Oavgörbara problem i elementär aritmetik

Erik Palmgren
Matematiska institutionen, Uppsala universitet

2 maj 2002

Många har nog slagits av hur svår talteori kan vara redan på den elementära nivån. Det är förhållandevis enkelt att råka skriva ner en Diofantisk ekvation där det inte finns någon känd metod för att bestämma om den har någon lösning. Fermats stora "sats" är ett bra exempel på detta, vilken man först efter 350 år lyckades bevisa. Ett objektivt skäl till att talteorin är svår är att den kan koda kombinatoriska problem. Detta upptäcktes, och användes på ett genialiskt sätt, av Kurt Gödel i hans berömda ofullständighetssats — det kanske viktigaste resultatet inom logiken.

Genom att bara använda symbolerna för addition, multiplikation, noll, ett och logiska symboler kan man formulera problem i elementär heltalsaritmetik som är principiellt olösbara. Ingen dator kan avgöra problemen, oavsett hur mycket tid eller minnesutrymme den har tillgång till. Vi kommer i denna artikel (avsnitt 1 – 4) att visa hur detta kan vara möjligt. Problem av samma form, fast om den vanliga (reella) tallinjen, visar sig i princip kunna avgöras av en dator (avsnitt 5). Den metod som ligger till grund för detta kan användas för att lösa problem i analytisk geometri på ett helt automatiskt sätt. Vi ger också exempel på hur den kan användas i ett populärt datoralgebrasystem.

1 Matematisk logik och Hilberts problem

Matematisk logik studerar väsentligen formerna för matematiska resonemang och beräkningar, och är en relativt ung gren av matematiken. Runt sekelskiftet 1900 genomgick matematiken den så kallade grundvalskrisen, som hade sitt upphov i paradoxer (Russells paradox, Burali-Fortis paradox) som upptäckts i alltför lättvindiga axiomatiseringar av mängdlära. För att säkerställa den nya axiomatiska metoden föreslog en av den tidens ledande matematiker, David Hilbert, att axiomatiska system och matematiska be-

vis själva skulle studeras med matematiska metoder. Syftet var att bevisa att inga paradoxer kunde uppstå i ett givet system, genom att undersöka formen hos de möjliga bevisen. För att detta skall vara meningsfullt måste naturligtvis de metoder med vilket detta görs själva vara fria från paradoxer eller motsägelser. Dödsstöten till detta Hilberts ambitiösa program kom när Kurt Gödel 1931 bevisade sin andra ofullständighetssats, som i grova drag säger att varje motsägelsefritt system som innehåller axiom för grundläggande talteori inte (ens) kan bevisa sin egen motsägelsefrihet. Ur den första ofullständighetssatsen framgick att redan elementär talteori kunde innehålla oavgörbara problem. Det senare resultatet stukade den rådande optimismen om matematikens obegränsade framstegsmöjligheter, som Hilbert hade formulerat i ordalagen *inom matematiken finns inget ovetbart.*

Vid den internationella matematikerkongressen 1900 presenterade Hilbert en lista av 23 problem för det kommande århundradet. Dessa problem har haft ett stort inflytande på matematikens utveckling. Tre av problemen kan, åtminstone i efterhand, hänföras till logiken. Det andra, konsistensproblemet, kan i en mening sägas vara om inte löst, så "upplöst", av Gödels satser. Hilberts program lever dock vidare i modifierad form inom en gren av logiken, bevisteori, där man studerar om en teori är motsägelsefri relativt en annan, och försöker finna teorier vars motsägelsefrihet är evident i någon mening. Viktiga bidrag till detta progam har gjorts av den svenske logikern Per Martin-Löf. Det första problemet, om Cantors kontinuumhypotes, löstes 1963 av Paul Cohen genom hans bevis att hypotesen är oberoende av Zermelo-Fraenkels axiom (ZFC) för mängdläran. Det 10:e problemet i listan motstod länge lösningsförsöken:

> **10. Bestämning av lösbarheten av en Diofantisk ekvation.** *En Diofantisk ekvation med ett godtyckligt antal obekanta och med rationella heltalskoefficienter är given: ange ett förfarande, med vilket det efter ett ändligt antal operationer kan bestämmas om ekvationen är lösbar i rationella heltal.* (Författarens översättning.)

De (rationella) heltalen är $\mathbf{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. I modernt språkbruk efterfrågar Hilbert en algoritm för att avgöra lösbarheten över $\mathbf{Z}$. Det Diofantiska problemet över $\mathbf{Z}$ med en obekant kan skrivas som

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0$$

där koefficienterna är heltal. Man kan förstås använda allmänna metoder för lösning av sådana ekvationer över komplexa tal, men för heltalslösningar finns

enklare knep att tillgå. En enkel uppskattning av storleken på lösningarna z ges av

$$|z| \leq 1 + |a_n|^{-1} \max(|a_n|, |a_{n-1}|, \ldots, |a_1|, |a_0|).$$

Det finns endast ändligt många heltal z som uppfyller olikheten. En metod för att för att avgöra lösbarheten i $\mathbf{Z}$ är helt enkelt att testa dessa heltal.

När det diofantiska problemet har flera variabler kan lösningarna oftast inte begränsas så att systematisk testning blir en giltig metod. Byggandes på tidigare resultat av Martin Davis, Hilary Putnam och Julia Robinson bevisade den ryske matematikern Juri Matijasevic 1971 att det allmänna problemet är algoritmiskt oavgörbart. Efter ett par inledande avsnitt som förklarar fundamentala logiska begreppp, skall vi försöka ge en idé om beviset och dess talteoretiska ingredienser.

2 Elementära problem

Ett av logikens bidrag under 1800-talet var att precisera och införa symbolism för det vardagliga språk som används när man resonerar om matematik. Detta banade vägen för utvecklingen av programmeringsspråken och för möjligheten att automatisera även teoretisk matematik. George Boole gav symbolism och en logisk kalkyl (den booleska algebran) för ord som *och, eller, inte, medför, om*. Senare behandlades även uttryck som *för alla* x, *det finns* y i den så kallade predikatkalkylen av Gottlob Frege. Vi skall introducera begreppet *elementärt* (eller första ordningens) matematiskt språk. "Elementärt" syftar här på att de objekt som kan diskuteras i språket alla är element i en fixerad mängd D. Godtyckliga delmängder av D faller utanför språkets uttrycksmöjligheter. Vi kommer huvudsakligen att betrakta fallen då D är någon av följande mängder, de naturliga talen $\mathbf{N} = \{0, 1, 2, 3, \ldots\}$, mängden av hela tal $\mathbf{Z}$, mängden av rationella tal $\mathbf{Q}$ eller mängden av reella tal $\mathbf{R}$. I samtliga dessa mängder har de aritmetiska operationerna $+$ och $\cdot$ en överensstämmande mening. Ett påstående som

$$\textit{Det finns något } x \in D \textit{ sådant att } y = x \cdot x \qquad (1)$$

har därför mening i var och en av de nämnda talsystemen, fast olika mening! För $D = \mathbf{N}, \mathbf{Z}$ säger påståendet att y är en perfekt kvadrat, för $D = \mathbf{R}$ betyder det endast att y är icke-negativ. Ett gemensamt elementärt språk bestäms av vokabulären $L = \{+, \cdot, 0, 1\}$. En *term* över L är ett välbildat uttryck som kan skrivas med hjälp av operationerna, $+$ och $\cdot$, samt variabelsymboler $x, y, z, u, v, \ldots$, konstanterna $0, 1$ och parenteser. En sådan term

kan betraktas som ett polynom med naturliga tal som koefficienter. Här är ett exempel på en term och motsvarande normaliserade polynom:

$$(x \cdot (x \cdot x)) + (((1 + 1) \cdot x) \cdot y) \qquad\qquad x^3 + 2xy$$

Ett *atomärt påstående* över L är en ekvation av formen $r = t$, där r och t är termer över L. De atomära påståenden kombineras sedan till *elementära påståenden* över L genom att använda olika logiska symboler. Vi påminner om de satslogiska konnektiven: $P \wedge Q$ betecknar P *och* Q, $P \vee Q$ betecknar P *eller* Q, $\neg P$ står för *icke* P. *Kvantifikatorerna* (eller *kvantorerna*) talar, som namnet antyder, om för hur många element x ett påstående $P(x)$ gäller. Deras räkneförmåga är högst begränsad: *universalkvantorn* $\forall$ kan bara säga att $P(x)$ gäller *för alla* x, medan *existenskvantorn* $\exists$ bara kan säga att $P(x)$ gäller *för minst ett* x. Sålunda uttrycks existenspåståendet *det finns något* $x \in D$ *så att* $P(x)$ med $(\exists x \in D)\, P(x)$, medan universalpåståendet *för alla* $x \in D$ *gäller* $P(x)$ uttrycks som $(\forall x \in D)\, P(x)$. Detta är alla tillåtna sätt att bilda elementära påstånden över L. Påståendet (1) blir i logisk symbolism

$$(\exists x \in D)\, y = x \cdot x. \qquad (2)$$

En variabel som inte ligger inom en kvantifierad variabels räckvidd kallas *parameter*. Den enda parametern i (2) är y. Nya begrepp kan definieras i termer av elementära påståenden med parametrar. Ordningsrelationen $<$ på $\mathbf{N}$ kan definieras så här

$$x < y \Leftrightarrow_{\text{def}} (\exists z \in \mathbf{N})\, x + 1 + z = y. \qquad (3)$$

Det vill säga $x < y$ gäller exakt då det finns något naturligt tal z som adderat till $x+1$ blir y. Vi överlämnar åt läsaren att dechiffrera följande definitioner:

$$\begin{aligned} r = \text{Rest}(m, n) &\Leftrightarrow_{\text{def}} (\exists k \in \mathbf{N})\, m = kn + r \wedge (0 = r \vee 0 < r) \wedge r < n \\ k = \text{Kvot}(m, n) &\Leftrightarrow_{\text{def}} (\exists r \in \mathbf{N})\, m = kn + r \wedge (0 = r \vee 0 < r) \wedge r < n \\ p \text{ primtal} &\Leftrightarrow_{\text{def}} p > 1 \wedge (\forall m \in \mathbf{N})\, [1 < m < p \rightarrow \text{Rest}(p, m) > 0] \end{aligned}$$

De naturliga talen tillsammans med operationerna addition och multiplikation och konstanterna 0 och 1 bildar vad man brukar kalla en *struktur*. Vi betecknar denna

$$\mathcal{N} = \langle \mathbf{N}; +, \cdot, 0, 1 \rangle$$

för att skilja den från mängden $\mathbf{N}$.

Mer allmänt kan man tala om en struktur

$$\mathcal{D} = \langle D; R_1, R_2, \ldots, f_1, f_2, \ldots, c_1, c_2, \ldots \rangle$$

vilken består av en mängd D och där R:n är relationer på D, f:n är operationer på D, och c:na är konstanter i D. Ett *elementärt påstående* över denna får då använda konstanterna, operationerna och variabler för att bygga upp termer, och de atomära påståendena får bildas med relationerna inklusive $=$. Till exempel, om R_1 är relationen $<$ så är olikheten $r < t$ ett atomärt påstående närhelst r och t är termer. Vi skriver $\mathcal{D} \models P(x_1, \ldots, x_n)$ för att beteckna att ett elementärt påstående P över $\mathcal{D}$ är sant för parametrarna $x_1, \ldots, x_n$. I denna notation skriver man inte ut D i kvantoruttryck som $\forall x \in D$ och $\exists x \in D$, eftersom dessa kan utläsas ifrån $\mathcal{D}$. Exempelvis betyder $\mathcal{N} \models (\exists z)\ x + z + 1 = y$ det samma som $(\exists z \in \mathbf{N})\ x + z + 1 = y$

En relation R på en mängd D brukar identifieras med en delmängd R av mängden D^n av n-tuppler $(x_1, \ldots, x_n)$ av element från D. Längden n är fixerad för varje relation. Exempelvis motsvarar likhetsrelationen $=$ på $\mathbf{N}$ mängden av par $(n = 2)$

$$\{(0,0), (1,1), (2,2), \ldots\}.$$

Man skriver omväxlande $R(x_1, \ldots, x_n)$ eller $(x_1, \ldots, x_n) \in R$ för att uttrycka att relationen R gäller mellan elementen $x_1, \ldots, x_n$. Relationen $R \subseteq D^n$ är *definierbar över* $\mathcal{D}$ om det finns ett elementärt påstående $P(x_1, \ldots, x_n)$ över $\mathcal{D}$ så att $R(x_1, \ldots, x_n)$ om, och endast om, $\mathcal{D} \models P(x_1, \ldots, x_n)$. Som vi såg ovan är till exempel $<$ definierbar över $\mathcal{N}$. En funktion $f : D^n \to D$ är definierbar om relationen $f(x_1, \ldots, x_n) = x_{n+1}$ är definierbar.

Det fundamentala begreppet vi skall intressera oss för är detta. Vi säger att en struktur $\mathcal{D}$ är *avgörbar* om det finns en algoritm som för varje elementärt påstående P över $\mathcal{D}$, utan parametrar, avgör om P är sann eller falsk. Vi kommer att betrakta strukturerna $\mathcal{N}$ samt $\mathcal{Z} = \langle \mathbf{Z}; +, \cdot, 0, 1 \rangle$, $\mathcal{Q} = \langle \mathbf{Q}; +, \cdot, 0, 1 \rangle$, $\mathcal{R} = \langle \mathbf{R}; <, +, \cdot, 0, 1 \rangle$ och några till. Det är i dessa fall klart att elementära påståenden kan ges en precis syntax så att de kan läsas in i en dator och behandlas algoritmiskt. I dessa strukturer har också relationerna och operationerna en överstämmande mening, så det är endast den underliggande mängden $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$ eller $\mathbf{R}$ som påverkar betydelsen hos ett påstående.

3 Definierbarhet i de hela och rationella talen

Gud skapade de naturliga talen — resten är människoverk.
Leopold Kronecker (1823-1891)

Vad som kan uttryckas i det elementära språket över en struktur $\mathcal{D}$ bestämmer förstås hur svåra problem som kan formuleras. De tre talsystemen $\mathbf{N}$, $\mathbf{Z}$ och $\mathbf{Q}$ visar sig vara i det närmaste likvärdiga i det avseendet. Ett elementärt påstående om $\mathcal{Z}$ kan översättas till ett ekvivalent påstående om $\mathcal{N}$ genom att betrakta ett heltal som ett par av naturliga tal (varvid deras skillnad är heltalet). Kom ihåg att negativa tal erkändes relativt sent som självständiga storheter inom matematiken! På motsvarande sätt kan en översättning från $\mathcal{Q}$ till $\mathcal{N}$ erhållas genom att betrakta trippler (m, n, k) av naturliga tal som rationella tal $(m-n)/(k+1)$. Detaljerna i detta är inte särskilt svåra att genomföra. Översättningar i motsatt riktning kräver djupare resultat i talteori. Om man kan finna ett elementärt påstående $N(x)$ så att $\mathcal{Z} \models N(x)$ gäller precis då x är ett naturligt tal, kan vi enkelt översätta ett påstående i $\mathcal{N}$ till ett ekvivalent påstående i $\mathcal{Z}$, genom att systematiskt begränsa kvantorernas och parametrarnas utsträckning med $N(x)$. En sådant påstående $N(x)$ kan erhållas genom att utnyttja Lagranges berömda sats att varje naturligt tal är summan av fyra perfekta kvadrater. Vi kan även använda ett annat resultat av Lagrange (och Fermat) om vad som brukar kallas *Pells ekvation*

$$x^2 - Dy^2 = 1.$$

(Denna ekvation spelar även en roll i lösningen av Hilberts 10:e problem.)

Sats *Om D ej är en perfekt kvadrat, så har Pells ekvation en lösning där x och y är positiva heltal.*

Man kan nu utan större besvär kontrollera att påståendet $N(z)$:

$$(\exists u\, z = u^2) \vee \exists x \exists y\, (x^2 = zy^2 + 1 \wedge \neg x = 0 \wedge \neg y = 0)$$

definierar de naturliga talen i $\mathcal{Z}$. En betydligt svårare översättning är den från $\mathcal{Z}$ till $\mathcal{Q}$, som först bevisades av Julia Robinson[1] (se Smoryński 1991).

[1] Julia (Bowman) Robinson (1919 – 1985) blev den första kvinnan att väljas in som ledamot i den matematiska sektionen av National Academy of Sciences (USA), liksom den första kvinnliga ordföranden för American Mathematical Society (1983). Flera läsvärda biografier finns tillgängliga, se till exempel Feferman (1994).

4 Oavgörbara problem

Alan Turing gav 1936 ett övertygande argument för att allt som kan beräknas eller avgöras algoritmiskt kan beräknas eller avgöras av hans Turingmaskiner. En Turingmaskin är en matematisk modell av en dator som består av en centralenhet och ett "sekundärminne" i form av ett oändligt långt band. Vid varje tidpunkt används dock bara en ändlig del av bandet. Indata skrivs på bandet. Maskinen startas, och om den stannar betraktas innehållet på bandet som utdata. En universell Turingmaskin tar en beskrivning e av en annan maskin M och indata x och simulerar sedan verkan av M med indata x. Man kan alltså betrakta e som (simulerings)programmet för M.

Sats (Turing 1936) *Stopproblemet är inte algoritmiskt avgörbart: det finns ingen Turingmaskin som avgör om en Turingmaskin e med indata x stannar efter ett ändligt antal steg.*

Beviset av Gödels ofullständighetssats innehöll många viktiga idéer om hur syntax och beräkningar kan kodas i talteori. En kodningsmetod som nu tillämpas i datorsammanhang är den så kallade ASCII-koden. Den tilldelar alla symboler som finns på ett vanligt tangentbord (och många därtill) ett tal mellan 0 och 255. En text `ZXCV` kan därmed betraktas som ett naturligt tal skrivet i basen 256:

$$\texttt{Z} \cdot 256^3 + \texttt{X} \cdot 256^2 + \texttt{C} \cdot 256^1 + \texttt{V} \cdot 256^0.$$

Vi skall se hur liknande kodningsidéer kan användas i bevis för oavgörbarhetsresultat.

Gödels kodningslemma *Det finns en funktion $\beta : \mathbf{N}^3 \to \mathbf{N}$ definierbar över $\mathcal{N}$ sådan att för varje ändlig följd $a_1, \ldots, a_n$ av naturliga tal, finns naturliga tal x och d så att*

$$\beta(x, d, i) = a_i \qquad (i = 1, \ldots, n).$$

Man skulle med idén ovan kunna betrakta talen $a_1, \ldots, a_n$ som tecken i ett tillräckligt stort alfabet $0, 1, 2, \ldots, d-1$. Sätt nu $x = a_n d^{n-1} + \cdots + a_2 d + a_1 d^0$ där $d = \max(a_1, \ldots, a_n) + 1$. Funktionen β skall alltså läsa av tecknen

ur "texten" x. Om vi antar att exponentialfunktionen x^y är definierbar över $\mathcal{N}$, så kan β konstrueras genom $\beta(x,d,i) = \text{Rest}(\text{Kvot}(x,d^i),d)$. Då gäller $\beta(x,d,i) = a_i$. Emellertid är det inte alls uppenbart att exponentialfunktionen är definierbar över $\mathcal{N}$. Gödel använde istället den *kinesiska restsatsen* för sitt bevis.

> Låter man $d = (an)!$, där $a = \max(a_1,\ldots,a_n)+1$, så är $1+d, 1+2d, \ldots, 1+nd$ parvis relativt prima. Den kinesiska restsatsen ger då ett naturligt tal x med $x \equiv a_i \pmod{1+id}$ för alla $i = 1,\ldots,n$. Vi har $a_i \leq d < 1+id$, så $a_i = \text{Rest}(x, 1+id)$. Man kan alltså låta
>
> $$\beta(x,d,i) = \text{Rest}(x, 1+id),$$
>
> som uppenbarligen är definierbar över $\mathcal{N}$.

Exponentialfunktionen kan nu definieras genom följande komplicerade formel (4), som egentligen inte säger något annat än att det finns ett talpar (x,d) som kodar följden

$$1, b, b^2, b^3, \ldots, b^n$$

och att y är sist i följden.

$$\begin{aligned} y = b^n \quad \Leftrightarrow_{\text{def}} \quad & (\exists x,d \in \mathbf{N})\, y = \beta(x,d,n+1) \wedge \\ & \beta(x,d,1) = 1 \wedge \\ & (\forall i < n)\, \beta(x,d,i+2) = b \cdot \beta(x,d,i+1). \end{aligned} \tag{4}$$

På liknande sätt kan andra rekursiva talföljder definieras, faktiskt alla så kallade *primitivt rekursiva funktioner*. Dessa funktioner kan identifieras med enkla datorprogram (i t.ex. BASIC) som endast använder addition, multiplikation, tilldelningssatser, villkorssatser (IF-THEN-ELSE), begränsade slingor (FOR-loopar), men inga hoppsatser (GOTO) — ett sådant program stannar alltid! Det är välkänt att det finns en primitivt rekursiv funktion $T(e,x,y)$ så att $T(e,x,y) = 1$ om y är en kod för en korrekt och avslutad beräkning i Turingmaskinen e med indata x, och $T(e,x,y) = 0$ i annat fall. Vi ser nu att stopproblemet kan uttryckas genom $(\exists y \in \mathbf{N})\, T(e,x,y) = 1$, så det är definierbart över $\mathcal{N}$. Av detta följer med hjälp av Turings sats att

Sats Strukturen $\mathcal{N}$ av naturliga tal med multiplikation och addition är oavgörbar.

Av denna sats följer dock inte att Hilberts 10:e problem är oavgörbart. För detta krävs betydligt större ansträngning som vi skall se nedan.

Hilberts tionde problem

Det finns två principiellt olika sätt att beskriva oändliga mängder med hjälp av algoritmer. Det ena är att en algoritm räknar upp elementen i mängden ett i taget men utan någon särskild ordning mellan dem (och möjligen med upprepningar). Det andra är att en algoritm givet ett element avgör om det tillhör mängden (i detta fall kallas mängden *avgörbar*). Elementen i grundmängden måste ha någon form av ändlig beskrivning. Det är ingen principiell inskränkning att anta att denna kan ges av naturliga tal.

En (icke-tom) mängd S kallas *algoritmiskt uppräknelig* om det finns en algoritm $n \mapsto s_n$ som räknar upp (eller genererar) alla element i S. Följande algoritm visar att mängden av primtal är algoritmiskt uppräknelig: Givet är ett naturligt tal n. Om $n < 2$ eller är jämnt delbar med något heltal $1 < k < n$ så svara $s_n = 2$ (ett säkert primtal), annars är n ett nytt primtal och vi svarar $s_n = n$. Följden som räknas upp av algoritmen

$$2, 2, 2, 3, 2, 5, 2, 7, 2, 2, 2, 11, 2, 13, \ldots$$

innehåller upprepningar av primtalet 2 men detta spelar ingen roll. Vi såg att vi här använde oss av en delalgoritm som avgör om ett tal är ett primtal.

Stoppmängden S är mängden av par (e, x) av Turingmaskiner e och indata x sådana att e stannar efter ändligt många steg givet x. Något överraskande är denna mängd algoritmiskt uppräknelig. Låt e_0 vara en maskin som stannar för varje indata. Följande algoritm genererar S: Givet (e, x, y), om $T(e, x, y) = 1$ svara (e, x), annars svara (e_0, x). Enligt Turings sats är denna mängd S inte avgörbar. Detta visar att det algoritmiskt sett är lättare att räkna upp elementen i en mängd än att avgöra om ett givet element tillhör mängden.

En relation på naturliga tal är *Diofantisk* om den bestäms av lösbarheten hos en viss Diofantisk ekvation. Ett enkelt exempel är $<$ som den definieras i (3) ovan. Mer formellt är en *Diofantisk relation* $R \subseteq \mathbf{N}^k$ en relation som kan definieras genom

$$\begin{aligned} R(x_1, \ldots, x_k) \quad \Leftrightarrow_{\text{def}} \quad & (\exists y_1, \ldots, y_n \in \mathbf{N}) \\ & p(x_1, \ldots, x_k, y_1, \ldots, y_n) = q(x_1, \ldots, x_k, y_1, \ldots, y_n) \end{aligned}$$

där p och q är polynom med koefficienter i $\mathbf{N}$. Alternativt kan man säga att R är projektionen av lösningarna till ekvationen $p = q$ i de första k koordinaterna. Det är i princip möjligt att generera lösningarna till ekvationen genom en systematisk testning av tupler (som dock aldrig slutar). Så

Diofantiska relationer och mängder är algoritmiskt uppräkneliga. Omvändningen verkar vid första anblicken osannolik. Men om man kan visa detta har man lösningen till Hilberts 10:e problem (som följer nedan).

Relationen R kallas *exponentiellt Diofantisk* om p och q får vara godtyckliga termer bildade av variabler, konstanter $0, 1$, addition, multiplikation och exponentiering x^y. Exempel: $p = x_1^{1+x_1 \cdot y_1} + x_2$. Martin Davis och Julia Robinson hade visat mycket betydelsefulla delresultat på vägen mot lösningen av Hilberts 10:e problem. Tillsammans med Hilary Putnam lyckades de 1961 visa

Sats *Varje algoritmiskt uppräknelig mängd (eller relation) är exponentiellt Diofantisk.*

Det svåraste problemet återstod att lösa: att visa att relationen $y = b^n$ är Diofantisk. (Definitionen (4) ovan är långt ifrån Diofantisk.) Uppfattningen bland vissa matematiker var att ovanstående sats inte hade något med Hilberts 10:e problem att göra, eftersom exponentialfunktionen spelade en så central roll. Juri Matijasevic lyckades dock 1971 ge en Diofantisk definition av exponentialfunktionen. En metod som ligger nära den i hans ursprungliga lösning är att använda sig av rötterna till Pells ekvation. Om $x_1, y_1 > 0$ är en lösning till $x^2 - Dy^2 = 1$, så är även alla x_n, y_n givna av

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n \qquad (n = 1, 2, 3, \ldots)$$

lösningar. Väljer man $x_1, y_1 > 0$ att vara den lösning där $x_1 + y_1\sqrt{D}$ är minimal, så kan man visa att (x_n, y_n) utgör samtliga positiva lösningar. Det är lätt att se att lösningarna växer exponentiellt med n. Speciella egenskaper hos Pells ekvation för $D = a^2 - 1$, där $a \geq 2$, gör det nu möjligt att erhålla y_n som en Diofantiskt definierbar funktion av n. Från denna funktion kan sedan den vanliga exponentialfunktionen erhållas. Konstruktionerna är intrikata, se (Smoryński 1991) för en detaljerad framställning.

Matijasevic sats (1971) *Varje algoritmiskt uppräknelig mängd (eller relation) är Diofantisk.*

Detta visar att stoppmängden S är Diofantisk. Huvudresultatet följer därmed av att stopproblemet är oavgörbart:

Följdsats *Hilberts 10:e problem är oavgörbart.*

Det är känt att man kan avgöra om ett system av linjära ekvationer har en heltalslösning. I kontrast till detta visar satsen att det inte är möjligt att avgöra om ett *system* $p_1 = 0, \ldots, p_m = 0$ av andragradsekvationer med flera obekanta är lösbart. Detta eftersom termer av högre totalgrad i en ekvation $q = 0$ kan ersättas av andragradstermer till kostnaden av att införa nya variabler och ekvationer. (Exempel: lösbarheten hos $x^3v + p = 0$ är ekvivalent med lösbarheten hos systemet $yz + p = 0, y - x^2 = 0, z - xv = 0$ där y och z är nya variabler.) Detta system är i sin tur ekvivalent med den enda ekvationen $p_1^2 + \cdots + p_n^2 = 0$, vilket betyder att Hilberts 10:e problem är oavgörbart för fjärdegradspolynom. För en andragradsekvation med flera obekanta är dock lösbarhet avgörbar. (Detta bygger på ett djupt resultat av C.-L. Siegel från 1972.) Vad gäller tredjegradsekvationer tycks problemet fortfarande vara öppet (Smoryński 1991).

Det finns även positiva konsekvenser av Matijasevics resultat. Exempelvis finns det ett polynom $r(x_1, \ldots, x_n)$ med heltalskoefficienter sådant att dess positiva värden, när $x_1, \ldots, x_n$ löper över de naturliga talen, är precis primtalen. (De negativa värdena ignoreras alltså.) Vi citerar ett sådant polynom från Davis mfl. (1976) (alla latinska bokstäver från a till z används som variabler):

$$
\begin{aligned}
(k+2)\Big\{ & 1 - \Big(\Big[wz + h + j - q \Big]^2 + \Big[(gk + 2g + k + 1)(h + j) + h - z \Big]^2 \\
& + \Big[16(k+1)^3(k+2)(n+1)^2 + 1 - f^2 \Big]^2 + \Big[2n + p + q + z - e \Big]^2 \\
& + \Big[e^3(e+2)(a+1)^2 + 1 - o^2 \Big]^2 + \Big[(a^2 - 1)y^2 + 1 - x^2 \Big]^2 + \Big[16r^2y^4(a^2 - 1) + 1 - u^2 \Big]^2 \\
& + \Big[((a + u^2(u^2 - a))^2 - 1)(n + 4dy)^2 + 1 - (x + cu)^2 \Big]^2 + \Big[(a^2 - 1)\ell^2 + 1 - m^2 \Big]^2 \\
& + \Big[ai + k + 1 - \ell - i \Big]^2 + \Big[n + \ell + v - y \Big]^2 + \Big[p + \ell(a - n - 1) + b(2an + 2a - n^2 - 2n - 2) - m \Big]^2 \\
& + \Big[q + y(a - p - 1) + s(2ap + 2a - p^2 - 2p - 2) - x \Big]^2 + \Big[z + p\ell(a - p) + t(2ap - p^2 - 1) - pm \Big]^2 \Big) \Big\}
\end{aligned}
$$

Det finns alltså en aritmetisk formel för att generera alla primtal utan att använda division! Hur kan man komma på ett sådant polynom? Vi såg ovan att mängden P av primtal är algoritmiskt uppräknelig. Därmed är denna mängd Diofantisk, och det finns alltså ett polynom $p(x_1, \ldots, x_n, y)$ med heltalkoefficienter så att y är ett primtal om, och endast om, $(\exists x_1, \ldots, x_n \in$

$\mathbf{N})\, p(x_1, \ldots, x_n, y) = 0$. Nu kan r definieras genom ett enkelt trick:

$$r(x_1, \ldots, x_n, y) = (y+1)(1 - p(x_1, \ldots, x_n, y)^2) - 1.$$

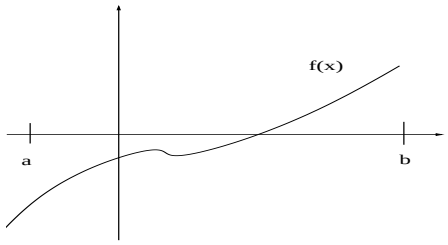
Det finns fortfarande flera olösta problem inom området. Hilberts 10:e problem där lösningar i rationella tal tillåts är fortfarande öppet (2000). Det är ej heller känt om $\mathbf{Z}$ kan definieras genom en Diofantisk relation över $\mathbf{Q}$.

5 Avgörbara problem

Elementär analytisk geometri. Många intressanta problem i analytisk geometri kan formuleras enbart med hjälp av polynom, olikheter och logiska symboler (se exempel (a) – (g) nedan). Det elementära språket över strukturen $\mathcal{R} = \langle \mathbf{R}; +, \cdot, <, 0, 1 \rangle$ är därför synnerligen uttrycksfullt. Den polske logikern Alfred Tarski bevisade under 1930-talet att $\mathcal{R}$ är avgörbar. Detta kan tyckas vara ett förvånande resultat eftersom de reella talen har en mycket mer komplicerad uppbyggnad än till exempel heltalen. Tarskis metod för bevisa avgörbarheten var så kallad *kvantorelimination*. Vi skall kort beskriva några av idéerna i denna.

Att direkt kontrollera om $(\forall x \in \mathbf{R})\, P(x)$ är sann skulle innebära att testa om påståendet $P(x)$ gäller för varje $x \in \mathbf{R}$. Detta är förstås inte möjligt, eftersom $\mathbf{R}$ är oändlig. Om man däremot, steg för steg, kan eliminera kvantorerna $\forall$ och $\exists$ genom omskrivning, finns det möjlighet att slutresultatet är avgörbart. Detta gäller i synnerhet om alla atomära påståenden utan variabler är avgörbara. För strukturen $\mathcal{R}$ räcker det att göra enkla aritmetiska kalkyler på s och t och jämföra leden: $s = t$, $s < t$ eller $s > t$.

Över de reella talen gäller som bekant satsen om mellanliggande värden: om f är en kontinuerlig funktion på intervallet $[a, b]$ med $f(a) \leq 0 \leq f(b)$ så har funktionen ett nollställe i intervallet (se figuren).



Betrakta fallet då f är ett polynom utan multipla nollställen. Om man vet att f har högst ett nollställe i intervallet $[a, b]$, så kan $(\exists x)\, f(x) = 0 \wedge a \leq x \leq b$ skrivas om som det ekvivalenta villkoret $f(a) \leq 0 \leq f(b) \vee f(b) \leq 0 \leq f(a)$. Detta eliminerar existenskvantorn $\exists x$. En välkänd algoritmisk metod enligt Sturm ger ett sätt att isolera samtliga rötter till $f(x) = 0$ inom rationella interval $[a_1, b_1], [a_2, b_2], \ldots, [a_n, b_n]$, så att varje interval innehåller högst en rot. Med hjälp av denna kan kvantorn i $\exists x\, f(x) = 0$ elimineras, även då x är obegränsad. Fallet med multipla rötter kan behandlas genom att betrakta derivatorna av f. Den generella metoden är ganska komplicerad. Stora förbättringar har gjorts av Tarskis ursprungliga algoritm. Tidsåtgången för den bästa kända algoritmen (se Marker 1996) är dock dubbelt exponentiell i antalet kvantorväxlingar $\forall\exists$ och $\exists\forall$ det finns i problemformuleringen när den ställs på Prenex-form.

Följande problem och satser kan tämligen direkt formuleras som elementära utsagor över $\mathcal{R}$ och är därmed enligt Tarskis resultat algoritmiskt lösbara eller bevisbara.

(a) Triangelolikheten och Cauchy-Schwarz' olikhet för det Euklidiska rummet $\mathbf{R}^n$.

(b) Bestäm det minsta tal c sådant att

$$(x - a)(b - x) \leq c(b - a)^2 \qquad (a \leq x \leq b).$$

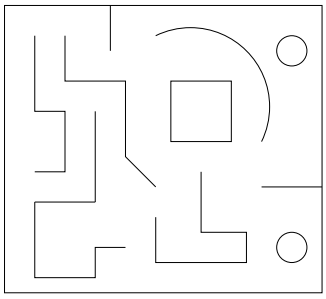
(c) Bestämning av asymptoter: Givet polynom $p(x)$ och $q(x)$, finn λ sådan att

$$(\forall \varepsilon > 0)\, (\exists d > 0)\, (\forall x > d)\, \left| \frac{p(x)}{q(x)} - \lambda \right| < \varepsilon.$$

(d) Givet en n-dimensionell sfär av radie 1. Hur många punkter kan placeras ut på denna så att inget par av punkter ligger närmare varandra än 1? (Euler, $n = 3$)

(e) Hur många enhetssfärer kan packas i en kubisk låda med sidan 10? (Sfärpackningsproblemet)

(f) Ett labyrintiskt konstgalleri skall övervakas. Antag att begränsningsytorna bestäms av ett system av algebraiska olikheter. Hur många övervakningskameror ($\leq n$) behövs för att varje skrymsle av galleriet skall kunna ses från någon av dem?



(g) Rörelseplanering: En robot skall patrullera galleriet i räta linjer. Hur många vändpunkter ($\leq n$) behövs för att den under vägen skall kunna övervaka alla utrymmen och samtidigt undvika att krocka med väggar eller föremål.

Vi påpekar att n skall vara ett fixerat positivt heltal i (a), (d), (f) och (g). För problem (d) och (e) är det möjligt att finna övre gränser för antalet sfärer och punkter genom att utnyttja Dirichlets lådprincip.

Kvantorelimination i praktiken. Kvantoreliminationsmetoden för reella tal kan programmeras för bruk på persondatorer. Det finns ett antal experimentella implementationer (Caviness och Johnson 1998). Sådana rutiner ingår även i det populära datoralgebrasystemet *Mathematica* (Strzebonski 2000). Exempelvis kan problemen (b), (c) och (d) (för cirklar) snabbt lösas automatiskt med hjälp av detta. Problem (b) löses genom följande kommandon i Mathematica 4.0:

```
In[3]:=
      << Developer`; << Experimental`;

In[4]:=
      Resolve[ForAll[{a, b, x}, a ≤ x && x ≤ b && c ∈ Reals,
        (x - a) * (b - x) ≤ c (b - a)^2], {c}]
Out[4]=
      c ≥ 1/4
```

Transcendenta funktioner. En intressant fråga är om Tarskis resultat kan utvidgas till transcendenta funktioner. Detta är ett aktivt forskningsområde

med anknytning till svåra algebraiska problem. Ett enkelt negativt resultat för strukturen $\langle \mathbf{R}; +, \cdot, <, 0, 1, \sin \rangle$ är följande. Nollställena till $\sin x$ är heltalsmultipplerna av π, så varje rationellt tal kan skrivas som kvoten av två sådana nollställen. Följande påstående definierar de rationella talen över $\mathbf{R}$: a rationell om och endast om

$$(\exists x, y, z \in \mathbf{R})\ \sin^2 x + \sin^2 y + (x - z^2 - 1)^2 + (ax - y)^2 = 0.$$

Därav följer, via Robinsons översättningsmetod, att strukturen oavgörbar. Det är ännu ett öppet problem om utvidgningen $\langle \mathbf{R}; +, \cdot, <, 0, 1, \exp \rangle$ med exponentialfunktionen exp är avgörbar. Macintyre och Wilkie (1996) har dock visat avgörbarhet under antagande att en förmodan inom transcendent talteori är sann (Schanuels förmodan). Märkligt nog är ingen generell algoritm känd (Marker 1996) för att avgöra ens variabelfria påståenden över $\langle \mathbf{R}; +, -, \cdot, <, 0, 1, \exp \rangle$, till exempel likheter som

$$e^{e^2-1} - e^5 = e^{6+e^{-3}}.$$

(Just denna är dock falsk.)

Litteratur

B.F. Caviness och J.R. Johnson (red.) (1998). *Quantifier Elimination and Cylindrical Algebraic Decomposition.* Springer-Verlag.

Martin Davis, Yuri Matiyasevich och Julia Robinson (1976). Hilbert's tenth problem. Diophantine equations: positive aspects of a negative solution. I: F.E. Browder (red.) *Mathematical developments arising from Hilbert problems.* American Mathematical Society, Rhode Island.

S. Feferman (1994). *Julia Bowman Robinson (1919 - 1985).* Bibliographical Memoirs, vol. 63. The National Academy Press, Washington D.C. Se även URL: `www.nap.edu/html/biomems/jrobinson.html`

Angus Macintyre och Alex J. Wilkie (1996). On the decidability of the real exponential field. I: P. Odifreddi (red.) *Kreiseliana: about and around Georg Kreisel.* A.K. Peters.

David Marker (1996). Model Theory and Exponentiation. *Notices of the American Mathematical Society* 43:7, pp 753-759.

Yuri Matiyasevich (1993). *Hilbert's Tenth Problem.* The MIT Press, London.

Craig Smoryński (1991). *Logical Number Theory I.* Springer-Verlag.

Adam Strzebonski (2000). Solving Algebraic Inequalities. *The Mathematica Journal 7:4,* pp 525 – 541.

Alfred Tarski (1951) *A Decision Method for Elementary Algebra and Geometry.* University of California Press. Nytryck i Caviness och Johnson (1998).