

Puzzelen

De puzzels in Java Magazine hebben sinds het begin van dit jaar een nieuw karakter en een nieuwe vorm. Wouter Klein Heerenbrink, die twee jaar lang de Puzzlers van Joshua Bloch en Neil Gafter bewerkte, neemt het heft in eigen handen. U bent gewaarschuwd: het wordt er niet eenvoudiger op. Voor deze puzzeltjes schrijf je een handig algoritme dat eens iets afwijkt van de standaardprocedures. Ze gebruiken basismethoden uit de programmeertaal. Het zijn programmeerpuzzels, die je (ook) met Java kunt oplossen. Antwoorden kunt u naar de redactie mailen. Onder de goede inzenders wordt telkens een IT-gerelateerd boek verloot.

Op je nummer worden gezet

In Zweden heeft iedere burger een uniek nummer, het persoonsnummer. In Nederland kennen wij dit nummer als het burgerservicenummer, voorheen sofi-nummer. Deze nummers worden in Nederland en Zweden voor vrijwel dezelfde doeleinden gebruikt, maar de opbouw van het nummer is totaal anders. De Zweedse variant is veel leuker om te analyseren, dus daar gaan we (I).

Als voorbeeldpersoon nemen we de koning van Zweden: Carl XVI Gustaf. Zijn nummer zou bijvoorbeeld het volgende kunnen zijn: 460430-1285. We definiëren het formaat van dit nummer als yymmdd-ccgr. Het eerste deel voor het streepje, yymmdd, is de geboortedatum van de persoon in kwestie. De koning van Zweden is geboren op 30 april 1946, wat maakt dat het eerste gedeelte van zijn nummer 460430 is.

Het tweede deel, de cc, wordt bepaald door de provincie (in het Zweeds län) waarin je geboren bent. Iedere provincie heeft een code toegewezen gekregen, grote provincies hebben meerdere codes. De koning is geboren in Stockholms län, dat het codebereik 00 tot en met 13 heeft. In het voorbeeld hebben we binnen het bereik gekozen voor de code 12. Verhuis je vanuit het buitenland naar Zweden dan is de tweede c een 9.

Vervolgens is er het getal g. Dit getal is een unieke toevoeging, zodat iedereen in een bepaalde provincie die op dezelfde datum geboren is, toch nog een uniek nummer kan hebben. Tevens zit er een betekenis verwerkt in dit getal. Mannen hebben een even getal, vrouwen een oneven. Een opmerkelijk detail aan deze betekenis, is dat bij een operationele geslachtswisseling het nummer wordt aangepast.

En dan is er nog het getal r. Dit getal is een zogenaamde checksum, een controlegetal dat verkregen wordt door een algoritme toe te passen op de voorgaande getallen in de code. Het algoritme levert uiteindelijk een ééncijferig getal op dat achter de code kan worden geplakt.

Dit controlegetal maakt het mogelijk om het persoonsnummer te valideren wanneer dit wordt ingevuld in bijvoorbeeld een belastingformulier. Het algoritme wordt dan uitgevoerd op het eerste deel van de code en wordt vergeleken met het laatste cijfer. Komen deze niet overeen dan is er een (type)fout gemaakt, komen ze wel overeen dan is de kans groot dat de code goed ingevoerd is (II).

Het algoritme

Het algoritme bestaat uit drie stappen. Voor het gemak van schrijven definiëren we de code zonder checksum als $F(n) = '460430-128'$ en definiëren we de Zweedse sleutel $G(n) = '212121-212'$.

Ter visualisatie schrijven we de waarden in tabel 1. Op de eerste rij schrijven we de waarden van n (1, 2, 3, ..., 9). Op de tweede rij plaatsen we de bijhorende waarden van $F(n)$ en op de derde rij de waarden van $G(n)$.

n	1	2	3	4	5	6	7	8	9
$F(n)$	4	6	0	4	3	0	1	2	8
$G(n)$	2	1	2	1	2	1	2	1	2
Stap 1	8	6	0	4	6	0	2	2	1+6=7

Tabel 1: Stap 1 vermenigvuldigt cijfer uit F en G.

Stap 1: Vermenigvuldig ieder cijfer uit de code $F(n)$ met het bijbehorende cijfer uit de sleutel $G(n)$. Voor $n=1$ vermenigvuldigen we 4 met 2 en krijgen we 8. Dit getal schrijven we op de vierde rij.



Op je nummer worden gezet

Er mag telkens maar één cijfer worden opgeschreven. Mocht het antwoord groter zijn dan 9 en dus uit twee cijfers bestaan dan tellen we de losse cijfers uit het antwoord bij elkaar op. Dit is van toepassing voor $n=9$, waar we 8 met 2 vermenigvuldigen wat 16 als antwoord heeft. We tellen de de losse cijfers 1 en 6 bij elkaar op en we krijgen 7.

Stap 2: Vervolgens tellen we alle getallen uit stap 1 bij elkaar op en het antwoord doen we modulo 10. In ons geval wordt dit $8+6+0+4+6+0+2+2+7=35$; $35 \text{ modulo } 10 = 5$ (III).

Stap 3: Tot slot trekken we 10 van het antwoord af en nemen we hier de absolute waarde van: $|5-10| = 5$. Dit cijfer 5 is de checksum en die voegen we achter de oorspronkelijke code F. Het resultaat is een geldig Zweeds persoonsnummer: 460430-1285.

Terug naar ons eigen land

Het Nederlandse burgerservicenummer heeft geen geslacht, geboortedatum of provincienummer in zich verwerkt. Dit komt voort uit de eis die opgenomen is in de wet betreffende regeling burgerservicenummer: "Ieder gegenereerd nummer moet informatieloos zijn." (IV).

Ook is er geen checksum opgenomen in de code zelf. Wel is er een proef om (type)fouten te detecteren. De proef is een variant op de 11-proef die gebruikt wordt om rekeningnummers te valideren. Bij deze proef vermenigvuldig je het eerste cijfer van het burgerservicenummer met 9, het tweede met 8 enz. Het voorlaatste cijfer vermenigvuldig je met 2 en, hier is de variatie op de 11-proef, het laatste cijfer met -1. Alle verkregen getallen tel je bij elkaar op. Dit resultaat is, wanneer geldig, deelbaar door elf. Met andere woorden: het antwoord modulo 11 is nul.

Met behulp van deze checksums kunnen logische of onbewuste fouten zoals een typfout of verspreking in de meeste gevallen worden ondervangen. Checksums worden ingezet op veel terreinen waarbij er gemakkelijk een fout kan worden gemaakt bij het invullen van cijfercodes. Denk hierbij dus aan servicenummers en rekeningnummers, maar ook aan communicatie tussen computers over computernetwerken. Over dit laatste binnenkort meer.

Puzzel

Een checksum implementeren is leuk, maar zo makkelijk kom je natuurlijk niet weg. Onderstaand een klein stukje van de lijst met namen van inwoners van Zweden met

hun naam, geboortedatum, geboorteplaats en geslacht.

Aan jou de taak om de kolom met geldige persoonsnummers te vullen. Hierbij mag je aannemen dat de gegeven lijst met namen, alle inwoners van Zweden bevat en er dus geen andere nummers in omloop zijn.

De volledige lijst is te vinden op de website van JavaMagazine, inclusief een lijst met de provincienummers.

- (I) We spreken hier over de variant van voor 1990. In 1990 is er een vernieuwing van het systeem doorgevoerd waardoor nieuwe nummers iets kunnen afwijken.
- (II) Hoe groot de kans is bewaren we voor een andere puzzel
- (III) Bij een modulo operatie voer je een deling uit met gehele getallen. Je trekt de waarde waardoor gedeeld wordt zo vaak mogelijk van de invoer af. Het getal dat overblijft wanneer aftrekken niet meer mogelijk is, is het resultaat van de modulo operatie. Bijvoorbeeld: 7 modulo 5 is 2; 10 modulo 10 is 0
- (IV) http://wetten.overheid.nl/BWBR0022835/Bijlage/geldigheidsdatum_15-06-2011

Genetische algoritmen

En de winnaar is... Andre Kampert.

Hij zag kans de puzzel op te lossen in slechts 480 iteraties. De eerstvolgende van de deelnemers had daar tenminste 1000 iteraties voor nodig. Eén deelnemer viel af, omdat hij het algoritme niet zelf had geschreven, maar gedownload en ingesteld.

Dat was natuurlijk niet de bedoeling.