Kommentar til

Gödel: Über formal unentschiedbare Sätze der Principia Mathematica und verwandter Systeme I, 1931

Denne afhandling af den 24-årige Kurt Gödel er blevet en klassiker. Det er vist den eneste moderne matematiske afhandling som menigmand kender af omtale. Dens resultater er af populærjournalistiken blevet tolket som "matematikkens sammenbrud", - ja, som "den menneskelige tankes fallit".

Gödel viser først at der i enhver matematisk teori, som omfatter aritmetikken, findes udsagn som hverken kan bevises eller modbevises. Hvis et sådant, eller dets negation, skal bevises, må der føjes flere aksiomer til teorien. Der er intet mærkeligt i dette. En smule overraskende er måske nok Gödel's andet resultat: for en teori som omfatter aritmetikken kan et bevis for at den er modsigelsesfri ikke formaliseres indenfor teorien selv. Der skal altså flere aksiomer til for at bevise at teorien er modsigelsesfri end der er i teorien selv. Men selv om man havde et bevis for at teorien er modsigelsesfri, som var formaliseret indenfor teorien selv, ville det ikke uden videre bevise at teorien er modsigelsesfri. Thi *hvis* der er en modsigelse i teorien, kan ethvert udsagn jo bevises (se § ...), og dermed også udsagnet "teorien er modsigelsesfri". Et bevis for at teorien er modsigelsesfri har kun værdi, hvis der kun gøres brug af helt harmløse aksiomer, dvs. aksiomer af endelig karakter. At teorien *kan* bevises at være modsigelsesfri, hvis man gør brug af flere aksiomer end der er i teorien, er Gentzen's bevis et eksempel på (§ ...).

Gödel's afhandling er på 26 sider og har 4 afsnit. Vi skitserer indholdet af hver af disse, og til sidst skitserer vi beviset for *Church´s uafgørbarhedssætning.*

§1 er den citerede indledning. Gödel nævner at det udsagn han vil konstruere, som hverken kan bevises eller modbevises, er inspireret af *Richard's paradoks.* Dette paradoks (Richard, 1905) kan formuleres således: Vi ordner alle sproglige udsagn, hvori der indgår netop eet (variabelt) naturlige tal n, på rad og række (efter et leksikografisk princip). Det k-te udsagn betegnes $E_k(n)$. Vi betragter nu udsagnet "$E_n(n)$ er falsk". Da dette er et sprogligt udsagn, hvori der indgår netop eet (variabelt) naturlige tal n, må det være blandt E_k-erne. Lad os antage at det har nummer q, altså

$$E_q(n) = \text{"}E_n(n) \text{ er falsk"}.$$

Men heraf følger (for n = q) at hvis $E_q(q)$ er sand er $E_q(q)$ falsk, og omvendt. Paradoksaliteten beror på, at i det naturlige sprog findes der ikke et sandhedsbegreb således at *ethvert* udsagn kan tildeles en sandhedsværdi sand/

falsk (se **tredie bog**, § ...).

I **§2** sættes det formale begrebsapparat op, og Gödel beviser at "i en teori som omfatter aritmetikken og som er ω-modsigelsesfri (se nedenfor), findes der et udsagn som hverken kan bevises eller modbevises".

Beviset føres ved at *aritmetisere metamatematikken:* enhver formal term, ethvert formalt udsagn og ethvert bevis tildeles et naturligt tal, herved kan ethvert metamatematisk udsagn omformes til et formalt aritmetisk udsagn.

Gödel giver hvert af tegnene i den formale teori et nummer (som er et naturligt tal). Enhver tegnstreng, altså specielt enhver term og ethvert udsagn, kun nu tildeles et naturligt tal, nemlig $2^{n_1}3^{n_2}5^{n_3}7^{n_4}11^{n_5}$ ... (k-te primtal)^{n_k}, hvis tegnstrengen består af k tegn, og det første tegn har nummer n_1, det andet tegn har nummer n_2, osv. Et naturligt tal n er ikke nødvendigvis et *Gödelnummer,* men hvis det er et sådant, vil vi betegne den tilsvarende tegnstreng G_n.

Tilsvarende kan ethvert bevis, som jo er en følge af udsagn, tildeles et nummer. Alle beviser kan opstilles på rad og række ved hjælp af et leksikografisk princip. Gödelnummeret til det k-te bevis betegnes B_k.

Gödel kalder en talteoretisk relation $R(x_1, ..., x_n)$ ($x_1, ..., x_n$ naturlige tal) *rekursiv,* hvis den kan dannes ved en særlig iterativ procedure. I årene 1932-36 fremkom der forskellige forslag til en definition af begrebet *effektiv beregnelighed* for en talteoretisk relation $R(x_1, ..., x_n)$. Dvs. en regneprocedure ved hjælp af hvilken man, for et givet sæt $(x_1, ..., x_n)$ af naturlige tal, kan afgøre om udsagnet $R(x_1, ..., x_n)$ er sandt eller falsk. Bl.a. *generel rekursiv* (Gödel, 1934), *λ-definerbar* (*Church* og *Kleene,* 1932-35) og *Turing-beregnelig* (*Turing,* l936). Det viste sig at disse definitioner er ensbetydende. Dette førte Church til at fremsætte den tese at "alle tilstrækkelig rummelige definitioner af effektiv beregnelighed er ensbetydende". En regneprocedure udformet som et computerprogram, i et tilstrækkelig rummeligt computersprog, er altså ensbetydende med disse definitioner. En relation som er rekursiv i Gödel's forstand kaldes idag *simpel* rekursiv.

Gödel beviser at for enhver rekursiv relation $R(x_1, ..., x_n)$, findes et formalt udsagn $r(x_1, ..., x_n)$ således at

a. ($R(x_1, ..., x_n)$ er sand) $\Rightarrow$ ($r(x_1, ..., x_n)$ er et teorem),

b. ($R(x_1, ..., x_n)$ er falsk) $\Rightarrow$ ($\neg r(x_1, ..., x_n)$ er et teorem),

her skal vi strengt taget i $r(x_1, ..., x_n)$ erstatte $x_1, ..., x_n$ med deres Gödelnumre.

Alle udsagn p(x), med een variabel x som gennemløber de naturlige tal, kan opstilles på rad og række ved hjælp af et leksikografisk princip. Gödel kalder et sådant udsagn "ein Klassenzeichen". Det n-te udsagn betegnes R(n). Funktionen "n $\rightarrow$ Gödelnummeret af R(n)" kan defineres i den formale teori. Udsagnet R(n)(m) betegnes [R(n):m].

Relationen R(m, n) = "B_m er ikke et bevis for [R(n):n]" er rekursiv, derfor findes der et formalt udsagn r(m, n) således at

a. "B_m er ikke et bevis for [R(n):n]" $\Rightarrow$ r(m, n) er et teorem

b. "B_m er et bevis for [R(n):n]" $\Rightarrow$ $\neg$r(m, n) er et teorem.

Udsagnet " $\forall$ m: r(m, n)" har een fri variabel, og er derfor R(q) for et naturligt tal q. Altså [R(q):n] = " $\forall$ m: r(m, n)", og derfor

$$[R(q):q] = \text{"} \forall m: r(m, q)\text{"}.$$

Gödel kalder en teori *ω-modsigelsesfri,* hvis der ikke findes et udsagn p(n) således at 1. $\neg(\forall n \in \mathbb{N}: p(n))$ er et teorem og 2. p(n) er et teorem for ethvert n $\in$ IN (husk at hvis " $\forall$ n $\in$ IN: p(n)" er et teorem er p(n) et teorem for ethvert n, men det omvendte gælder ikke, jvf. logisk aksiom ... side ...). Hvis teorien er ω-modsigelsesfri er den modsigelsesfri. Thi hvis den var kontradiktorisk, ville ethvert formalt udsagn jo være et teorem, derfor ville den ikke kunne være ω-modsigelsesfri.

Af det ovenstående følger nu:

1. Hvis teorien er modsigelsesfri kan [R(q):q] ikke bevises. Bevis: Hvis [R(q):q] var et teorem, ville findes et m' således at $B_{m'}$ var et bevis for [R(q):q]. Derfor ville, ifølge b, $\neg$r(m', q) være et teorem, men dette strider mod at " $\forall$ m: r(m, q)" er et teorem.

2. Hvis teorien er ω-modsigelsesfri kan $\neg$[R(q):q] ikke bevises. Bevis: Hvis teorien er ω-modsigelsesfri er den modsigelsesfri. Derfor kan, ifølge 1., [R(q):q] ikke bevises, dvs. for ethvert m er Bm ikke et bevis for [R(q):q]. Ifølge a er r(m, q) derfor et teorem for ethvert m. Hvis nu $\neg$[R(q):q] kunne bevises, ville $\neg(\forall$ m: r(m, q)) være et teorem, men dette strider imod at teorien er ω-modsigelsesfri.

Hvis teorien er modsigelsesfri kan " $\forall$ m: r(m, q)" altså ikke bevises, men r(m, q) er ikke destomindre et teorem for ethvert m. Hvis vi føjer "$\neg(\forall$ m: r(m, q))" til teorien som et nyt aksiom, får vi en teori som er stadig er modsigelsesfri (vis dette!), men ikke ω-modsigelsesfri, idet det både gælder at $\neg(\forall$ m: r(m, q)) er et teorem og at r(m, q) er et teorem for ethvert m.

For at bevise at [R(q):q] ikke kan bevises, forudsætter Gödel at teorien er modsigelsesfri, men for at bevise at $\neg$[R(q):q] ikke kan bevises, behøver Gödel en kraftigere forudsætning, nemlig at teorien er ω-modsigelsesfri. I 1936 viste *Rosser* at man kan nøjes med at forudsætte at teorien er modsigelsesfri.

I **§3** viser Gödel at enhver rekursiv relation er *aritmetisk,* dvs. den kan dannes alene ved hjælp af aritmetikkens tegn. Det følger heraf at i udsagnet " $\forall$ m: r(m, q)", kan r(m, q) erstattes med et aritmetisk udsagn s(m), således at hvis " $\forall$ m: r(m, q)" kan bevises da kan også " $\forall$ m: s(m)" bevises, og omvendt. Hvis altså teorien (som omfatter aritmetikken) er ω-modsigelsesfri, findes et aritmetisk udsagn som hverken kan bevises eller modbevises. Desuden viser Gödel at udsagnet " $\forall$ m: r(m, q)", kan erstattes med et udsagn " $\forall$ x: F(x)" i den rene prædikat kalkyle (se § ...), således at hvis " $\forall$ m: r(m, q)" kan bevises da kan også " $\forall$ x: F(x)" bevises, og omvendt. Hvis altså teorien (som omfatter aritmetikken) er ω-modsigelsesfri, findes et udsagn i den rene prædikat kalkyle som hverken kan bevises eller modbevises.

I **§4** skitseres beviset til "ein merkwürdiges Resultat". Nemlig at "for en teori som omfatter aritmetikken kan et bevis for at den er modsigelsesfri ikke formaliseres indenfor teorien selv". Et komplet bevis er langt og indviklet - det blev udarbejdet af Hilbert og Bernay i 1939.

At "teorien er modsigelsesfri" betyder at "der findes et udsagn som ikke kan bevises". Dette metamatematiske udsagn kan formaliseres som " $\exists$ k: G_k udsagn $\wedge$ $\forall$ m: B_m ikke bevis for k". Gödel argumenterer nu for at beviset i §3, for at "hvis teorien er modsigelsesfri da kan udsagnet [R(q):q] ikke bevises", kan omformes til et *formalt* aritmetisk bevis. Udsagnet "($\exists$ k: G_k udsagn $\wedge$ $\forall$ m: B_m ikke bevis for k) $\Rightarrow$ ($\exists$ m: B_m bevis for [R(q):q])" er altså et teorem. Men da det allerede er bevist at *hvis* teorien er modsigelsesfri da kan udsagnet [R(q):q] ikke bevises, følger heraf at præmissen " $\exists$ k: G_k udsagn $\wedge$ $\forall$ m: B_m ikke bevis for k" ikke kan bevises hvis teorien er modsigelsesfri.

Til sidst nævner Gödel at forudsætningen i §3, at teorien er ω-modsigelsesfri, nu kan erstattes med "at teorien er kontradiktorisk kan ikke bevises i teorien selv", thi da er udsagnet " $\exists$ k: G_k udsagn $\wedge$ $\forall$ m: B_m ikke bevis for k" et eksempel på et udsagn som hverken kan bevises eller modbevises.

I 1936 beviste Church at enhver teori, som omfatter aritmetikken eller den rene prædikat kalkyle, er *uafgørbar.* Dvs. der findes ikke en effektiv beregningsprocedure, ved hjælp af hvilken man kan afgøre om et givet udsagn har et bevis eller ej. I tilfældet hvor teorien omfatter aritmetikken, går Church således til værks:

Alle computerprogrammer som udfører beregninger med naturlige tal, således at såvel input som (evt.) output er ét tal, opstilles på rad og række ved hjælp af et leksikografisk ordningsprincip. Det k-te program betegnes P_k. Hvis tallet m er input i P_k, kan der ske 1. programmet går ikke igang, 2. programmet går igang men kører i det uendelige, 3. programmet går igang og standser efter endelig mange udregninger, output er et tal n. Lad udsagnet T(k, m, n) være givet ved: T(k, m, n) er sandt hvis og kun hvis 3. gælder, og lad funktionen t: IN → IN være givet ved

$t(m) = n + 1$ hvis T(m, m, n) er sandt, og

$t(m) = 0$ hvis der ikke findes noget n således at T(m, m, n) er sandt.

Denne funktion er ikke beregnelig. Bevis: Hvis t var beregnelig, ville der findes et tal q således at P_q beregner den, dvs. T(q, m, t(m)) ville være sandt for ethvert m. Specielt ville T(q, q, t(q)) være sandt. Men dette ville medføre at $t(p) = t(p) + 1$ (da n er entydigt bestemt ved k og m, hvis T(k, m, n) er sandt).

Det følger umiddelbart heraf at det ikke, ved en beregnings-procedure, kan afgøres om udsagnet " ∃ n: T(m, m, n) er sandt" (med én variabel m) er sandt eller falsk.

Ifølge et tidligere nævnt resultat af Gödel, findes et formalt udsagn r(m, n) således at

a. T(m, m, n) er sandt ⇒ r(m, n) er et teorem

b. T(m, m, n) er falsk ⇒ ¬r(m, n) er et teorem,

og af dette følger at

(∃ n: T(m, m, n) er sandt) ⇔ ((∃ n: r(m, n)) er et teorem).

Hvis teorien var både ω-modsigelsesfri og afgørbar, ville sandhedsværdien af udsagnet "(∃ n: r(m, n)) er et teorem" (med én fri variabel m) være beregnelig. Og dette ville medføre at sandhedsværdien af udsagnet " ∃ n:T(m, m, n) er sandt" var beregnelig, men det har vi lige vist at den ikke er.