Om Gödels sætning

Thomas Bolander og Helge Elbrønd Jensen

7. marts 2005

Resumé

Gödels sætning er en af det 20. århundredes mest berømte matematiske sætninger. Den er kendt langt ud over de professionelle matematikeres kreds og indgår i mange diskussioner om logik og erkendelse. Gödels sætning er i sin formelle form vanskelig tilgængelig, hvilket har gjort det svært for ikke-specialister at beskæftige sig mere indgående med sætningen. I denne artikel vil vi give en kort, let læselig fremstilling af hovedlinierne i beviset for Gödels sætning, og søge at formidle en mere grundlæggende forståelse af sætningens indhold og betydning.

1 Indledning

Gödels sætning siger i sin mest markante formulering, at "Der findes sande sætninger, der ikke kan bevises". I denne formulering har sætningen en umiddelbar appel, der rækker langt udover de professionelle matematikeres kredse, og som har givet anledning til mange fortolkninger af dens indhold.

I den ovennævnte formulering indgår to centrale begreber, nemlig "sand" og "bevis", hvis betydning må præciseres, før man kan gå videre med sætningen.

Begrebet "sand" indgår faktisk ikke i den korrekte formulering af Gödels sætning, der udelukkende opererer med formler (sætninger) og formelle beviser, og som under en række forudsætninger siger, at "Der findes formler, der hverken kan bevises eller modbevises". At en formel φ kan modbevises betyder her, at negationen af φ, skrevet $\neg\varphi$, kan bevises.

Begrebet "sand" kan relateres hertil, hvis det i en konkret sammenhæng antages, at enhver formel enten er sand eller falsk. Antag nemlig, at dette er tilfældet og antag yderligere, at der findes en formel φ for hvilken hverken φ eller $\neg\varphi$ kan bevises. Der gælder da, at enten φ eller $\neg\varphi$ er sand, og da ingen af dem kan bevises, er der altså en sand formel der ikke kan bevises.

Det andet centrale begreb i Gödels sætning: "bevis", hvilket som antydet ovenfor skal forstås som "formelt bevis", er en term der vil blive præciseret senere. Men det skal på dette sted fremhæves, at det vi forstår ved bevis i sædvanlig matematisk praksis, ikke direkte kan identificeres med formelt bevis. Formelle beviser er sekvenser af formler formuleret i et symbolsprog og konstrueret i henhold til rent formelt givne regler.

I den matematiske folklore fortolkes Gödels sætning ofte derhen, at den endegyldigt beviser en begrænsning i mulighederne for erkendelse af sandhed gennem logiske argumenter. I en endnu mere radikal fortolkning siges Gödels sætning at vise grænserne for den rationelle tankegangs rækkevidde. At det kun er en mindre del af virkeligheden, man kan erkende ved logisk tænkning, er givetvis korrekt. Det har bare ikke noget at gøre med Gödels sætning.

Gödels sætning er en matematisk sætning om grænser for det opnåelige ved formalisering af matematik. Det er ikke en sætning "om matematikken" som den udøves i praksis, men en sætning om formelle systemer, hvor alle objekter og begreber, herunder begrebet "bevis", udtrykkes i et rent formelt symbolsprog, der ikke tillægger de indgående størrelser mening eller betydning udover hvad der ligger i de formelle regler. Sådanne formelle systemer kan under passende forudsætninger anskues som en mulig aksiomatisk, formel beskrivelse af velkendte matematiske objekter, som f.eks. de naturlige tal. Men man må ikke uden videre identificere en sådan formel beskrivelse og behandling med den sædvanlige matematiske fremgangsmåde, som denne har etableret sig gennem fagets udvikling og fremstår i den aktuelle forsknings- og undervisningsmæssige praksis. Det vil under alle omstændigheder aldrig være muligt, at give et matematisk bevis for at visse matematiske sætninger ikke kan bevises indenfor sædvanlig matematisk praksis, da "sædvanlig matematisk praksis" ikke er et begreb vi kan fastfryse og præcisere matematisk.

Gödels sætning, der stammer fra 1930, er et helt fundamentalt resultat inden for både den matematiske grundlagsforskning, den matematiske logik og blandt andet de dele af datalogien, der beskæftiger sig med logikprogrammering og kunstig intelligens. Sætningen har en status og en fascinationskraft, der gør den kendt af alle matematikere. Men det er de færreste der har set et bevis for sætningen og har stiftet bekendtskab med baggrunden for, at sætningen er rigtig.

Det er dette vi vil søge at råde bod på med denne artikel. Formålet er at give en kort, let læselig fremstilling af hovedlinierne i beviset for Gödels sætning. Dette gøres ved at undlade en detaljeret behandling af formaliseringen, men i stedet tage udgangspunkt i en overordnet beskrivelse af et formelt system. For et sådant vil vi opstille en række antagelse, der er umiddelbart naturlige for et system, der skal kunne afspejle en realistisk matematisk struktur.

Herefter bevises det, at hvis et formelt system af den behandlede art er konsistent, og hvis yderligere enhver formel kan bevises eller modbevises, da ledes man uundgåeligt frem til et paradoks (en modstrid). Dette viser, at der for et konsistent system nødvendigvis må findes en formel, der hverken kan bevises eller modbevises.

2 Formelle systemer

Et **formelt system** er karakteriseret ved størrelserne **alfabet**, **formler**, **variable**, **konstanter**, **aksiomer**, **slutningsregler** og **beviser**.

Et **alfabet** er en endelig mængde af symboler. **Variable**, **konstanter** og **formler** er alle endelige tegnstrenge over alfabetet. Da alfabetet er endeligt kan der kun være et nummerabelt antal variable, konstanter og formler. Til at betegne et formelt systems variable benytter vi sædvanligvis

$$x, x_1, x_2, x_3, \ldots$$

og til at betegne konstanter benyttes

$$k_1, k_2, k_3, \ldots.$$

Hvis φ er en formel hvori variablene $x_1, \ldots, x_n$ optræder som delstrenge, skriver vi også $\varphi(x_1, \ldots, x_n)$ for denne formel. For ethvert valg af konstanter eller variable $t_1, \ldots, t_n$ lader vi da $\varphi(t_1, \ldots, t_n)$ være strengen, der opstår ved at substituere enhver forekomst af delstrengen x_i i φ med t_i. Hvis altså eksempelvis $\varphi(x)$ betegner formlen $x + 3 = 2 * x$, da betegner $\varphi(3)$ formlen $3 + 3 = 2 * 3$.

Vi vil antage, at alle betragtede formelle systemer indeholder symbolet $\neg$ i deres alfabet, og at $\neg\varphi$ er en formel i systemet når blot φ er det. $\neg\varphi$ kaldes som nævnt for **negationen** af φ.

Aksiomerne er en på forhånd udvalgt mængde af formler. En **slutningsregel** er et princip, der på rent formel, mekanisk måde angiver, hvordan man fra en eller flere formler kan udlede en ny formel.

Et formelt bevis, eller blot et **bevis**, er en endelig sekvens af formler. Et bevis starter med et eller flere **aksiomer**, og enhver formel skal (hvis den ikke selv er et aksiom) fremkomme af de foregående formler i sekvensen ved brug af en slutningsregel.

Vi siger, at en formel φ kan **bevises**, hvis φ optræder som sidste formel i et bevis. Vi siger, at φ kan **modbevises**, hvis negationen $\neg\varphi$ kan bevises.

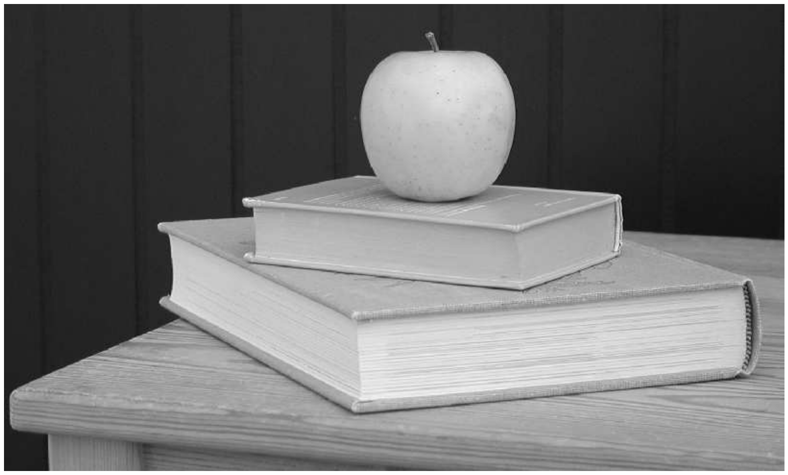
Vi illustrerer meningen med de ovennævnte størrelser ved et eksempel.

Eksempel 1 (Det lille overfyldte bord). I dette eksempel vil vi konstruere et formelt system som beskriver situationen på Figur 1. **Alfabetet** består i dette tilfælde af bogstaverne i det danske alfabet, cifrene `0`–`9` samt tegnene

$$) \quad (\quad , \quad \wedge \quad \rightarrow \tag{1}$$

Konstanterne er

`lillebog`, `storebog`, `æble`



Figur 1: Det lille overfyldte bord

og **variablene** er

$\texttt{x1}, \texttt{x2}, \texttt{x3}, \ldots$

Formlerne er af typen

på(t_1,t_2)

og

over(t_1,t_2)

hvor t_1 og t_2 er vilkårlige konstanter eller variable. Til formlerne regner vi desuden alle sædvanlige kombinationer af formler af ovenstående type ved brug af symbolerne (1).

Aksiomerne er:

(A1) på(storebog,lillebog)

(A2) på(lillebog,æble)

(A3) på(x1,x2) $\rightarrow$ over(x1,x2)

(A4) over(x1,x2) $\wedge$ over(x2,x3) $\rightarrow$ over(x1,x3)

Slutningsreglerne er:

(SR1) udfra $\varphi \rightarrow \psi$ og φ sluttes ψ.

(SR2) udfra φ og ψ sluttes $\varphi \wedge \psi$.

(SR3) udfra $\varphi(x_1, \ldots, x_n)$ sluttes $\varphi(k_1, \ldots, k_n)$ for et vilkårligt valg af konstanter $k_1, \ldots, k_n$.

Vi vil herefter bevise formlen

`over(storebog,æble)`.

Et bevis kan se ud som følger, hvor vi har opstillet de enkelte formler i sekvensen i en nummereret liste. Til højre for hver formel er det angivet, hvordan den pågældende formel er fremkommet.

1.	`på(storebog,lillebog)`	aksiom (A1)
2.	`på(lillebog,æble)`	aksiom (A2)
3.	`på(x1,x2)` $\rightarrow$ `over(x1,x2)`	aksiom (A3)
4.	`på(storebog,lillebog)` $\rightarrow$ `over(storebog,lillebog)`	(SR3) på 3.
5.	`over(storebog,lillebog)`	(SR1) på 1.,4.
6.	`på(lillebog,æble)` $\rightarrow$ `over(lillebog,æble)`	(SR3) på 3.
7.	`over(lillebog,æble)`	(SR1) på 2.,6.
8.	`over(storebog,lillebog)` $\wedge$ `over(lillebog,æble)`	(SR2) på 5.,7.
9.	`over(x,y)` $\wedge$ `over(y,z)` $\rightarrow$ `over(x,z)`	aksiom (A4)
10.	`over(storebog,lillebog)` $\wedge$ `over(lillebog,æble)` $\rightarrow$ `over(storebog,æble)`	(SR3) på 9.
11.	`over(storebog,æble)`	(SR1) på 8.,10.

3 Repræsenterbarhed

I et formelt system har de indgående tegn, formler og symboler ikke i sig selv nogen bestemt mening eller betydning. Man kan naturligvis—som i eksemplet med det overfyldte bord—vælge betegnelserne ud fra en tilsigtet mening, men når det drejer sig om hvad "der gælder" i det formelle system, er det udelukkende aksiomerne og slutningsreglerne der kan benyttes.

I et formelt system er der ikke indbygget noget begreb "sandhed". Det eneste der angiver, hvad et formelt system kan udtrykke, eller hvad "der gælder", er de formler der kan bevises.

Et formelt system, der kun indeholder få bevisbare formler, kan ikke udtrykke synderligt meget, og kan dermed ikke være en realistisk formalisering af interessante matematiske strukturer.

Hvis hensigten f.eks. er at formalisere behandlingen af de naturlige tal, skal det formelle system være i stand til at udtrykke velkendte egenskaber for tal og udtrykke, om et givet objekt (tal) har denne egenskab eller ej. Lad os betragte egenskaben "primtal" og lad os antage, at $0, 1, 2, \ldots$ er konstanterne i det formelle system. At det formelle system er i stand til at "tale om" egenskaben at være et primtal, skal forstås på den måde, at der findes en formel $\varphi(x)$, så der for ethvert naturligt tal i gælder

$$i \text{ er et primtal} \Leftrightarrow \varphi(i) \text{ kan bevises.}$$

Vi kan da sige, at formlen $\varphi(i)$ udtrykker egenskaben, at "tallet i er et primtal".

Vi benytter glosen **repræsenterbarhed** i forbindelse med præciseringen af, hvad et formelt system kan udtrykke eller "tale om". Lad R være en n-plads relation, dvs. en mængde af n-tupler. R siges at være **repræsenterbar** i et formelt system, hvis der eksisterer en formel $\varphi(x_1, x_2, \ldots, x_n)$ i systemet, så følgende er opfyldt for alle valg af konstanter $k_1, \ldots, k_n$ fra systemet

$$(k_1, \ldots, k_n) \in R \Leftrightarrow \varphi(k_1, \ldots, k_n) \text{ kan bevises.}$$

Vi illustrerer igen meningen ved et eksempel.

Eksempel 2 (Det lille overfyldte bord). Lad os vende tilbage til eksemplet fra før. Vi fik her bevist formlen `over(storebog,æble)`, idet denne optrådte som den sidste formel i det givne bevis. Ser vi kun på de første 7 linier af beviset, får vi et bevis for formlen `over(lillebog,æble)`, idet denne er den sidste formel i det afkortede bevis. Tilsvarende udgør de første 5 linier af beviset et bevis for formlen `over(storebog,lillebog)`.

Vi har altså fået bevist følgende tre formler

`over(storebog,lillebog)`, `over(lillebog,æble)`, `over(storebog,æble)`.

Det er ikke vanskeligt at vise, at for ethvert andet valg af konstanter k_1, k_2 kan `over(`k_1`,`k_2`)` *ikke* bevises. Det betyder, at formlen `over(x1,x2)` repræsenterer relationen som indeholder følgende par

`(storebog,lillebog)`, `(lillebog,æble)`, `(storebog,æble)`.

Af figuren i foregående eksempel ses det, at formlen `over(x1,x2)` netop repræsenterer relationen "på bordet ligger `x2` over `x1`". For ethvert valg af konstanter k_1, k_2 udtrykker formlen `over(`k_1`,`k_2`)` således, at "på bordet ligger k_2 over k_1".

I det følgende vil vi for nemheds skyld udelukkende betragte formelle systemer hvori konstanterne er $0, 1, 2, \ldots$. Det betyder, at alle de relationer, som de betragtede formelle systemer kan repræsentere, er relationer over de naturlige tal.

4 Grellings paradoks

Beviset for Gödels sætning bygger på et af de klassiske paradokser, nemlig Grellings paradoks, der fremkommer som følger:

Et sprogligt udtryk kaldes **heterologisk**, hvis det *ikke* har den egenskab som det selv udtrykker. Eksempelvis er begrebet "æble" heterologisk, da begrebet "æble" jo ikke selv er et æble. Derimod er begrebet "et abstrakt begreb" ikke heterologisk, da "et abstrakt begreb" selv er et abstrakt begreb.

Vi stiller nu det spørgsmål, om det sproglige udtryk "heterologisk" er heterologisk eller ej.

Antag først, at "heterologisk" selv er heterologisk. Det betyder, at "heterologisk" ikke har den egenskab, som det selv betegner, hvorfor "heterologisk" ikke er heterologisk. Dette er en modstrid.

Antag dernæst, at "heterologisk" ikke er heterologisk. Det betyder, at "heterologisk" har den egenskab, som det selv betegner, altså at det er heterologisk. Dette er igen en modstrid.

Uanset om vi antager, at "heterologisk" selv er heterologisk eller ej, ledes vi til en modstrid. Vanskelighederne opstår når begrebet heterologisk så at sige bringes til at udtale sig om sig selv. Går vi ud fra, at ethvert klart forståeligt spørgsmål af den betragtede art kan besvares med "ja" eller "nej", men ikke med begge dele, kommer vi ud i en åbenlys selvmodsigelse.

5 Gödels sætning

I indledningen formulerede vi Gödels sætning som "Der findes formler, der hverken kan bevises eller modbevises". Dette er som nævnt et resultat der angår formelle beviser i formelle systemer. Men det gælder kun for formelle systemer af en vis *styrke*. Et formelt systems styrke måles i hvilke relationer, der er repræsenterbare i systemet, det vil sige, hvilke begreber systemet er i stand til at "tale om".

Et formelt system kaldes **konsistent**, hvis ingen formel φ både kan bevises og modbevises. Konsistens er naturligvis en meget central egenskab for et formelt system. Hvis vi eksempelvis både kan bevise og modbevise, at 7 er et primtal, er vi ikke blevet meget klogere med hensyn til egenskaberne for tallet 7. Et formelt system kaldes **fuldstændigt**, hvis enhver formel enten kan bevises eller modbevises. **Ufuldstændighed** er da det samme som eksistensen af en formel, der hverken kan bevises eller modbevises.

Vi vil vise at ethvert konsistent formelt system af tilstrækkelig styrke er ufuldstændigt. Den grundlæggende idé i beviset for dette er, at hvis vi antager at et formelt system både er fuldstændigt og af tilstrækkelig styrke, da kan vi indenfor systemet udtrykke en formaliseret variant af Grellings paradoks. Det er ikke åbenlyst, hvordan noget sådant skal gøres, men en kort fremgangsmåde er følgende:

Antag, at vi er givet et formelt system, og lad x betegne en af systemets variable. Da alfabetet for systemet er endeligt, er der kun et nummerabelt antal formler som indeholder variablen x, og disse kan vi derfor nummerere

$$\varphi_0(x), \varphi_1(x), \varphi_2(x), \ldots$$

Vi kalder en formel $\varphi_n(x)$ **heterologisk**, hvis $\neg\varphi_n(n)$ kan bevises. Hvis $\varphi_n(x)$ er heterologisk, kalder vi n for et **heterologisk tal**.

Det formelle system siges nu at være **af tilstrækkelig styrke**, hvis mængden af heterologiske tal er repræsenterbar i det.

Denne egenskab er helt afgørende, og med den som forudsætning er vi nu klar til at bevise vores version af Gödels sætning, som opstår ved en "formalisering" af Grellings paradoks.

Sætning 3 (Gödels sætning). *I ethvert konsistent formelt system af tilstrækkelig styrke eksisterer der en formel, som hverken kan bevises eller modbevises.*

Bevis. Antag, med henblik på at opnå en modstrid, at vi har et formelt system som både er konsistent, fuldstændigt og af tilstrækkelig styrke. Da systemet er af tilstrækkelig styrke, indeholder det en formel $\varphi_h(x)$, som repræsenterer mængden af heterologiske tal. Det betyder, at der for alle naturlige tal n gælder

$$n \text{ er et heterologisk tal} \Leftrightarrow \varphi_h(n) \text{ kan bevises.}$$

Da der desuden gælder

$$n \text{ er et heterologisk tal} \Leftrightarrow \neg\varphi_n(n) \text{ kan bevises}$$

får vi i alt

$$\varphi_h(n) \text{ kan bevises} \Leftrightarrow \neg\varphi_n(n) \text{ kan bevises.}$$

Dette gælder for alle n. Lader vi specielt $n = h$ fås

$$\varphi_h(h) \text{ kan bevises} \Leftrightarrow \neg\varphi_h(h) \text{ kan bevises.}$$

Vi har altså konstrueret en formel $\varphi_h(h)$, som kan bevises netop hvis den kan modbevises. Da systemet er fuldstændigt, vil $\varphi_h(h)$ enten kunne bevises eller modbevises. Men da må den *både* kunne bevises og modbevises. Dette er en modstrid, da systemet er antaget at være konsistent.

Hvis et formelt system altså både er konsistent og af tilstrækkelig styrke, kan det ikke samtidig være fuldstændigt. Det betyder, at der i ethvert konsistent system af tilstrækkelig styrke eksisterer en formel, som hverken kan bevises eller modbevises. □

At ovenstående bevis reelt set blot er en variant af Grellings paradoks, kan direkte ses som følger:

Udtrykket eller begrebet "heterologisk" i det sproglige paradoks svarer i det formelle system til formlen $\varphi_h(x)$, som repræsenterer mængden af heterologiske tal. At spørge om begrebet "heterologisk" selv er heterologisk eller ej, svarer da til at spørge om formlen $\varphi_h(x)$ er heterologisk eller ej. Men dette svarer igen til at spørge om $\neg\varphi_h(h)$ kan bevises eller ej. Forsøget på at afklare dette spørgsmål leder til en modstrid på nøjagtig samme måde som i den sproglige formulering.

I den sproglige formulering giver modstriden anledning til et paradoks. I den formelle variant giver modstriden et bevis for, at på de givne præmisser er konsistens og fuldstændighed i modbyrdes konflikt: Hvis den ene af disse to betingelser er opfyldt, er det en logisk umulighed, at den anden betingelse også er opfyldt.

Det er værd at bemærke, at ovenstående bevis faktisk er konstruktivt, i den forstand at vi for ethvert konsistent system af tilstrækkelig styrke får udpeget en konkret formel, som hverken kan bevises eller modbevises. I ovenstående bevis viste vi, at $\varphi_h(h)$ er en formel, der kan bevises netop hvis den kan modbevises. Til dette benyttede vi hverken konsistens eller fuldstændighed af det betragtede system. Hvis vi yderligere antager, at det formelle system er konsistent, kan $\varphi_h(h)$ ikke både bevises og modbevises. Men så er eneste mulighed, at $\varphi_h(h)$ hverken kan bevises eller modbevises. Givet et konsistent system af tilstrækkelig styrke får vi altså udpeget $\varphi_h(h)$ som en konkret formel, der hverken kan bevises eller modbevises.

6 Tilstrækkelig styrke

Vi har nu set, at ethvert konsistent system, hvori mængden af heterologiske tal er repræsenterbar, må være ufuldstændigt. Så længe vi ikke har grund til at tro, at formelle systemer i almindelighed skulle være i stand til at repræsentere mængden af heterologiske tal, er resultatet ikke særlig afgørende. Problemet er blot, at det nu viser sig, at mængden af heterologiske tal faktisk *er* repræsenterbar i alle formelle systemer, blot disse formaliserer ikke-trivielle dele af matematikken. Det betyder, at det for alle sådanne systemer må gælde, at de enten er inkonsistente, eller at der eksisterer en formel, som hverken kan bevises eller modbevises. Begge alternativer er meget alvorlige begrænsninger for et formelt system.

I dette afsnit ser vi lidt nærmere på hvad det er der gør, at mængden af heterologiske tal kan blive repræsenterbar i et formelt system.

Lad der være givet et formelt system. Lad os betragte mængden af tegnstrenge over systemets alfabet. Vi kan danne sekvenser af sådanne tegnstrenge. Blandt disse sekvenser finder vi specielt alle systemets formelle beviser, da et formelt bevis er en endelig sekvens af formler, og enhver formel er en tegnstreng. Da alfabetet er endeligt, er mængden af endelige sekvenser nummerabel. Det betyder, at vi kan karakterisere samtlige sekvenser ved numrene $0, 1, 2, \ldots$.

Hørende til ethvert formelt system kan vi nu definere en 3-plads relation B over de naturlige tal ved følgende betingelse

$$(p, q, r) \in B \Leftrightarrow \text{sekvens nr. } p \text{ er et formelt bevis, og sidste element i sekvensen er formlen } \neg\varphi_q(r).$$

Der gælder altså, at et trippel (p, q, r) er med i B hvis og kun hvis sekvens nr. p er et bevis for formlen $\neg\varphi_q(r)$.

I de fleste realistiske formelle systemer, som formaliserer ikke-trivielle dele af matematikken, vil relationen B for systemet være repræsenterbar i systemet selv. Lad os give et kort heuristisk argument for dette.

Vi har antaget at beviser generelt er sekvenser af formler opfyldende simple, mekaniske regler. Derfor er det en relativt simpel affære, at afgøre om en sekvens udgør et formelt bevis eller ej: vi skal blot checke om strengene i sekvensen er formler, der overholder slutningsreglerne for systemet (eller er aksiomer). Givet en sekvens, som er et formelt bevis, er det naturligvis desuden let at checke om det sidste element i sekvensen er en bestemt formel $\neg\varphi_q(r)$. Det betyder, at det er relativt simpelt at afgøre om et trippel (p, q, r) tilhører relationen B eller ej. Derfor vil allerede formelle systemer, som kun er i stand til at repræsentere simple relationer over de naturlige tal, kunne repræsentere en relation som B.

Ofte indeholder formelle systemer symbolet $\exists$ i deres alfabet, som bruges når systemet skal udtrykke eksistenspåstande. Givet en formel $\varphi(x)$ og en variabel x er $\exists x\varphi(x)$ da en formel som udtrykker: “Der eksisterer et x, så $\varphi(x)$ er opfyldt”. Hvis der for enhver formel $\varphi(x)$ i systemet gælder

$$\exists x\varphi(x) \text{ kan bevises} \Leftrightarrow \text{der eksisterer et } n \in \mathbb{N} \text{ så } \varphi(n) \text{ kan bevises}$$

kalder vi $\exists$ for en **stærk eksistenskvantor** i det formelle system. Hvis et formelt system på korrekt vis skal kunne udtrykke eksistenspåstande vedrørende naturlige tal, må det naturligvis indeholde en sådan stærk eksistenskvantor.

Både repræsenterbarheden af relationen B og tilstedeværelsen af en stærk eksistenskvantor synes altså at være rimelige krav at stille til et formelt system. Men, som nedenstående lemma viser, vil ethvert formelt system opfyldende disse to krav være af tilstrækkelig styrke, det vil sige, være af den type som Gödels sætning viser ikke både kan være konsistente og fuldstændige.

Lemma 4. *Mængden af heterologiske tal er repræsenterbar i ethvert formelt system, der opfylder følgende to betingelser:*

1. *Systemet indeholder en stærk eksistenskvantor.*
2. *Relationen B er repræsenterbar i systemet.*

Bevis. Antag relationen B er repræsenteret af en formel $\varphi_b(x_1, x_2, x_3)$ i systemet. Lad $\psi(x)$ være formlen

$$\exists x_1\varphi_b(x_1, x, x).$$

Vi vil vise at $\psi(x)$ repræsenterer mængden af heterologiske tal. Vi skal altså vise, at der gælder

$$n \text{ er et heterologisk tal} \Leftrightarrow \psi(n) \text{ kan bevises.}$$

Dette følger af følgende række af ækvivalenser:

$$\begin{aligned}
&\psi(n) \text{ kan bevises} \\
\Leftrightarrow\ &\exists x_1 \varphi_b(x_1, n, n) \text{ kan bevises} \\
\Leftrightarrow\ &\text{der eksisterer } m \in \mathbb{N} \text{ så } \varphi_b(m, n, n) \text{ kan bevises} \\
\Leftrightarrow\ &\text{der eksisterer } m \in \mathbb{N} \text{ så } (m, n, n) \in B \\
\Leftrightarrow\ &\text{der eksisterer } m \in \mathbb{N} \text{ så sekvens nr. } m \text{ er bevis for formlen } \neg\varphi_n(n) \\
\Leftrightarrow\ &\neg\varphi_n(n) \text{ kan bevises} \\
\Leftrightarrow\ &n \text{ er et heterologisk tal.}
\end{aligned}$$

Hermed er beviset fuldført. □

7 Det oprindelige bevis

Det vi her kalder for Gödels sætning kaldes sædvanligvis for "Gödels første ufuldstændighedsteorem". Resultatet er det ene af to hovedresultater i Gödels berømte artikel fra 1931 ([5]). I dette afsnit vil vi prøve at skitsere gangen i det oprindelige bevis for det første ufuldstændighedsteorem, og sammenholde det med vores version af resultatet. Men først vil vi kort skitsere hvad det andet ufuldstændighedsteorem går ud på.

Lad der være givet et formelt system af tilstrækkelig styrke. I beviset for Lemma 4 så vi, at det er muligt at konstruere en formel $\psi(x)$, sådan at $\psi(n)$ kan bevises netop når $\neg\varphi_n(n)$ kan bevises. Vi kan derfor sige, at $\psi(n)$ er en formel, som udtrykker: "formlen $\neg\varphi_n(n)$ kan bevises". Under få yderligere betingelser for systemet kan vi på tilsvarende måde konstruere en formel som udtrykker: "der eksisterer ikke en formel som både kan bevises og modbevises". Denne formel kaldes for systemets **konsistenspåstand**, da den udtrykker netop det, som er betingelsen for systemets konsistens. Gödels andet ufuldstændighedsteorem siger nu følgende: "Hvis et formelt system af tilstrækkelig styrke er konsistent, da kan konsistenspåstanden ikke bevises i systemet". Med andre ord, hvis et formelt system er konsistent, kan systemet ikke bevise sin egen konsistens.

Vi vender nu tilbage til det første ufuldstændighedsteorem, og skitserer de enkelte skridt i Gödels oprindelige bevis for resultatet:

1. Introduktion af systemet P

Gödels artikel begynder med introduktionen af et konkret formelt system P, som han igennem resten af artiklen tager udgangspunkt i. Systemet P er et forsøg på en formalisering af en del af matematikken indeholdende teorien for de naturlige tal.

2. Introduktion af Gödel-nummerering

Der introduceres herefter en nummerering af alle tegnstrenge over alfabetet for P. Dette svarer til den ordning af formlerne indeholdende variablen x, som vi introducerede i afsnit 5. Nummereringer af denne type er senere kommet til at hedde **Gödel-nummereringer**. En formel φ's nummer i en sådan nummerering kaldes da for formlens **Gödel-nummer**.

3. Introduktion af rekursive relationer

Nu introduceres i artiklen et vigtigt begreb, som ikke forekommer i vores version af Gödels sætning. Dette er begrebet "rekursivitet". Gödel definerer en klasse af **rekursive relationer**. Disse er relationer over de naturlige tal. I moderne sprog kan de rekursive relationer karakteriseres som de, der kan beregnes af en computer, hvis vi forestiller os denne udstyret med ubegrænset lagerplads.

4. Bevis for at relationen B for ethvert realistisk system er rekursiv

Gödel betragter i sin artikel udelukkende formelle systemer, som kan fremkomme af systemet P ved at tilføje en række nye aksiomer. Vi kalder et sådan system for **realistisk**, hvis mængden af Gödel-numre for systemets aksiomer er rekursiv.

Gödel viser, at for ethvert realistisk formelt system er systemets relation B, som vi introducerede i Afsnit 6, rekursiv. At vise netop dette udgør langt over halvdelen af Gödels oprindelige bevis for det første ufuldstændighedsteorem. Det er et meget langt og teknisk bevis.

Der er dog intet overraskende i resultatet. Vi har antaget, at mængden af Gödel-numre for aksiomerne er en rekursiv mængde. Det betyder, at en computer vil være stand til at afgøre om en tegnstreng er et aksiom eller ej. Da samtidigt slutningsreglerne antages at være simple formelle, mekaniske regler, må computeren også være i stand til at afgøre om en slutningsregel er anvendt korrekt eller ej. Men da bliver computeren i stand til at afgøre om en sekvens er et formelt bevis eller ej. Da det desuden for en computer er en let sag at checke om den sidste formel i et bevis er en bestemt på forhånd given formel, bliver computeren således i stand til at beregne relationen B. Dette viser, at B må være rekursiv.

5. Bevis for at enhver rekursiv relation er repræsenterbar i P

Gödel skitserer et bevis for, at enhver rekursiv relation er repræsenterbar i P. Heraf følger trivielt, at enhver rekursiv relation må være repræsenterbar i ethvert realistisk system, da disse alle er udvidelser af P. Dette er også et ret teknisk resultat at bevise.

Lad der være givet et vilkårligt realistisk formelt system. Da enhver rekursiv relation er repræsenterbar i systemet, og da vi ved fra punkt 4 at relationen B er rekursiv, må specielt B være repræsenterbar. For ethvert realistisk system gælder det altså, at relationen B for systemet er repræsenterbar.

6. Bevis for det første ufuldstændighedsteorem

Gödel definerer begrebet **ω-konsistens**, som er et lidt stærkere krav end almindelig konsistens. Gödel kan nu formulere og bevise sit berømte ufuldstændighedsresultat:

> I ethvert realistisk, ω-konsistent system eksisterer der en formel, som hverken kan bevises eller modbevises.[1]

Dette resultat kan vises indenfor vores ramme på følgende måde. Antag, at der eksisterer et realistisk formelt system, som både er ω-konsistent og fuldstændigt. Da systemet er realistisk, er relation B repræsenterbar i det (se punkt 5 ovenfor). Det kan desuden let vises, at ethvert fuldstændigt, ω-konsistent system indeholder en stærk eksistenskvantor. Dette giver os, ifølge Lemma 4, at systemet er af tilstrækkelig styrke. Sætning 3 viser derfor, at systemet ikke kan være konsistent. Men da ω-konsistens er en stærkere betingelse en almindelig konsistens, og da vi har antaget ω-konsistens, giver dette en modstrid. Det er derfor bevist, at der ikke kan eksistere et realistisk system, som både er ω-konsistent og fuldstændigt.

Gödel benytter grundlæggende den samme idé i sit bevis som i vores. Gödels bevis er dog ikke gennemført i så streng parallel til paradokserne som vores. Han nævner blot slægtskabet med paradokserne i artiklens indledning. Han sammenligner selv sit bevis med Richards paradoks, som er en variant af Grellings paradoks. Han bemærker også, at valget af paradoks ikke er afgørende. Et hvilket som helst "semantisk" paradoks vil kunne benyttes i et bevis for ufuldstændigheden.

Ethvert realistisk, ω-konsistent system er af tilstrækkelig styrke, men ikke omvendt. Vores version af Gödels sætning omfatter derfor flere systemer end Gödels oprindelige sætning. Til gengæld har Gödels oprindelige sætning to store fordele fremfor vores version. Den første er, at den oprindelige sætning får udpeget et *konkret* formelt system P, for hvilket ufuldstændighedsresultatet gælder. Den anden fordel er, at det bliver vist, at vi ikke på nogen realistisk måde kan tilføje noget til P, som kan gøre det fuldstændigt. For alle udvidelser af P, som blot er realistiske, gælder samme ufuldstændighedsresultat. Det er disse to ting som gør, at Gödels oprindelige bevis er så meget længere og tungere end vores.

Lad os afslutte dette afsnit med at give den oprindelige formulering af Gödels sætning:

[1] Gödel bruger dog ikke selv udtrykket "realistisk".

Satz VI: Zu jeder ω-widerspruchsfreien rekursiven Klasse κ von *Formeln* gibt es rekursive *Klassenzeichen* r, so daß weder v Gen r noch Neg(v Gen r) zu Flg(κ) gehört (wobei v die *freie Variable* aus r ist).

8 Hvor kommer ufuldstændigheden fra?

Vi vil i dette afsluttende afsnit prøve at få en forståelse for, hvad det er der gør, at formelle systemer nødvendigvis er underlagt den begrænsning, som Gödels sætning demonstrerer.

Lad os først tage et nærmere kig på den konkrete formel, som i Gödels sætning bliver vist hverken at kunne bevises eller modbevises. Givet et konsistent formelt system af tilstrækkelig styrke har vi set, i beviset for Sætning 3, at dette er en formel $\varphi_h(h)$, hvor $\varphi_h(x)$ repræsenterer mængden af heterologiske tal. Formlen $\varphi_h(h)$ udtrykker altså "h er et heterologisk tal", hvilket er det samme som at sige, at "formlen $\varphi_h(x)$ er heterologisk". $\varphi_h(x)$ er heterologisk netop hvis $\varphi_h(h)$ kan modbevises. Formlen $\varphi_h(h)$ kan derfor også opfattes som udtrykkende "formlen $\varphi_h(h)$ kan modbevises". $\varphi_h(h)$ er således en formel der udtrykker en egenskab *ved sig selv*. Den udtrykker egenskaben, at den selv kan modbevises. Med andre ord udtrykker den: "denne formel kan modbevises". Det er i denne forstand en **selv-refererende** formel. Netop heri ligger en nøgle til forståelse af kernen i Gödels sætning.

Ethvert formelt system kan ses som et forsøg på at give en **model** af en del af virkeligheden. Det så vi tidligere et eksempel på med det lille overfyldte bord. Et formelt system er også en del af virkeligheden, og kan som sådan være en del af det, der forsøges modelleret.

Formelle systemer, der modellerer en del af virkeligheden som ikke omfatter systemerne selv, er generelt uproblematiske. Problemerne opstår først når en model forsøger at modellere en virkelighed indeholdende modellen selv. Dette er f.eks. tilfældet hvis vi ønsker at skabe en model for *hele* virkeligheden. En sådan må nødvendigvis også modellere modellen selv, da modellen er en del af virkeligheden.

Lad os kalde formelle systemer, som direkte eller indirekte modellerer sig selv, for **refleksive systemer**. I et sådant system eksisterer der altså formler som udtrykker egenskaber ved systemet selv. Et eksempel på dette er, når vi i et formelt system kan repræsentere relationen B ved en formel $\varphi_b(x_1, x_2, x_3)$. Enhver formel af typen $\varphi_b(p, q, r)$ udtrykker da en egenskab ved systemet selv, nemlig egenskaben at sekvens nr. p er et bevis for formlen $\neg\varphi_q(r)$.

Hvis et refleksivt system er tilstrækkelig stærkt, vil der, udover at være formler som udtrykker egenskaber ved andre formler i systemet, også være formler som udelukkende udtrykker egenskaber *ved sig selv*. Dette er altså formler af en vis patologisk natur, da disse ikke modellerer en ydre virkelighed, men kun sig selv. Visse af disse formler er i stand til at udtrykke egenskaber ved sig selv, som er modsatte af dem de egentlig har.

Dette gør, at det ville være kontradiktorisk hvis de kunne bevises eller modbevises. De forårsager således ufuldstændigheden af det system, de er en del af. Eksempler på sætninger i vores naturlige sprog som udtrykker modsatte egenskaber ved sig selv er

> “denne sætning er falsk”

og

> “begrebet heterologisk er heterologisk”.

Tilsvarende har vi i formelle systemer en formel som f.eks. $\varphi_h(h)$, som udtrykker “denne formel kan modbevises”. Hvis den kunne bevises, måtte det den udtrykte være sandt. Men den udtrykker netop, at den kan modbevises, og dette giver en modstrid, hvis vi antager at vi befinder os indenfor et konsistent system. Hvis vi modsat antager, at den kan modbevises, så må det den udtrykker være falsk. Men da får vi, at den *ikke* kan modbevises, hvilket igen er en modstrid. En sådan formel kan således hverken bevises eller modbevises.

Formelle systemer, som er tilstrækkeligt stærke og som modellerer en virkelighed indeholdende sig selv, vil altså nødvendigvis være ufuldstændige pga. de patologiske selv-refererende sætningers tilstedeværelse. Dette viser sig også at gælde for formelle systemer, der formaliserer ikke-trivielle dele af matematikken. Grunden til dette er, at et formelt system er et relativt simpelt matematisk objekt: det består essentielt kun af en mængde tegnstrenge (aksiomer) og en række formelle, mekaniske regler (slutningsregler), der giver mulighed for at udlede nye tegnstrenge fra gamle. Det at et formelt system er et relativt simpelt matematisk objekt betyder, at allerede i forholdsvist beskedne dele af matematikken vil vi kunne beskrive formelle systemer og deres egenskaber. Det betyder så igen, at formelle systemer, som formaliserer disse dele af matematikken, nødvendigvis kommer til at modellere den del af matematikken, der omhandler formelle systemer. På denne måde kommer disse systemer uundgåeligt til at være refleksive. Hermed følger de selv-refererende sætninger og den ufuldstændighed de giver anledning til.

Man skal være forsigtig med at sige, at Gödels sætning endegyldigt viser grænserne for matematisk erkendelse eller lignende. Som vi har set, hænger Gödels sætning uløseligt sammen med selv-reference og problemerne i at have modeller, som modellerer sig selv. Gödels sætning handler måske mest af alt om det komplekse forhold som består imellem model og modelleret virkelighed, når disse to ikke er komplet adskillelige. Der er i forbindelse med dette et væld af uløste problemer, som studeres indenfor mange forskellige emneområder. Det gælder f.eks. indenfor aksiomatisk mængdelære i matematikken og indenfor kunstig intelligens i datalogien. Netop det, at man endnu ikke har en speciel god overordnet forståelse for disse forhold bør gøre, at man skal være forsigtig med at drage for vidtrækkende konklusioner af de resultater, man indtil videre har. Der er helt klart fundamentale erkendelser omkring det generelle forhold imellem "noget, der repræsenterer" og "det, det repræsenterer" som endnu venter på at blive afklaret. Før disse fundamentale erkendelser er på plads, får vi næppe den fulde forståelse for rækkeviden og betydningen af Gödels sætning eller den fulde forståelse for paradokser som Grellings paradoks.

Litteratur

[1] Steven J. Bartlett, editor. *Reflexivity—A Source-Book in Self-Reference*. North-Holland, Amsterdam, 1992.

[2] Arthur Charlesworth. A proof of Gödel's theorem in terms of computer programs. *Mathematics Magazine*, 54(3):109–121, 1981.

[3] Solomon Feferman. Penrose's gödelian argument. *Psyche*, 2(7), 1995.

[4] Paul Finsler. Formal proofs and undecidability. 1926. Reprinted in [15].

[5] Kurt Gödel. Über formal unentscheidbare Satze der Principia Mathematica und verwandter Systeme I. *Monatschefte für Mathematik und Physik*, 38:173–198, 1931. Reprinted in [6].

[6] Kurt Gödel. *Collected works. Vol. I.* Oxford University Press, 1986. Publications 1929–1936, Edited and with a preface by Solomon Feferman.

[7] Justus Hartnack. *Erkendelsens grundlag: paradokser indenfor logikkens og matematikkens filosofi*. C. A. Reitzel, 1993.

[8] C. A. R. Hoare and D. C. S. Allison. Incomputability. *ACM Computing Surveys*, 4(3):169–178, 1972.

[9] Douglas R. Hofstadter. *Gödel, Escher, Bach: an eternal golden braid*. Basic Books Inc. Publishers, New York, 1979.

[10] S. C. Kleene. *Introduction to Metamathematics*. North-Holland, 1964.

[11] Ernest Nagel and James R. Newman. *Gödel's proof*. New York University Press, New York, 1958.

[12] W. V. Quine. Paradox. *Scientific American*, 20(4):84–96, 1962. Reprinted in [1].

[13] Jules Richard. The principles of mathematics and the problem of sets. 1905. Reprinted in [15].

[14] Raymond M. Smullyan. *Gödel's incompleteness theorems.* Oxford University Press, 1992.

[15] Jean van Heijenoort. *From Frege to Gödel. A source book in mathematical logic, 1879–1931.* Harvard University Press, 1967.