Szita-formula

4. előadás, Nagy Zoltán Lóránt,

2022.10.07.

1. Szita-formula, motiváció és tétel

Ismétlés

Az alábbiakban a „Dobjuk ki a rosszat!” számolási alapötletet fogjuk általánosítani. Emlékeztetőül: egy H **halmaz elemszámát** $|H|$ **jelöli.**

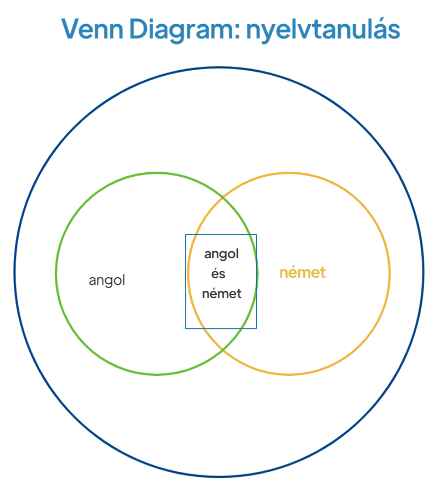
Motivációs feladat **1.** Hány olyan négy hosszú kockadobás-sorozat van, amiben van legalább egy hatos? Hány olyan négy hosszú kockadobás-sorozat van, amiben egy hatos sincs?

Bizonyítás. Megszámoljuk az összes dobássorozatot, és kivonjuk azok számát, amiben nincsen hatos. Ez utóbbi könnyebb: minden dobás 5-féle értékű lehet egymástól függetlenül, tehát itt a szorzási szabályt használva 5^4 hatost nem tartalmazó sorozatot kapunk. A hatost tartalmazó sorozatok száma tehát $6^4 - 5^4$.
Ezt megfogalmazhatjuk halmazokkal is:
• legyen H az összes négy hosszúságú kockadobás-sorozat halmaza, és
• legyen $A \subset H$ azon négy hosszú kockadobás-sorozatok halmaza, amelyekben nincs hatos.
Ekkor a jó dobássorozatok száma $|H \setminus A| = |H| - |A|$, ahol az előbbiek alapján $|H| = 6^4$, $|A| = 5^4$. □

Ezt általánosítjuk arra az esetre, ha több mint egy „rossz” halmaz elemeit akarjuk kidobni. Hétköznapi példa erre

Feladat **2.** Egy osztályban 37 gyerek tanul. Közülük angolul 27-en, németül 15-en, mindkét nyelven 10-en beszélnek. Hány olyan gyerek van, aki sem angolul, sem németül nem tud?

Bizonyítás. A szokásos megoldás, hogy kiszámoljuk a csak németül, illetve csak angolul beszélő gyerekek számát; ez 5, illetve 17. Ezután ezeket, valamint a mindkét nyelven beszélő gyerekek számát levonjuk az osztálylétszámból. Azaz az egyik nyelven sem beszélő gyerekek száma 37-5-17-10=5. Persze ezt az eredeti adatokkal is kifejezhettük volna ilyenformán: $37-(27-10)-(15-10)-10$; átrendezve $37-27-15+10=5$ adja a megoldást.
A megközelítést szemléltethetjük halmazábrán, az úgynevezett Venn-diagramon.(1 ábra). Ha most H az osztály tanulóinak halmaza, $A \subset H$ az angolul, $B \subset H$ a németül tudó tanulók halmaza, akkor mi a $H \setminus (A \cup B)$ halmaz méretére vagyunk kíváncsiak. Itt a „rossz” $A \cup B$ halmazt azon tanulók alkotják, akik *valamelyik* nyelvet beszélik. A fentiek alapján a $|H \setminus (A \cup B)| = |H| - |A| - |B| + |A \cap B|$ képlet adódik.
A képlet igazságát könnyű belátni: az osztály létszámából ($|H|$) kivontuk az angolul és a németül tudó tanulók számát ($|A|$, illetve $|B|$); de mivel ezzel kétszer vontuk le az angolul és németül is tudók számát ($|A \cap B|$) – pedig csak egyszer kellett volna –, korrigáljuk a többletet ezen szám hozzáadásával. □



1. Kérdés. *Miként változik a helyzet, ha* 3*, vagy még több nyelv szerepelne a feladatban?*

Azt vehetjük észre, hogy a fenti eljárás működni fog ilyenkor is, elegendő tudni a nyelvek minden részhalmazára vonatkozóan azt, hogy hányan tanulják, és ebből az ismeretből ki tudjuk számítgatni azt is, hogy hányan nem tanulnak egyáltalán nyelvet. A kiszámítási módot az alábbi általános tétel fogalmazza meg.

Az alábbiakban a H halmaz az alaphalmaz (ami az összes esetet, vizsgálandó objektumot tartalmazza), és az $A_1, A_2, \ldots, A_n \subset H$ részhalmazok a "rossz eseteket" tartalmazzák. Így a „jó esetek” halmaza $H \setminus (A_1 \cup \ldots \cup A_n)$.

1. Tétel (Logikai szitaformula)**.**

$$\begin{aligned}|H \setminus (A_1 \cup \ldots \cup A_n)| = |H| - (|A_1| + |A_2| + \ldots + |A_n|)\\ + (|A_1 \cap A_2| + \ldots + |A_{n-1} \cap A_n|)\\ - (|A_1 \cap A_2 \cap A_3| + \ldots + |A_{n-2} \cap A_{n-1} \cap A_n|)\\ \ldots\\ + (-1)^k(|A_1 \cap \ldots \cap A_k| + \ldots\ldots + |A_{n-k+1} \cap \ldots \cap A_n|) + \ldots\\ + (-1)^n|A_1 \cap \ldots \cap A_n|, \quad (1)\end{aligned}$$

ahol tehát a k-adik zárójeles kifejezésen az $A_1, \ldots, A_n$ halmazok összes *lehetséges k-as metszete pontosan egyszer szerepel. (Ilyen k-as metszetből $\binom{n}{k}$ darab van.) Szummás alakban fölírva:*

$$|H \setminus (A_1 \cup \ldots \cup A_n)| = |H| + \sum_{k=1}^{n}(-1)^k \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} |A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}|.$$

(Itt az első szummában a k futóindex azt jelenti, hogy hány halmazt metszünk össze, a $(-1)^k$ az alkalmas előjelet biztosítja, a második szummában pedig az összegzés az indexek összes lehetséges k-asára $(i_1, \ldots, i_k)$ megy.)

1. Következmény (A logikai szitaformula $n = 3$ "rossz esemény", "rossz esethalmaz" esetén)**.**

$$|H \setminus (A_1 \cup A_2 \cup A_3)| = |H| - (|A_1| + |A_2| + |A_3|)+$$

$$+(|A_1 \cap A_2| + |A_1 \cap A_3| + |A_2 \cap A_3|) - |A_1 \cap A_2 \cap A_3|$$

2. Következmény (A logikai szitaformula abban az esetben, ha bármely $k \geq 1$ különböző A_i halmaz metszete egyforma méretű, vagyis $|A_|| = |A_2| = \ldots = |A_n|$, továbbá $|A_1 \cap A_2| = |A_1 \cap A_3| = \ldots$, stb.)**.**

$$|H \setminus (A_1 \cup \ldots \cup A_n)| = |H| - \binom{n}{1}|A_1| + \binom{n}{2}|A_1 \cap A_2| - \ldots (-1)^n \binom{n}{n}|A_1 \cap A_2 \cap \ldots \cap A_n|$$

vagy szummás alakban:

$$|H \setminus (A_1 \cup \ldots \cup A_n)| = |H| + \sum_{k=1}^{n}(-1)^k \binom{n}{k}|A_1 \cap A_2 \cap \ldots \cap A_k|.$$

Világos, hogy a fenti formulából közvetlenül következik az is, ha az A_i halmazok uniójának komplementere helyett az A_i halmazok uniójának a méretét akarjuk meghatározni: hiszen

$$|H \setminus (A_1 \cup \ldots \cup A_n)| + |(A_1 \cup \ldots \cup A_n)| = |H|.$$

3. Következmény. *Tetszőleges $A_1, \ldots, A_n$ halmazokra*

$$|A_1 \cup \ldots \cup A_n| = |A_1| + \ldots + |A_n| - (|A_1 \cap A_2| + \ldots + |A_{n-1} \cap A_n|) + \ldots$$

$$+(-1)^{k+1}(|A_1 \cap \ldots \cap A_k| + \ldots + |A_{n-k+1} \cap \ldots \cap A_n|) + \ldots + (-1)^n|A_1 \cap \ldots \cap A_n|.$$

Szummás alakban ugyanez:

$$|A_1 \cup \ldots \cup A_n| = \sum_{k=1}^{n}(-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} |A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}|.$$

2. Szita-formula alkalmazások

2.1. 1. alkalmazás: a motivációs feladat újragondolása

Motivációs Feladat 1: Hány olyan négy hosszú kockadobás-sorozat van, amiben egy hatos sincs?

Ezt persze szorzási szabállyal meg tudtuk könnyen oldani, de természetes lehet a következő megközelítés is:
1. lépés: *Határozzuk meg az összes esetet (H halmaz) és a rossz esetek kategóriáit (A_i) !*
H: az összes dobássorozat.
A_i: az i. helyen hatos áll. $(i = 1, 2, 3, 4)$.
2. lépés: *Határozzuk meg az $|A_i|$, $|A_i \cap A_j|$, illetve általában többszörös metszetek elemszámait.*
Könnyű: $|H| = 6^4$. $|A_i| = 6^3$ $(i \in \{1, 2, 3, 4\})$.
$|A_i \cap A_j| = 6^2$ $(i < j$ és $i, j \in \{1, 2, 3, 4\})$
$|A_i \cap A_j \cap A_k| = 6^1$ $(i < j < k$ és $i, j, k \in \{1, 2, 3, 4\})$
$|A_1 \cap A_2 \cap A_3 \cap A_4| = 6^0 = 1$
Valóban: ha egyes helyeken rögzítjük a hatost, a többi helyen bármi állhat, akkor a megmaradt helyek száma száma a döntések száma, minden döntés megmarad 6-féle.
3. lépés: *Írjuk fel a szita-formulát.* Használhatjuk a 2. következmény szerinti alakot is most, mert a k-as metszetek mérete csak az összemetszett rossz halmazok k számától függ.

$$\#\text{ hatosmentes sorozatok } = 6^4 - \binom{4}{1}6^3 + \binom{4}{2}6^2 - \binom{4}{3}6^1 + \binom{4}{4}6^0.$$

Diszkusszió: Egy más alakú eredményt kaptunk, mint a fejezet elején. Honnan látszódik, hogy ugyanazt a számot jelölik? A binomiális tétel segít, $(x + y)^4$ kifejtése $x = 6, y = -1$ szereposztással!

A szita-formula alkalmazási receptje

- **1. lépés:** *Határozzuk meg az összes esetet (H halmaz) és a rossz esetek kategóriáit (A_i)*

  Ha jól dolgozunk, akkor a jó eseteket megkapjuk így: $H \setminus (A_1 \cup A_2 \cup \ldots \cup A_n)$.

- **2. lépés:** *Határozzuk meg az $|A_i|$, $|A_i \cap A_j|$, illetve általában a k-as metszetek elemszámait (vagyis amikor k halmazt metszünk össze).*

  Elképzelhető, hogy a k-as metszet mérete csak k-tól függ, a halmazok indexeitől nem. (Ekkor a 2. köv. egyszerűbb alakja alkalmazható.) De az is lehet, hogy nem ez a helyzet.

- **3. lépés:** *Írjuk fel a szita-formulát.*

  Ez meg is adja a jó esetek számát.

2.2. n-hez relatív prímek száma: az Euler féle φ függvény

1. Állítás. *Legyen az $n > 1$ pozitív egész szám prímtényezős felbontása $n = p_1^{\alpha_1} \cdot p_2^{\alpha_2} \cdot \ldots p_s^{\alpha_s}$. Ekkor az n-hez relatív prím n-nél kisebb számok számát megadja a*

$$\varphi(n) := n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\ldots\left(1 - \frac{1}{p_s}\right).$$

Bizonyítás. **1. lépés:** *Határozzuk meg az összes esetet (H halmaz) és a rossz esetek kategóriáit (A_i)*
Legyen H az 1 és n közti egész számok halmaza. Egy szám akkor rossz, ha nem relatív prím n-hez, azaz osztható valamelyik p_i-vel. Legyen A_1 az 1 és n közti, p_1-gyel osztható számok halmaza, A_2 az 1 és n közti, p_2-vel osztható számok halmaza, és így tovább, végül A_s az 1 és n közti, p_s-vel osztható számok halmaza.
Az 1 és n közti, n-hez relatív prím számok halmaza tehát $H \setminus (A_1 \cup \ldots \cup A_k)$.

2. lépés: *Határozzuk meg az $|A_i|$, $|A_i \cap A_j|$, illetve általában a k-as metszetek elemszámait*
A_i halmazban a p_i prímmel osztható számok szerepelnek, melyekből $\frac{n}{p_i}$ darab van.
Ha néhány halmazt összemetszünk, akkor a metszetben a halmazoknak megfelelő prímek mindegyikével osztható számok lesznek. Ezek éppen azok a számok, amik a prímek szorzatával oszhatóak.

3. lépés: *Írjuk fel a szita-formulát.* Eszerint a következőt kapjuk:

$$\phi(n) = n - \frac{n}{p_1} - \frac{n}{p_2} - \ldots - \frac{n}{p_s} + \frac{n}{p_1 p_2} + \ldots + \frac{n}{p_{s-1}p_s} - \ldots + (-1)^s \frac{n}{p_1 p_2 \ldots p_s},$$

ami az

$$n\left(1 - \frac{1}{p_1}\right)\left(1 - \frac{1}{p_2}\right)\ldots\left(1 - \frac{1}{p_s}\right)$$

alakba is írható (gondoljuk meg!) □

2.3. Karácsonyi húzás - fixpontmentes permutációk száma

Feladat **3.** Az osztályba n gyerek jár. Karácsonykor kihúzzák egymás nevét: mindenki pontosan egyvalakiét. Akkor sikeres egy húzás ha mindenkinél van cetli és senki sem saját magát húzta. Mennyi a lehetséges sikeres húzások száma?

1. Megjegyzés. *A húzásra gondolhatunk úgy, hogy az osztálynapló sorrendjében minden diáknak van egy sorszáma, és minden ilyen sorszámhoz azt a sorszámot rendeljük, akit húzott. Ez a számok egy permutációját adja. A fixpont jelentené azt, hogy valaki önmagát húzta.*

Bizonyítás. **1. lépés:** *Határozzuk meg az összes esetet (H halmaz) és a rossz esetek kategóriáit (A_i).*
Az összes kiosztások száma $|H| = n!$, ha a sikertelen húzásokat is számoljuk. Ez a szorzási szabályból következik: Az első gyerek n-féle cetlit húzhat, a következő már csak $n-1$-et, és egyesével fogynak a kihúzható cetlik minden egyéni húzás után. A rossz húzás jellemzője, hogy van olyan gyerek, aki magát húzta. Jelölje A_i az olyan cetlikiosztások eseteit (vagyis az esetek halmazát) amikor az i. gyerek önmagát húzta. Így annyiféle rossz eset fordul elő, ahányan vannak az osztályban. Jelölje H az összes kiosztások halmazát. Ezzel az 1. lépést elvégeztük. Valóban szükséges lesz szitálni, mivel könnyű látni, hogy vannak olyan kiosztások amik többféle szempontból is rosszak, legalább két A_i A_j halmazban is számolnánk őket, de mi minden rossz esetet csak egyszer akarunk az összes esetből levonni, vagyis $H \setminus (A_1 \cup \ldots A_n)$ méretére vagyunk kíváncsiak.

2. lépés: *Határozzuk meg az $|A_i|$, $|A_i \cap A_j|$, illetve általában a k-as metszetek elemszámait.*
$|A_i| = (n-1)!$, hiszen ha az i. diák a saját cetlijéthúzza, a többiek szabadon húzhatnak a maradék $n-1$ cetli közül. Hasonlóan bármelyik $A_i \cap A_j$ kettes metszet elemszáma $(n-2)!$, hiszen az i. és j. diák cetlijének kivétele után szabadon oszthatunk cetliket. Ezt általánosíthatjuk: ha rögzített k ember (pl. $1., 2., \ldots, k$) önmagát húzza, a többiekről pedig nincsen megkötésünk, akkor egy k-as metszet elemszámát kell megszámolnunk ($A_1 \cap A_2 \cap \ldots \cap A_k$), és erre a $(n-k)!$ eredmény adódik.

3. lépés: *Írjuk fel a szita-formulát.*
A szitaformula 3. következmény szerinti alakját használhatjuk, és eszerint a sikeres húzások keresett száma:

$$|H \setminus (A_1 \cup \ldots \cup A_n)| =$$
$$= n! - \binom{n}{1}(n-1)! + \binom{n}{2}(n-2)! - \ldots + (-1)^k \binom{n}{k}(n-k)! \pm \ldots + (-1)^n \binom{n}{n}.$$

□

2. Megjegyzés. *Ezzel a feladat megoldásaként kaptunk egy formulát, de még jó lenne értelmezni azt. Az eredményből kiemelve $n!$-t a következőt kapjuk n elem fixpontmentes permutációinak számára:*

$$n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \ldots + (-1)^n \frac{1}{n!}\right) = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

Analízisből majd tanuljuk, hogy a $\sum_{k=0}^{n} \frac{x^k}{k!}$ kifejezés értéke e^x-hez tart, amint n tart a végtelenhez (ahol $e \approx 2,71$ az ún. természetes logaritmus alapja, az Euler-féle szám). Így az eredményünkben kapott szorzót $x = -1$ helyettesítésel kapjuk és amely eszerint közelítőleg $e^{-1} = 1/e \approx 0,37$. Ez azt jelenti, hogy annak az esélye, hogy egy véletlenül választott permutáció fixpontmentes legyen, vagy a karácsonyi húzás sikeres, körülbelül 37%.

2.4. $X \to Y$ szürjektív függvények száma

2. Állítás. *Legyen a két halmazunk $X = \{x_1, \dots, x_n\}$ ill. $Y = \{y_1, \dots, y_m\}$. Ekkor az $X \to Y$ szürjektív leképezések száma*

$$\sum_{i=0}^{m-1} (-1)^i \binom{m}{i} (m-i)^n.$$

Bizonyítás. Először vegyük észre, hogy azok a rossz leképezések, amelyek esetén valamelyik y_j elemet egyik X-beli elemhez sem rendelték hozzá. Minden $j \in \{1, 2, \dots, m\}$-hoz van tehát rossz eseteknek egy-egy A_j halmaza, és ezek össze is metszenek.

Tehát az összes lehetséges $X \to Y$ leképezések száma $|H| = m^n$ (szorzási szabály, minden elem X-ben $|Y| = m$ helyettesítési értéket vehet fel a függvény szerint).
A_j elemszáma hasonlóképpen $(m-1)^n$, hiszen az egyik felvehető értéket kitiltottuk.
Ha i darab különböző indexű A halmazt metszünk össze, akkor annak mérete $(m-i)^n$, hiszen pontosan i felvehető értéket tiltottuk ki.
Innen a szitaformulával a képlet azonnal adódik, ha figyelmebe vesszük hogy i halmazt pontosan $\binom{m}{i}$ féle módon választhatunk m rossz $A_1, \dots A_m$ halmazunk közül. □

3. A szita-formula bizonyítása

1. bizonyítás: direkt. Nézzük meg, hogy H egyes elemeit hányszor veszi tekintetbe a bal illetve a jobb oldal. Ha az $x \in H$ elem nincs benne $A_1 \cup \dots \cup A_k$-ban, akkor a bal oldalon és a jobbon egyaránt 1-szer vesszük tekintetbe.
Ha x az A_i-k közül legalább egyben szerepel, akkor a baloldalon, $H \setminus (A_1 \cup A_2 \cup \dots \cup A_n)$-ben egyszer sem számoljuk. Lássuk a jobboldalt: Ha x az A_i-k közül pontosan $s > 0$ darabban van benne, – átindexelés után feltehető, hogy ezek $A_1, \dots, A_s$ – akkor a jobb oldalon ennyiszer számoltuk:

$$1 - s + \binom{s}{2} - \binom{s}{3} \pm \dots + (-1)^s \binom{s}{s} = \sum_{i=1}^{s} (-1)^i \binom{s}{i} \tag{2}$$

Valóban, ez esetben a jobboldalon az egyelemű metszetekben s-szer, a kételemű metszetekben $\binom{s}{2}$-szer, és általában a k-elemű metszetekben $\binom{s}{k}$-szor számoltuk.

Ez a kifejezés a binomiális tétel miatt egyenlő $(1-1)^s = 0$-val. Nekünk pedig pont ennyiszer kell számolnunk a bal oldalon az ilyen elemeket. □

2. bizonyítás: teljes indukció. A rossz esetek kategóriáinak n száma szerint

Kezdőlépés:

- $n = 1: \ |H \setminus A_1| = |H| - |A_1|$ ✓
- $n = 2: \ |H \setminus (A_1 \cup A_2)| = |H| - |A_1| - |A_2| + |A_1 \cap A_2|$ ✓

Indukciós lépés: Feltesszük, hogy $n \geq 2$-re igaz, és ekkor belátjuk, hogy $n+1$-re is teljesül.

Tehát tegyük fel, hogy teljesül ez:

$$|H \setminus (A_1 \cup \dots \cup A_n)| = |H| + \sum_{k=1}^{n} (-1)^k \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} |A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}|,$$

vagy Következmény 3. szerint ennek megfelelő

$$|A_1 \cup \dots \cup A_n| = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} |A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}|. \tag{3}$$

Jelöljük el B-vel ezt a halmazt: $B := (A_1 \cup \dots \cup A_n)$.
A bizonyítandó bal oldala így írható: $|H \setminus (B \cup A_{n+1})|$. Ez éppen olyan alakú, mint amit a kéttagú szitaformulára már láttunk! Vagyis:

$$|H \setminus (B \cup A_{n+1})| = |H| - |B| - |A_{n+1}| + |B \cap A_{n+1}| \tag{4}$$

Vessük ezt össze a bizonyítandó jobb oldalával!

$$|H| + \sum_{k=1}^{n+1} (-1)^k \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n+1} |A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}|. \tag{5}$$

$|H|$ jelenléte stimmel. $-|B| = -|A_1 \cup \ldots \cup A_n|$, ami épp olyan alakú az indukciós feltétel, vagyis (3) egyenlőség alapján, mint amit (5)-ben látunk, de az indexek úgy futnak, hogy az összes olyan halmaz metszeteit gyűjti össze ez a tag előjelezve, ami NINCS összemetszve A_{n+1}-gyel.
Mi a helyzet az A_{n+1}-gyel összemetszett tagokkal? Egyrészt $|A_{n+1}|$ megfelelő előjellel megvan nekünk (4)-ben, eddig jó. Másrészt azt kell észrevenni, hogy $|B \cap A_{n+1}|$ éppen a hiányzó tagokat gyűjti a megfelelő előjellel.
Valóban,

$$B \cap A_{n+1} = (A_1 \cap A_{n+1}) \cup (A_2 \cap A_{n+1}) \cup \ldots \cup (A_n \cap A_{n+1}).$$

Így amikor $B \cap A_{n+1}$ számosságát akarjuk becsülni, használhatjuk az indukciós feltételt, (3) egyenlőséget, ám most a következő n halmazra:

$$A_1^* = (A_1 \cap A_{n+1}), A_2^* = (A_2 \cap A_{n+1}), \ldots, A_n^* := (A_n \cap A_{n+1})$$

Az így definiált halmazok metszethalmazai épp azok a tagok lesznek a bizonyítandó (5) kifejezésben amiben A_{n+1} szerepel a metsző hamazok között, és az előjelezés miatt a szükséges előjellel jelennek meg. Ezzel a bizonyítást befejeztük. □

Az $n = 3$ esetre lépést $n = 2$ről az ÁBRA mutatja be.