TECHNISCHE UNIVERSITÄT DRESDEN

Fakultät Informatik, Institut für Künstliche Intelligenz, Professur Computational Logic

THEORETISCHE INFORMATIK UND LOGIK

23. Vorlesung: Gödels 1. Unvollständigkeitssatz

Hannes Straß

TU Dresden, 7. Juli 2022

Kurt Gödel

Der 1. Gödelsche Unvollständigkeitssatz

Was Gödel 1931 zeigte, war – grob gesagt – folgendes:

Satz: Jedes konsistente formale System, in dem eine gewisse Menge elementarer Arithmetik dargestellt werden kann, ist unvollständig in Bezug auf die Beweisbarkeit von Sätzen der elementaren Arithmetik:
Es gibt solche Sätze, die weder bewiesen noch widerlegt werden können.

Um das zu verstehen, müssen wir einiges klären:

- Was ist ein formales System?
- Was ist konsistent?
- Was ist „eine gewisse Menge elementarer Arithmetik"?
- Was genau bedeutet unvollständig hier?

Formale Systeme

Ein formales System ist ganz allgemein ein Beweissystem, bestehend aus:

- Einer formalen Sprache, in der Aussagen formuliert werden können;
- einer Menge von Axiomen, d.h. als wahr vorgegebener Aussagen;
- einem effektiven Verfahren, mit dem man aus gegebenen Aussagen neue Schlüsse ableiten kann.

Beweisbare Sätze heißen Theoreme des formalen Systems.

Anmerkung: Auch die Axiome sind Theoreme, wenn auch mit sehr kurzen Beweisen.

Beispiel: Die Prädikatenlogik liefert formale Systeme:

- Sprache: Die Sprache der prädikatenlogischen Sätze.
- Axiome: Eine gegebene Theorie, z.B. die Theorie der kommutativen Monoide.
- Ableitungsverfahren: Resolutionskalkül

Formale Systeme: Wichtige Eigenschaften

Formale Systeme können viele Formen haben – zum Beispiel beinhalten sie auch jedes System, in dem mathematische Beweise formal geführt werden können (z.B. ZFC: Mengenlehre nach Zermelo-Fraenkel mit Auswahlaxiom).

Die Details sind relativ egal, aber wir wollen doch ein paar Anforderungen stellen:

Grundeigenschaft Formaler Systeme: Die Menge der Theoreme eines formalen Systems ist rekursiv aufzählbar (semi-entscheidbar).

Negation: Für jeden Satz S gibt es in der Sprache eines formalen Systems auch einen negierten Satz $\neg S$, so dass gilt:

- S ist genau dann ein Theorem, wenn $\neg\neg S$ ein Theorem ist.
- Wenn S und $\neg S$ Theoreme sind, dann sind alle Sätze Theoreme.

Eigenschaften formaler Systeme

Syntaktische Eigenschaften:

- Konsistenz: Ein formales System ist konsistent, wenn es keinen Satz S gibt, so dass S und $\neg S$ beweisbar sind.
- Vollständigkeit: Ein formales System ist (negations-)vollständig, wenn für jeden Satz S entweder S oder $\neg S$ beweisbar sind.

Wenn man z.B. in Prädikatenlogik arbeitet, dann kann man Mengen von Sätzen eine Semantik (Modelltheorie) geben und weitere Eigenschaften fordern:

Semantische Eigenschaften:

- Korrektheit: Ein formales System ist korrekt, wenn jeder beweisbare Satz auch semantisch wahr (tautologisch) ist.
- Vollständigkeit: Ein formales System ist (semantisch) vollständig, wenn jeder semantisch wahre Satz auch beweisbar ist.

Verwechslungsgefahr

Achtung! Es gibt zwei Arten von Vollständigkeit:

- Negationsvollständigkeit: Die syntaktische Eigenschaft, dass jeder Satz bewiesen oder widerlegt werden kann.
- Semantische Vollständigkeit: Die semantische Eigenschaft, dass jede Tautologie bewiesen werden kann.

Gödels Vollständigkeitssatz bezieht sich auf die zweite Art von Vollständigkeit, Gödels Unvollständigkeitssätze auf die erste!

Beispiele

Beispiel: Angenommen, ein Satz F kann in einem formalen System $\mathcal{S}$ bewiesen werden und wir wissen, dass $\mathcal{S}$ konsistent ist. Folgt daraus, dass F semantisch wahr ist?

Beispiel: Aus der (monotonen) Logik bekannte Zusammenhänge gelten auch hier:

- Ein Satz F ist genau dann in $\mathcal{S}$ beweisbar, wenn $\mathcal{S}$ bei Hinzunahme des Axioms $\neg F$ inkonsistent wird.
- Weder F noch $\neg F$ sind in $\mathcal{S}$ beweisbar gdw. $\mathcal{S}$ sowohl bei Hinzunahme von F als auch bei Hinzunahme von $\neg F$ konsistent ist.

Eine gewisse Menge Arithmetik

Satz: Jedes konsistente formale System, in dem eine gewisse Menge elementarer Arithmetik dargestellt werden kann, ist unvollständig in Bezug auf die Beweisbarkeit von Sätzen der elementaren Arithmetik:
Es gibt solche Sätze, die weder bewiesen noch widerlegt werden können.

Was ist mit „eine gewisse Menge elementarer Arithmetik“ gemeint?

- Das System sollte Sätze über Beziehungen von konkreten natürlichen Zahlen ausdrücken können.
- Dabei sollten die elementaren Operationen $+$, $-$ und $\times$ sowie die Relation $=$ unterstützt werden.
- Das System sollte bezüglich der gängigen Semantik dieser arithmetischen Ausdrucksmittel korrekt sein.

Oft kommt man mit noch weniger aus, aber diese Eigenschaften reichen in der Regel.

Arithmetik in Prädikatenlogik

Man kann die benötigte Menge von Arithmetik durch eine prädikatenlogische Theorie axiomatisieren:

- Konstante $\mathtt{0}$
- Funktionssymbole s (unär, „Nachfolger“), $+$ und $\times$ (binär, infix)
- Prädikatssymbol $\approx$ (binär, infix)

Darstellung natürlicher Zahlen als Nachfolger von 0:

$$\mathtt{0} \mathrel{\hat=} 0, \quad s(\mathtt{0}) \mathrel{\hat=} 1, \quad s(s(\mathtt{0})) \mathrel{\hat=} 2, \quad \ldots$$

Sätze zur Axiomatisierung der Grundrechenarten:

- $\forall x.\neg(s(x) \approx \mathtt{0})$
- $\forall x.(x + \mathtt{0} \approx x)$
- $\forall x, y.(x + s(y) \approx s(x + y))$
- $\forall x.(x \times \mathtt{0} \approx \mathtt{0})$
- $\forall x, y.(x \times s(y) \approx (x \times y) + x)$
- ... (Mit obigem kann man schon einiges ausrechnen, aber es gibt noch mehr zu sagen, z.B. eine komplette Gleichheitstheorie.)

Gödels Beweis des 1. Satzes

Idee

Wie kann man zeigen, dass irgendein Satz weder beweisbar noch widerlegbar ist?

Ein einfaches Szenario:

- Annahme: Sie glauben nur Dinge, die wirklich wahr sind; und was Sie glauben, ist (in seiner Gesamtheit) konsistent.
- Behauptung: Es gibt wahre Sätze, die Sie nicht glauben.
- Beweis: „Ich mache jetzt eine Behauptung, die Sie mir nicht glauben“ ist eine Behauptung, die Sie nicht glauben können:
  - Wenn Sie sie glauben, dann muss sie wahr sein, d.h. Sie glauben sie nicht – Widerspruch.
  - Also glauben Sie sie nicht. Dann ist die Behauptung wahr. □

Behauptungen dieser Form nennt man Gödelsätze.

Gödelsätze formal machen

Natürliche Sprache eignet sich nicht für stichhaltige Argumente, da sie keine klar definierte mathematische Interpretation hat.
(Zum Beispiel können wir darin sagen: „Dieser Satz ist falsch.“)

Gödels Beweis (vereinfacht) für formale Systeme:

- Annahme: Das gegebene System $\mathcal{S}$ ist korrekt und konsistent.
- Behauptung: Es gibt wahre Sätze, die nicht beweisbar sind.
- Beweis: Gödel definiert eine mathematische Formel F, welche ausdrückt:

  „F ist in System $\mathcal{S}$ nicht beweisbar.“

  – Wenn F beweisbar wäre, dann wäre es wahr und daher nicht beweisbar – Widerspruch.
  – Also ist F nicht beweisbar, und damit wahr. □

Gödels Zahlen

„F ist in System $\mathcal{S}$ nicht beweisbar.“

Die Herausforderung ist, solche Gödelsätze zu definieren:
Offenbar kann man F nicht als Teilausdruck in F verwenden!

Gödel definiert daher numerische Bezeichner für Formeln – sogenannte Gödelzahlen – und kodiert seine Sätze anders:

„Die Zahl m_F ist Element der Menge der Gödelzahlen nicht beweisbarer Sätze.“

wobei m_F die Gödelzahl für F ist.

Man muss dafür eine Menge technischer Ergebnisse zeigen, z.B.

- Ein solcher Satz F, welcher seine eigene Gödelzahl m_F verwendet, existiert überhaupt.
- Die Menge der Gödelzahlen nicht beweisbarer Sätze ist arithmetisch darstellbar.

Quotes und Quines

Eine verwandte Form von Selbstbezüglichkeit ist auch in der Informatik bekannt:

Ein Quine ist ein Programm, das bei einer leeren Eingabe seinen eigenen Quellcode ausgibt.

Auch hier denkt man vielleicht zuerst, dass dies nicht möglich wäre, weil das Programm dazu seinen eigenen Code enthalten müsste – eine unendliche Rekursion . . .

Es ist aber gar nicht so schwer:

Beispiel:
Gib den folgenden Satz zweimal aus, beim zweiten mal in Anführungszeichen:
"Gib den folgenden Satz zweimal aus, beim zweiten mal in Anführungszeichen: "

Gödels 1. Satz berechnungstheoretisch zeigen

Unentscheidbarkeit und Unvollständigkeit

Wir wissen:

Die Theoreme formaler Systeme sind rekursiv aufzählbar, d.h. semi-entscheidbar.

- Speziell gilt das auch für die Theoreme, die sich auf arithmetische Aussagen beziehen.

Es gibt Mengen (Sprachen), die nicht semi-entscheidbar sind.

- Zum Beispiel das Komplement des Halteproblems für Turingmaschinen.

$\rightsquigarrow$ Falls man Instanzen des Wortproblems einer nicht-semi-entscheidbaren Sprache auf die Wahrheit arithmetischer Sätze reduzieren könnte, dann würde daraus schon Unvollständigkeit folgen.

Zahlen vs. Sprachen

Die Theorie der Berechenbarkeit handelt von Sprachen, d.h. Mengen von Wörtern.

Es ist leicht möglich, jedem Wort $w \in \Sigma^*$ eine natürliche Zahl $n_w \in \mathbb{N}$ zuzuordnen – man spricht von einer Gödelzahl für das Wort.

(Dies ist möglich, da die Menge aller Wörter abzählbar ist, siehe Formale Systeme WS 2021/2022; natürlich gibt es viele mögliche Zuordnungen.).

Wir folgern:

- Es gibt eine Eins-zu-eins-Beziehung zwischen formalen Sprachen $\mathbf{L} \subseteq \Sigma$ (über einem gegebenen Alphabet Σ) und Mengen natürlicher Zahlen.
- Wir können also von der Entscheidbarkeit/Unentscheidbarkeit einer Menge natürlicher Zahlen sprechen.
- Es gibt unentscheidbare Mengen natürlicher Zahlen, z.B. die Menge der Zahlen n_w, welche ein Wort w kodieren, das eine Turingmaschine kodiert ($w = \mathrm{enc}(\mathcal{M})$), welche auf der leeren Eingabe hält.

Eine erste Beweisidee

Mögliche Strategie zum Beweis von Gödels 1. Satz:

- Definiere eine Gödel-Nummerierung für alle Wörter, die eine Turingmaschine kodieren.
- Finde arithmetische Ausdrücke auf diesen Zahlen, mit denen man konkrete syntaktische Eigenschaften der kodierten TM testen kann.
- Finde arithmetische Formeln, die ausdrücken, dass die kodierte TM auf dem leeren Wort hält.

$\rightsquigarrow$ Programmiere eine universelle TM in Arithmetik.

(Machbar, aber aufwändig!)

Ein schweres arithmetisches Problem

Die Aufgabe wird viel einfacher, wenn man ein unentscheidbares Problem verwendet, welches sich schon auf arithmetische Aussagen bezieht.

Eine Diophantische Gleichung ist ein arithmetischer Ausdruck der Form $D[x_1, \ldots, x_n] = 0$, wobei $D[x_1, \ldots, x_n]$ ein Funktionsterm über den Symbolen $x_1, \ldots, x_n, \mathtt{0}, s, +, -, \times$ ist. Eine Lösung der Gleichung ist eine Liste von natürlichen Zahlen $z_1, \ldots, z_n$, für welche die Gleichung erfüllt wird.

Beispiele für diophantische Gleichungen:

- $y = 3x + 5$, d.h. $3 \times x - y + 5 = 0$ (unendlich viele Lösungen)
- $x^2 + y^2 = -1$, d.h. $x \times x + y \times y + 1 = 0$ (keine Lösung)
- $x^2 = 2y^4 - 1$ (genau zwei Lösungen: $\langle x, y\rangle = \langle 1, 1\rangle$ und $\langle x, y\rangle = \langle 239, 13\rangle$)
- $(x + 1)^2 + (y + 1)^2 = (z + 1)^2$ (unendlich viele Lösungen; vgl. Pythagoräische Tripel)
- $(x + 1)^4 + (y + 1)^4 = (z + 1)^4$ (keine Lösung; Fermatsche Vermutung)

Hilberts Zehntes Problem und MRDP

Hilberts 10. Problem: Man finde ein Verfahren um zu ermitteln, ob eine gegebene diophantische Gleichung lösbar ist.

Wir wissen heute, dass dies unentscheidbar ist:

- Jede diophantische Gleichung $D[x_1, \ldots, x_n, y] = 0$ definiert eine Menge natürlicher Zahlen $\{k \mid D[x_1, \ldots, x_n, k] = 0 \text{ ist lösbar}\}$.
- Eine so definierte Menge heißt diophantische Menge.
- Alle diophantischen Mengen sind rekursiv aufzählbar: Teste systematisch alle möglichen Belegungen von $D[x_1, \ldots, x_n, y] = 0$ und gib bei jeder gefundenen Lösung den Wert von y aus.

Satz (Matiyasevich/Robinson/Davis/Putnam): Jede rekursiv aufzählbare Menge natürlicher Zahlen ist diophantisch.

Intuitiv: Diophantische Gleichungen sind Turing-mächtig.

Konsequenzen

Wir folgern:

- Hilberts 10. Problem ist wirklich unentscheidbar:
  - Wäre es entscheidbar, dann könnte man jede diophantische Menge entscheiden. (Kontrollfrage: Wie?)
  - Wir wissen aber bereits, dass es rekursiv aufzählbare Mengen natürlicher Zahlen gibt, die nicht entscheidbar sind.
- Nicht-Lösbarkeit diophantischer Gleichungen ist nicht semi-entscheidbar:
  - Lösbarkeit ist bereits semi-entscheidbar.
  - Wenn Nicht-Lösbarkeit auch semi-entscheidbar wäre, dann wäre beides entscheidbar.

Beweis von Gödels 1. Unvollständigkeitssatz

Sei $\mathcal{S}$ ein formales System, welches folgende Bedingungen erfüllt:

(1) $\mathcal{S}$ ist konsistent.

(2) Jede wahre Aussage der Form $D[z_1, \ldots, z_n] = 0$ für eine beliebige diophantische Gleichung und konkrete Zahlenwerte $z_1, \ldots, z_n$ ist beweisbar.

(Einfach durch Ausrechnen des Wertes.)

(3) Sätze der Form $\exists x_1, \ldots, x_n.D[x_1, \ldots, x_n] = 0$ sind darstellbar und alle beweisbaren Sätze dieser Form sind wahr.

Wegen (1) und (2) gilt:
Beweist $\mathcal{S}$ einen Satz der Form $\neg\exists x_1, \ldots, x_n.D[x_1, \ldots, x_n] = 0$, so ist dieser wahr.

Aber: $\mathcal{S}$ kann nicht alle wahren Sätze dieser Form beweisen:

- Die Menge aller beweisbaren Sätze dieser Form ist semi-entscheidbar.
- Die Menge aller wahren Sätze dieser Form ist nicht semi-entscheidbar.

Mit (3) gilt: Es gibt Sätze F der Form $\exists x_1, \ldots, x_n.D[x_1, \ldots, x_n] = 0$,
so dass $\mathcal{S}$ weder F noch $\neg F$ beweisen kann. □

Der 1. Gödelsche Unvollständigkeitssatz

Satz (Gödel, 1931): Jedes konsistente formale System, in dem eine gewisse Menge elementarer Arithmetik dargestellt werden kann, ist unvollständig in Bezug auf die Beweisbarkeit von Sätzen der elementaren Arithmetik.

Relevante Begriffe:

- Formales System: Ein implementierbares Verfahren, mit dem man Theoreme endlich beweisen kann.
- Konsistent: Man kann niemals eine Aussage und ihr Gegenteil beweisen.
- Gewisse Menge Arithmetik: Kodierung konkreter natürlicher Zahlen und deren korrekte Addition, Subtraktion, Multiplikation und Vergleich.
- Unvollständig: Es gibt Sätze, die weder bewiesen noch widerlegt werden können.

Beispiele

Beispiel: Natürliche Zahlen und einfache Rechenregeln können mit einer prädikatenlogischen Theorie definiert werden. Mit Resolution kann man daraus korrekte neue Schlüsse ziehen. Laut Gödels erstem Satz kann man auf diese Art aber niemals alle wahren Aussagen der Arithmetik beweisen, außer wenn die Theorie widersprüchlich ist.

Keine prädikatenlogische Theorie kann die elementare Arithmetik vollständig beschreiben.

Beispiel: Die moderne Mathematik basiert auf der Mengenlehre von Zermelo-Fraenkel unter Hinzunahme des Auswahlaxioms. Dieses formale System heißt ZFC. Es ist klar definiert, was ein korrekter mathematischer Beweis in ZFC ist. Laut Gödels erstem Satz gibt es also wahre Aussagen über elementarer Arithmetik, die nicht in ZFC bewiesen werden können.

Zusammenfassung und Ausblick

Formale Systeme, die elementare Rechnungen auf natürlichen Zahlen ausführen können, sind entweder inkonsistent oder (negations-)unvollständig.

Diese Unvollständigkeit hat damit zu tun, dass nicht alle Mengen natürlicher Zahlen Turing-erkennbar sind.

Was erwartet uns als nächstes?

- Gödels zweiter Unvollständigkeitssatz
- Zusammenfassung und Ausblick
- Probeklausur und Prüfung

Bildrechte

Folie 4: Fotografie um 1926, gemeinfrei