

Burzliwe początki cyfrowego tysiąclecia

Tomasz IDZIASZEK



Rozpoczynając na początku lat 80. XX wieku studia doktoranckie na Uniwersytecie w Erlangen, Karlheinz Brandenburg raczej nie przypuszczał, że wyniki jego pracy przyczynią się do zrewolucjonizowania branży muzycznej na całym świecie. A zaczęło się od tego, że jego promotor, profesor Dieter Seitzer, rozważał zagadnienie przesyłania muzyki liniami telefonicznymi. Ze względu na znaczne pogorszenie jakości dźwięku analogowy przesył nie wchodził w grę, należało więc użyć technologii ISDN (*Integrated Services for Digital Network*) umożliwiającej jednoczesne przesyłanie tymi liniami sygnału cyfrowego (co w niedalekiej przyszłości wykorzystano do dostarczenia Internetu do masowego odbiorcy). Problemem była zbyt mała przepustowość tego kanału wynosząca 128 kb/s (kilobitów na sekundę). Świeżo wypromowany standard płyty kompaktowej opisywał kodowanie dźwięku, w którym na jedną minutę potrzebne jest około 10 MB (pisaliśmy o tym w *Delcie* 10/2011), zatem wymagana przepustowość wynosiła 1411,2 kb/s, czyli nieco ponad 11 razy więcej niż umożliwiała ISDN. Próbuąc opatentować swój pomysł, Seitzer został odesłany z kwitkiem z biura patentowego, gdyż „nie patentuje się rzeczy niemożliwych”.

Brandenburg, próbując zrealizować wizję Seitzera, postanowił zmniejszyć ilość danych potrzebnych do zapisu dźwięku, czyli dokonać jego *kompresji*. Niestety, metody kompresji bezstratnej (które redukują liczbę bitów potrzebną do przesłania redundantnych informacji, ale w taki sposób, że nadal jest możliwe wierne odtworzenie pierwotnych danych) nie za dużo mogą pomóc w przypadku cyfrowej muzyki (średnio można przyjąć, że pozwalają zmniejszyć rozmiar danych o połowę). Potrzebna była *kompresja stratna*, polegająca na usunięciu części informacji, ale w taki sposób, by zostało to niezauważone przez odbiorcę.

Aby podjąć decyzję, które dane mogą być usunięte, Brandenburg skorzystał z *modelu psychoakustycznego*, który opisuje, jak działają uszy i mózg w odbieraniu dźwięku. Przykładowo, choć człowiek słyszy dźwięki w zakresie



Zadania

Redaguje Tomasz TKOCZ

M 1462. Na płaszczyźnie dany jest zbiór punktów \mathcal{S} . Mówimy, że punkt P jest widoczny z punktu A , jeśli odcinek AP jest zawarty w \mathcal{S} . Zbiór \mathcal{S} jest widoczny z punktu A , jeśli każdy jego punkt jest widoczny z A .

Wykazać, że zbiór \mathcal{S} jest widoczny z każdego punktu trójkąta ABC , jeśli jest widoczny z każdego wierzchołka tego trójkąta.

Rozwiązanie na str. 8

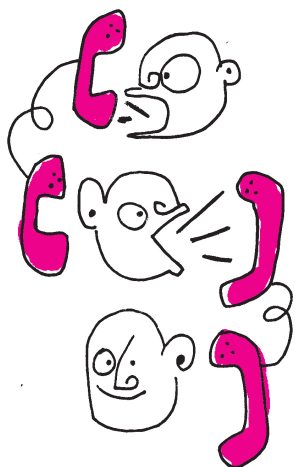
M 1463. Udowodnić, że dla liczby całkowitej dodatniej n oraz liczb rzeczywistych x, y spełniających $x + y = 1$, prawdziwa jest tożsamość

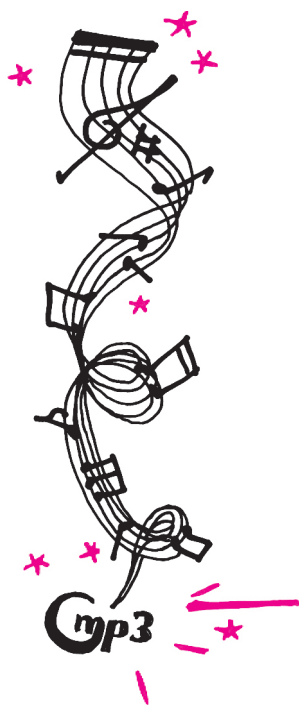
$$x^{n+1} \sum_{k=0}^n \binom{n+k}{k} y^k + y^{n+1} \sum_{k=0}^n \binom{n+k}{k} x^k = 1.$$

Rozwiązanie na str. 9

M 1464. Każda spośród n osób zna dokładnie jedną wiadomość i każda z tych n wiadomości jest inna. Co jakiś czas pewna osoba dzwoni do innej i przekazuje jej wszystko, co wie (nie poznając żadnej informacji w zamian). Ile co najmniej rozmów telefonicznych musi się odbyć, zanim wszystkie osoby będą znać wszystkie informacje?

Rozwiązanie na str. 11





od 12 Hz do 20 kHz, to maksymalna czułość występuje w zakresie od 2 kHz do 4 kHz. Zatem można rozdzielić sygnał dźwiękowy na kilka pasm obejmujących różne zakresy częstotliwości i zakodować z mniejszą dokładnością pasma odbierane przez ludzkie ucho z mniejszą czułością. Model opisuje też zjawisko *maskowania*, które polega na tym, że głośnie dźwięk potrafi zagłuszyć cichsze dźwięki o podobnej częstotliwości, a także cichsze dźwięki, które występują zaraz po nim lub tuż przed nim. Dzięki temu te cichsze dźwięki również można zapisać, używając mniejszej liczby bitów.

W 1988 roku powołano międzynarodową grupę standaryzacyjną MPEG (*Moving Picture Experts Group*) zrzeszającą inżynierów z różnych ośrodków naukowych, których celem było wypracowanie algorytmów kompresji obrazu i dźwięku. W ramach jednego z jej zespołów kontynuował swe badania Brandenburg. W wywiadach przywołuje on różne anegdoty z tego okresu, jak choćby to, że dobierając parametry algorytmu kompresji, przesłuchał niemal tysiąc razy piosenkę Suzanne Vega *Tom's Dinner*, lub to, że kolejne nieudane próby zaimplementowania teoretycznie poprawnego algorytmu spowodowane były błędem w używanym przez niego kompilatorze. Ostatecznie prace zespołu zaowocowały przyjętym trzy lata później standardem kompresji dźwięku nazwanym oficjalnie MPEG-1 Audio Layer III, który urzeczywistnił marzenie Seitzera, gdyż umożliwił kompresję dźwięku z przepustowością 128 kb/s bez znacznego pogorszenia jego jakości.

Początkowo opracowany standard nie przyjął się w przemyśle jako zbyt skomplikowany i wymagający zbyt dużej mocy obliczeniowej, jak na możliwości ówczesnego sprzętu (do użytku wszedł prostszy standard MPEG-1 Audio Layer II, aktualnie nadal stosowany do przesyłu audycji radia cyfrowego). Ale gdy w 1997 roku firma Intel wypuściła na rynek procesor Pentium MMX, komputery osobiste stały się wystarczająco silne, by podołać zadaniu kompresji. Jedną z decyzji, które musieli podjąć koledzy Brandenburga, było przyjęcie trzyliterowego rozszerzenia plików, w których zapisywane byłyby dane zakodowane ich algorytmem. Tak narodził się, dziś znany chyba wszystkim użytkownikom komputerów, skrót *MP3*.



Zadania

Przygotował Andrzej MAJHOFER

F 883. Linie widma promieniowania elektromagnetycznego wysyłanego przez atomy deuteru są nieznacznie przesunięte w stosunku do analogicznych linii emitowanych przez standardowy wodór (z protonem jako jądrem atomu). Jaką prędkość v ruchu przypisze atomowi astronom, jeśli badał widmo spoczywającego deuteru w przekonaniu, że to wodór? Masa elektronu $m_e = 0,511 \text{ MeV}/c^2$, masa protonu $m_p = 938 \text{ MeV}/c^2$, masa deuteronu (jądra deuteru) $m_d = 1876 \text{ MeV}/c^2$, a prędkość światła $c = 3 \cdot 10^5 \text{ km/s}$.

Rozwiązanie na str. 3

F 884. Cząsteczki chemiczne poza ruchem postępowym wykonują także oscylacje i rotacje. Energie wzbudzeń oscylacyjnych i rotacyjnych zależą od składu atomowego i stanu elektronów. Dla cząsteczki dwuatomowej energie oscylacji wynoszą $E_V = \hbar\omega(\nu + 1/2)$, gdzie liczba $\nu = 0, 1, 2, 3, \dots$ numeruje kolejne stany oscylacyjne, a energie stanów rotacyjnych to $E_R = BJ(J + 1)$, dla $J = 0, 1, 2, 3, \dots$, $\hbar = h/2\pi$, zaś h oznacza stałą Plancka. Wielkość B nazywa się stałą rotacyjną cząsteczki w danym stanie atomowym. Dla cząsteczki tlenu węgla $^{12}\text{C}^{16}\text{O}$ w stanie podstawowym powłok elektronowych odpowiednie stałe wynoszą: $\omega = 2169,8 \text{ hc/cm}$, $B = 1,931 \text{ hc/cm}$, gdzie c oznacza prędkość światła. Odległość równowagowa jąder węgla i tlenu wynosi $R = 1,128 \cdot 10^{-8} \text{ cm}$. Jakiej zmianie ulegnie widmo promieniowania, gdy zamienimy atomy węgla i tlenu na ich cięższe izotopy, otrzymując cząsteczkę $^{14}\text{C}^{18}\text{O}$?

Rozwiązanie na str. 5

