר על
מינימום קשתות.
מינימום קשתות = מינימום
פרמטרים בטבלאות ה-CPT.



פירוק והרכבה

כשהכל נכשל, השתמשו בחוק
השרשרת לפרק ל-Full Joint,
ואל תשכחו את הכלל
בהסקה מותנית - לנרמל את
המכנה!

סוף כיתת האמן. אתם מוכנים.