O wartości informacji

Prof. dr inż. Juliusz Lech KULIKOWSKI

Komu kiedyś zdarzyło się spędzić godzinę lub więcej w kolejce po gazetę, ten miał okazję zastanowić się, co decyduje o wartości informacji. Czy nominalna cena informacji (w tym przypadku — 5 zł) jest równoważna jej wartości? Wśród podstawowych atrybutów informacji, do jakich zalicza się m.in. jej miarę ilościową, treść, jakość itp., wartość informacji jest atrybutem stosunkowo mało poznanym. Owszem, podejmowano w przeszłości próby wyjaśnienia istoty wartości informacji (M. M. Bongard, A. A. Charkiewicz, N. Stanułow i in.), lecz nie wyszły one poza poziom ogólnych koncepcji. Koncepcja, którą niżej przedstawiam Czytelnikom, po raz pierwszy była zreferowana na IV kongresie IFAC w Warszawie w 1969 r. oraz w bardziej rozwiniętej postaci — na drugim Międzynarodowym Sympozjum Teorii Informacji w Cachkadzor (Armeńska SRR w 1970 r.). Nie zamierzam ukrywać, że spotkała się ona z przyjęciem życzliwym, ale także ograniczonym do wąskiego grona zainteresowanych osób. Mimo to problem, jak zobiektywizować pojęcie wartości informacji, uważam za wciąż aktualny, dlatego ośmielam się zaprzątnąć nim także uwagę Czytelników.

Ekonomia polityczna określa wartość towaru jako równoważnik ilości pracy społecznie niezbędnej do tego, by towar wyprodukować. Czy odnosi się to także do wartości informacji? Jeśli nawet informacja jest przedmiotem wymiany handlowej, to mimo wszystko jest to towar szczególny. Konsumpcja informacji nie niszczy jej, tak jak na przykład konsumpcja artykułów spożywczych, lecz raczej — powiela informację i przedłuża jej istnienie. Ponadto, wartość informacji nie jest związana żadną prostą zależnością funkcyjną z jej miarą ilościową. Proste buchalteryjne metody wyliczania wartości informacji nie na wiele zatem się przydadzą. Ocenę wartości informacji powinniśmy wiązać raczej z jej wartością użytkową, ocenianą przez odbiorcę informacji. Należy jednak zastrzec się, że nie idzie tu o wartość rozumianą bardzo wąsko, w kategoriach korzyści czysto materialnych. Na przykład, jeśli kiedykolwiek nastąpi potwierdzenie hipotezy o istnieniu cywilizacji pozaziemskich, to z punktu widzenia statystycznej teorii komunikacji przyniesie nam ono zaledwie 1 bit informacji (przyjmując, że obecnie hipotezy istnienia i braku takich cywilizacji trzeba uznać za jednakowo prawdopodobne). Utylitarne znaczenie takiej informacji przypuszczalnie będzie również niewielkie (wykluczam tu możliwość wylądowania Kosmitów na Ziemi), natomiast wartość użytkowa informacji o cywilizacji pozaziemskiej rozumiana szeroko, z uwzględnieniem jej aspektów przyrodniczych i światopoglądowych byłaby ogromna. Każda sensowna teoria wartości informacji musi liczyć się z jej względnością i subiektywnością ocen. Mimo to pewne czynniki mają niewątpliwy i bezpośredni związek z wartością informacji dla każdego, kto tę wartość ocenia. Można do nich przykładowo zaliczyć:

— treść informacji,
— aktualność informacji,
— jej wiarygodność,
— jej zrozumiałość,

itp.. Dochodzimy zatem do ważnego spostrzeżenia, że wartość informacji ,,nie jedno ma imię'', lecz wyraża szereg różnych jej cech jakościowych, które można wprawdzie poddawać kwantyfikacji, lecz trudno sprowadzić je do jednego parametru liczbowego (skalara).

Z drugiej strony natomiast, jeśli mówimy o ,,wartości'' czegokolwiek, to na ogół milcząco zakładamy, że:

— wartości można dodawać,
— wielokrotne dodawanie wartości można zastąpić przez mnożenie ich przez liczby naturalne, a ogólniej— rzeczywiste,
— wartości można porównywać.

Spostrzeżenia te naprowadzają nas na pewien teoretyczny model wartości informacji, który niżej opiszemy. Wartość informacji w tym modelu jest wektorem, którego poszczególne składowe odpowiadają czynnikom decydującym w opinii użytkownika o wartości informacji. Dodawanie wektorów i mnożenie ich przez liczby rzeczywiste możliwe jest, jeżeli uważamy wektory za elementy przestrzeni liniowej. Aksjomaty takiej przestrzeni zakładają bowiem, że jeśli x, y są dwoma jej dowolnymi elementami (wektorami), a, b zaś są dwiema dowolnymi liczbami rzeczywistymi, to określone jest zarówno dodawanie elementów oznaczane jako $x \oplus y$, jak i mnożenie ich przez liczbę rzeczywistą (oznaczane jako $a \cdot x$), przy czym zachodzą następujące tożsamości:

1. $a \cdot (x \oplus y) = a \cdot x \oplus a \cdot y = a \cdot y \oplus a \cdot x$,
2. $(a+b) \cdot x = a \cdot x \oplus b \cdot x$,
3. $a \cdot (b \cdot x) = (ab) \cdot x, \quad 1 \cdot x = x$.

W przestrzeni liniowej wyróżniony jest ponadto pewien element $\mathbf{0}$, zwany elementem zerowym, taki, że dla dowolnego x

4. $\mathbf{0} \cdot x = \mathbf{0}$

Reguły te są spełnione przez przestrzenie wektorów o ustalonej liczbie składowych rzeczywistych; wektory takie można by zatem uznać za reprezentantów wartości informacji, gdybyśmy nadali właściwą interpretację ich składowym oraz gdybyśmy wyjaśnili, na czym polega możliwość porównywania wektorów. Należy jednocześnie zauważyć, że podobnie jak sumowanie probabilistycznych miar ilości informacji może być stosowane tylko w odniesieniu do komunikatów parami statystycznie niezależnych, tak też dodawanie wektorowych bądź skalarnych miar wartości informacji może być stosowane tylko w odniesieniu do komunikatów niezależnych treściowo lub statystycznie. Powtórzenie tego samego komunikatu i skierowanie go do tego samego odbiorcy w jego odczuciu nie zwiększa zatem wartości użytkowej zawartej w informacji, chyba że przyczynia się do podniesienia wiarygodności informacji, lecz i w tym przypadku wzrost wartości informacji nie będzie proporcjonalny do liczby powtórzeń. Dalsze rozważania ograniczymy zatem do komunikatów treściowo i statystycznie niezależnych, nie precyzując zresztą tych pojęć, gdyż dla naszych celów wystarczy pojmować je w sposób zgodny z intuicją. Przytoczymy przykład, który ułatwi nam dalsze rozważania. Przypuśćmy, że interesuje nas pewien fakt historyczny, o którym chcemy zasięgnąć bliższej informacji. Możemy ten cel osiągnąć dwojako: 1° pytając kolegów lub domowników o interesujące nas szczegóły, 2° wypożyczając w czytelni książkę historyczną na odpowiedni temat. Pierwsze rozwiązanie gwarantuje nam szybki dostęp do informacji, niestety, mało wiarygodnej. Drugie rozwiązanie, odwrotnie, dostarczy nam wiarygodnej informacji, lecz ze znacznym opóźnieniem. Które rozwiązanie powinniśmy wybrać? Zapytajmy inaczej: w którym przypadku wartość użytkowa informacji będzie większa?

Zaczniemy od sparametryzowania interesujących nas cech informacji: jej wiarygodności i aktualności. Możemy na przykład założyć, że jeśli p oznacza prawdopodobieństwo tego, iż odpowiedź na określoną kategorię zapytań otrzymana z pewnego źródła będzie nieprawidłowa, to *współczynnik wiarygodności* k_w informacji uzyskiwanych z tego źródła na dany temat będzie wyliczany z wzoru

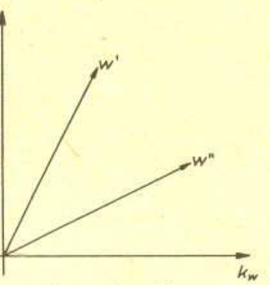
$$k_w = -\log p.$$

‹ określony współczynnik wiarygodności przyjmuje wartość
vną 0, jeśli informacje z danego źródła i na dany temat są
eguły nieprawdziwe, natomiast dąży do nieskończoności,
dy prawdopodobieństwo uzyskania nieprawdziwej odpowiedzi
y do zera. Można też powiedzieć, że tak określony
ółczynnik wiarygodności wyraża ilość informacji zawartej
twierdzeniu, że odpowiedź otrzymana z danego źródła
dany temat jest nieprawdziwa.

półczynnik aktualności informacji k_a możemy natomiast
eślić przy pomocy wzoru

$$k_a = \frac{t_o}{t},$$

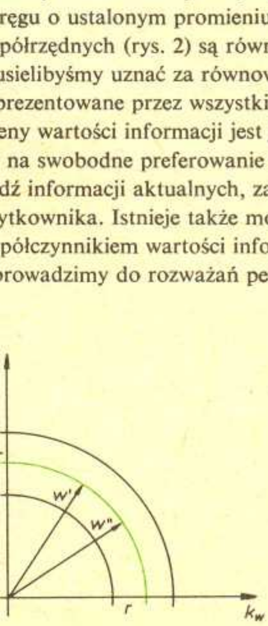
którym t_o oznacza pewien normatywny czas opóźnienia danej
ormacji (tj. upływ czasu między zadaniem pytania i
skaniem odpowiedzi), zaś t jest faktycznym czasem
óźnienia tejże informacji. Współczynnik aktualności jest równy
u, jeśli opóźnienie informacji równe jest nieskończoności,
równy jedności, jeśli opóźnienie to jest równe normatywnemu
sowi opóźnienia, które z założenia powinno być większe od
a i jest równy nieskończoności, jeśli opóźnienie jest równe
o.



Rys. 1

rys. 1 przedstawiono parę wektorów, których składowe
vne są wartościom współczynników k_w, k_a; wektor w'
powiada informacji mało wiarygodnej, lecz natychmiastowej,
ktor w'' natomiast odpowiada informacji wysoce wiarygodnej,
z otrzymanej ze znacznym opóźnieniem. Ponawiamy nasze
tanie: w którym przypadku wartość informacji jest większa,
yli który z dwóch wektorów jest „większy"?

miarę „wielkości" wektora można, oczywiście, przyjąć jego
igość. Oznacza to, iż wszystkie wektory wychodzące
oczątku układu współrzędnych, których końce znajdują się na
ręgu o ustalonym promieniu r zakreślonym z początku układu
półrzędnych (rys. 2) są równoważne. W konsekwencji,
isielibyśmy uznać za równowartościowe informacje
orezentowane przez wszystkie takie wektory. Taki sposób
eny wartości informacji jest jednak zbyt sztywny, nie pozwala
na swobodne preferowanie bądź informacji wiarygodnych,
dź informacji aktualnych, zależnie od woli i potrzeby
ytkownika. Istnieje także możliwość operowania „ważonym"
półczynnikiem wartości informacji. Jeśli mianowicie
rowadzimy do rozważań pewien współczynnik liczbowy c



Rys. 2

taki, że $0 \leqslant c \leqslant 1$ i określimy „ważony" współczynnik wartości informacji:

$$k = c \cdot k_w + (1-c) \cdot k_a,$$

to współczynnik k może posłużyć do „uporządkowania" wektorów (k_w, k_a). Wektorami wzajemnie równoważnymi stają się w tym przypadku takie wektory, którym odpowiada ta sama wartość współczynnika k, a więc wektory, których końce znajdują się na prostych o równaniach:

$$k_a = \frac{c}{c-1} k_w + \frac{k}{1-c}.$$

Współczynnik liczbowy c wyraża w tym przypadku pewien arbitralnie przyjęty stopień preferencji przyznawanej wiarygodności informacji w stosunku do jej aktualności. Jak łatwo zauważyć, jeżeli przyjmiemy $c = 1$, to aktualność informacji w ogóle nie będzie brana pod uwagę jako czynnik decydujący o jej wartości użytkowej. Odwrotnie, przyjęcie, że $c = 0$ oznacza, że nie przypisuje się znaczenia wiarygodności informacji, byle była ona dostarczana szybko.

Również i ten sposób porównywania wartości informacji, traktowanych wektorowo nie jest wystarczająco elastyczny, choć dodatkowy parametr c zwiększa jego elastyczność w stosunku do poprzedniego sposobu.

Jako kolejny wariant modelu formalnego wartości informacji rozważymy zatem tzw. *liniową przestrzeń półuporządkowaną*. Idea takiej przestrzeni została wysunięta jeszcze w latach trzydziestych przez L. B. Kantorowicza, a następnie została rozwinięta przez szkołę matematyków leningradzkich.

Liniowa przestrzeń półuporządkowana, którą będziemy krótko nazywali przestrzenią typu K, jest dowolną przestrzenią liniową, w której została określona relacja „dodatniości" jej elementów, oznaczana jak następuje:

$$x \succ \boldsymbol{0}$$

(czyt. „x jest większe od $\boldsymbol{0}$"), przy czym $\boldsymbol{0}$ jest elementem zerowym przestrzeni liniowej, na przykład wektorem o zerowych składowych. Relacja „dodatniości" powinna przy tym spełniać następujące założenia:

1° jeśli $x \succ \boldsymbol{0}$, to nie zachodzi $x = \boldsymbol{0}$,

2° jeśli $x \succ \boldsymbol{0}$ i $y \succ \boldsymbol{0}$, to $x \oplus y \succ \boldsymbol{0}$,

3° dla każdego elementu x w przestrzeni typu K istnieje taki element $x' \succ \boldsymbol{0}$, że $x' \oplus (-1) \cdot x \succ \boldsymbol{0}$, czyli krócej taki, że $x' \succ x$ (czyt. „x' jest większe od x"),

4° jeśli $x \succ \boldsymbol{0}$ i $a > 0$, to $a \cdot x \succ \boldsymbol{0}$.

Dla wyżej podanych założeń dochodzi jeszcze jedno, dotyczące podzbiorów ograniczonych w przestrzeni typu K. Zbiór S elementów tej przestrzeni nazywamy zbiorem ograniczonym z góry, jeśli istnieje taki element ξ w tejże przestrzeni, że ξ jest większy od każdego elementu zbioru S; element ξ nazywamy wówczas ograniczeniem górnym zbioru S. Jeśli ponadto ξ ma tę własność, że nie jest on większy od żadnego innego elementu przestrzeni będącego także ograniczeniem górnym zbioru S, to nazywamy go kresem górnym zbioru S i oznaczamy przez sup S. W przestrzeni typu K obowiązuje założenie, że

5° każdy podzbiór S przestrzeni typu K ograniczony z góry, ma w niej kres górny sup S.

Założenie to odgrywa istotną rolę szczególnie przy rozważaniu zagadnień zbieżności ciągów elementów w przestrzeni typu K.

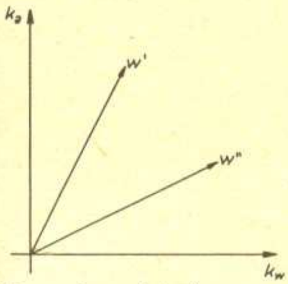
Powróćmy do przykładu przestrzeni wektorów dwuwymiarowych wyrażających wartość informacji poprzez wskaźniki jej wiarygodności i aktualności. Za wektory dodatnie w sensie aksjomatyki przestrzeni typu K możemy przyjąć wszystkie takie

Tak określony współczynnik wiarygodności przyjmuje wartość równą 0, jeśli informacje z danego źródła i na dany temat są z reguły nieprawdziwe, natomiast dąży do nieskończoności, kiedy prawdopodobieństwo uzyskania nieprawdziwej odpowiedzi dąży do zera. Można też powiedzieć, że tak określony współczynnik wiarygodności wyraża ilość informacji zawartej w stwierdzeniu, że odpowiedź otrzymana z danego źródła na dany temat jest nieprawdziwa.

Współczynnik aktualności informacji k_a możemy natomiast określić przy pomocy wzoru

$$k_a = \frac{t_o}{t},$$

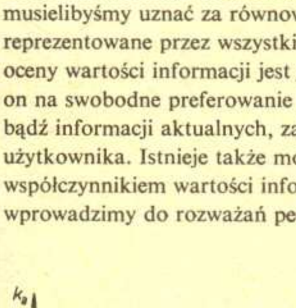
w którym t_o oznacza pewien normatywny czas opóźnienia danej informacji (tj. upływ czasu między zadaniem pytania i uzyskaniem odpowiedzi), zaś t jest faktycznym czasem opóźnienia tejże informacji. Współczynnik aktualności jest równy zeru, jeśli opóźnienie informacji równe jest nieskończoności, jest równy jedności, jeśli opóźnienie to jest równe normatywnemu czasowi opóźnienia, które z założenia powinno być większe od zera i jest równy nieskończoności, jeśli opóźnienie jest równe zero.



Rys. 1

Na rys. 1 przedstawiono parę wektorów, których składowe równe są wartościom współczynników k_w, k_a; wektor w' odpowiada informacji mało wiarygodnej, lecz natychmiastowej, wektor w'' natomiast odpowiada informacji wysoce wiarygodnej, lecz otrzymanej ze znacznym opóźnieniem. Ponawiamy nasze pytanie: w którym przypadku wartość informacji jest większa, czyli który z dwóch wektorów jest ,,większy"?

Za miarę ,,wielkości" wektora można, oczywiście, przyjąć jego długość. Oznacza to, iż wszystkie wektory wychodzące z początku układu współrzędnych, których końce znajdują się na okręgu o ustalonym promieniu r zakreślonym z początku układu współrzędnych (rys. 2) są równoważne. W konsekwencji, musielibyśmy uznać za równowartościowe informacje reprezentowane przez wszystkie takie wektory. Taki sposób oceny wartości informacji jest jednak zbyt sztywny, nie pozwala on na swobodne preferowanie bądź informacji wiarygodnych, bądź informacji aktualnych, zależnie od woli i potrzeby użytkownika. Istnieje także możliwość operowania ,,ważonym" współczynnikiem wartości informacji. Jeśli mianowicie wprowadzimy do rozważań pewien współczynnik liczbowy c



Rys. 2

taki, że $0 \leqslant c \leqslant 1$ i określimy ,,ważony" współczynnik wartości informacji:

$$k = c \cdot k_w + (1-c) \cdot k_a,$$

to współczynnik k może posłużyć do ,,uporządkowania" wektorów (k_w, k_a). Wektorami wzajemnie równoważnymi stają się w tym przypadku takie wektory, którym odpowiada ta sama wartość współczynnika k, a więc wektory, których końce znajdują się na prostych o równaniach:

$$k_a = \frac{c}{c-1} k_w + \frac{k}{1-c}.$$

Współczynnik liczbowy c wyraża w tym przypadku pewien arbitralnie przyjęty stopień preferencji przyznawanej wiarygodności informacji w stosunku do jej aktualności. Jak łatwo zauważyć, jeżeli przyjmiemy $c = 1$, to aktualność informacji w ogóle nie będzie brana pod uwagę jako czynnik decydujący o jej wartości użytkowej. Odwrotnie, przyjęcie, że $c = 0$ oznacza, że nie przypisuje się znaczenia wiarygodności informacji, byle była ona dostarczana szybko.

Również i ten sposób porównywania wartości informacji, traktowanych wektorowo nie jest wystarczająco elastyczny, choć dodatkowy parametr c zwiększa jego elastyczność w stosunku do poprzedniego sposobu.

Jako kolejny wariant modelu formalnego wartości informacji rozważymy zatem tzw. *liniową przestrzeń półuporządkowaną*. Idea takiej przestrzeni została wysunięta jeszcze w latach trzydziestych przez L. B. Kantorowicza, a następnie została rozwinięta przez szkołę matematyków leningradzkich.

Liniowa przestrzeń półuporządkowana, którą będziemy krótko nazywali przestrzenią typu K, jest dowolną przestrzenią liniową, w której została określona relacja ,,dodatniości" jej elementów, oznaczana jak następuje:

$$x \succ \mathbf{0}$$

(czyt. ,,x jest większe od $\mathbf{0}$"), przy czym $\mathbf{0}$ jest elementem zerowym przestrzeni liniowej, na przykład wektorem o zerowych składowych. Relacja ,,dodatniości" powinna przy tym spełniać następujące założenia:

1° jeśli $x \succ \mathbf{0}$, to nie zachodzi $x = \mathbf{0}$,
2° jeśli $x \succ \mathbf{0}$ i $y \succ \mathbf{0}$, to $x \oplus y \succ \mathbf{0}$,
3° dla każdego elementu x w przestrzeni typu K istnieje taki element $x' \succ \mathbf{0}$, że $x' \oplus (-1) \cdot x \succ \mathbf{0}$, czyli krócej taki, że $x' \succ x$ (czyt. ,,x' jest większe od x"),
4° jeśli $x \succ \mathbf{0}$ i $a > 0$, to $a \cdot x \succ \mathbf{0}$.

Dla wyżej podanych założeń dochodzi jeszcze jedno, dotyczące podzbiorów ograniczonych w przestrzeni typu K. Zbiór S elementów tej przestrzeni nazywamy zbiorem ograniczonym z góry, jeśli istnieje taki element ξ w tejże przestrzeni, że ξ jest większy od każdego elementu zbioru S; element ξ nazywamy wówczas ograniczeniem górnym zbioru S. Jeśli ponadto ξ ma tę własność, że nie jest on większy od żadnego innego elementu przestrzeni będącego także ograniczeniem górnym zbioru S, to nazywamy go kresem górnym zbioru S i oznaczamy przez sup S. W przestrzeni typu K obowiązuje założenie, że

5° każdy podzbiór S przestrzeni typu K ograniczony z góry, ma w niej kres górny sup S.

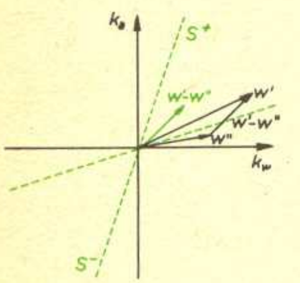
Założenie to odgrywa istotną rolę szczególnie przy rozważaniu zagadnień zbieżności ciągów elementów w przestrzeni typu K.

Powróćmy do przykładu przestrzeni wektorów dwuwymiarowych wyrażających wartość informacji poprzez wskaźniki jej wiarygodności i aktualności. Za wektory dodatnie w sensie aksjomatyki przestrzeni typu K możemy przyjąć wszystkie takie wektory, które sprowadzone do początku układu współrzędnych mieszczą się wewnątrz pewnego arbitralnie wybranego sektora kątowego S^+ (zob. rys. 3). O dwóch wektorach w', w'' można wówczas orzec, że na przykład $w' \succ w''$ wtedy i tylko wtedy, gdy różnica wektorowa $w' - w''$ jest wektorem dodatnim, należącym do sektora S^+ (rys. 4). Jeśli różnica ta jest natomiast wektorem należącym do sektora S^- wektorów „ujemnych" będących symetrycznym względem początku układu O odbiciem wektorów dodatnich tworzących sektor S^+, to wektor w'' uznajemy za większy od wektora w'. Mamy tu zatem względnie prostą zasadę porównywania wzajemnego wektorów wyrażających wartość informacji. Pozostaje jednak wątpliwość, jak należy traktować takie pary wektorów, których różnica nie należy ani do sektora S^+, ani do sektora S^-, jak to ilustruje rys. 5. O takich parach wektorów trzeba orzec, że są wzajemnie „nieporównywalne".

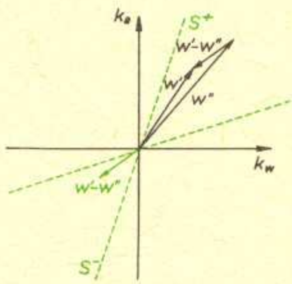
Konsekwencją tego jest to, że niektóre pary komunikatów trzeba uznać za nieporównywalne z punktu widzenia wartości użytkowej zawartych w nich informacji. Na pierwszy rzut oka może wydać się, że istnienie par wektorów wzajemnie nieporównywalnych jest istotnym niedostatkiem naszego modelu wartości informacji. Tak jednak nie jest, gdyż także w życiu walory użytkowe niektórych informacji trzeba uznać za nieporównywalne. Model formalny nie musi rościć sobie pretensji, by poprawiać rzeczywistość, wystarczy, że ją wiernie opisuje i pozwala na tej podstawie wyciągać wnioski.

Dobierając stosownie do potrzeby rozwartość sektora S^+ możemy w znacznych granicach regulować stopień nieporównywalności wektorów w przestrzeni typu K, w czym przejawia się elastyczność modelu, tak istotna w jego praktycznych zastosowaniach.

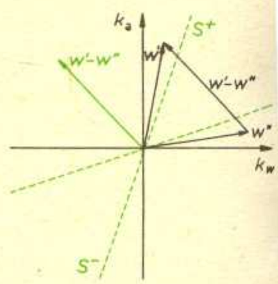
Przestrzeń typu K może być przestrzenią o dowolnie dużej, skończonej liczbie wymiarów (może to być także przestrzeń nieskończenie wielowymiarowa, lecz ten przypadek nie ma dla nas praktycznego znaczenia). Sektor S^+ wektorów dodatnich przybiera w ogólności postać wielowymiarowego „stożka" wychodzącego z punktu O (jednakże nie zawierającego tego punktu, zgodnie z założeniem 1°). Pozostałe zasady porównywania wektorów w takiej wielowymiarowej przestrzeni nie ulegają zmianie. Porównywanie takich wieloaspektowych wartości informacji może być nieco bardziej kłopotliwe, jednakże daje się ono względnie łatwo zaprogramować na komputery. Dzięki temu wieloaspektowa wartość informacji może być w sposób automatyczny uwzględniana w komputerowych bazach danych, stając się dodatkowym narzędziem selekcji informacji pod kątem indywidualnych potrzeb użytkowników.



Rys. 3



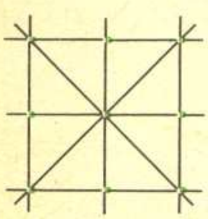
Rys. 4



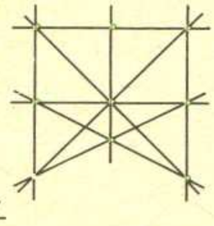
Rys. 5

Zadania, których nie umiemy rozwiązać

Najpierw te, które umiemy:

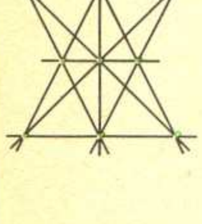


1. Jak można posadzić 9 drzew w 8 rzędach po trzy w każdym? Odpowiedź widoczna jest na rysunku.



2. A 9 drzew w 9 rzędach po trzy w każdym? Też łatwo (rysunek).

3. No to może ogólnie: ustawiamy m drzew w r rzędach po s drzew w każdym. Jaka jest największa wartość stosunku rs/m?



Tego zadania nie umiemy rozwiązać ani my, ani Czytelnicy *Journal of Recreational Mathematics* 14 (1), 1981—2. Można ustawić n^2 drzew w sposób podobny do tego z rysunku do zadania 1. Największą wartość wyrażenia rs/n^2 dostajemy wówczas dla $n = 3$, mianowicie 8/3. Ale ustawienie „w twierdzenia Pappusa" (rysunek) jest lepsze (dla $n = 3$ daje stosunek $rs/n^2 = 10/3$) i przypuszczalnie to jest szukane maksimum.

Dlaczego nazwaliśmy to ustawieniem „w twierdzenia Pappusa"? To jasne, jeżeli tylko przypomnimy to twierdzenie:

Jeżeli punkty A, B, C są współliniowe i punkty D, E, F też, to punkty przecięcia przeciwległych boków sześciokąta $AFBDCE$ także są współliniowe.