

Logika dla informatyków

Andrzej M. Borzyszkowski

Instytut Informatyki
Uniwersytet Gdański

inf.ug.edu.pl/~amb

Spełnialność

Definicja

Formuła φ jest prawdziwa jeśli $\models \varphi$

uwaga: to jest to samo co tautologia – formuła, której znaczeniem jest **T** przy każdym wartościowaniu formuł atomowych (tj. modelu).

Definicja

Formuła φ jest spełnialna, jeśli dla pewnego wartościowania φ ma wartość **T**

Twierdzenie

Formuła φ jest tautologią wtedy i tylko wtedy, gdy $\neg\varphi$ nie jest spełnialna.

Oczywiście, jeśli $\neg\varphi$ jest **T** dla pewnego modelu, to φ dla tego modelu nie jest **T**, czyli φ nie jest tautologią i na odwrót.

Wniosek

Sprawdzanie tautologii i sprawdzanie spełnialności są równoważne.

SAT solvers

Pierwsze podejście

- Ograniczamy się do formuł zadanych gramatyką:

$$\varphi ::= p \mid (\neg\varphi) \mid (\varphi \wedge \varphi)$$

tzn. tylko negacja i koniunkcja.

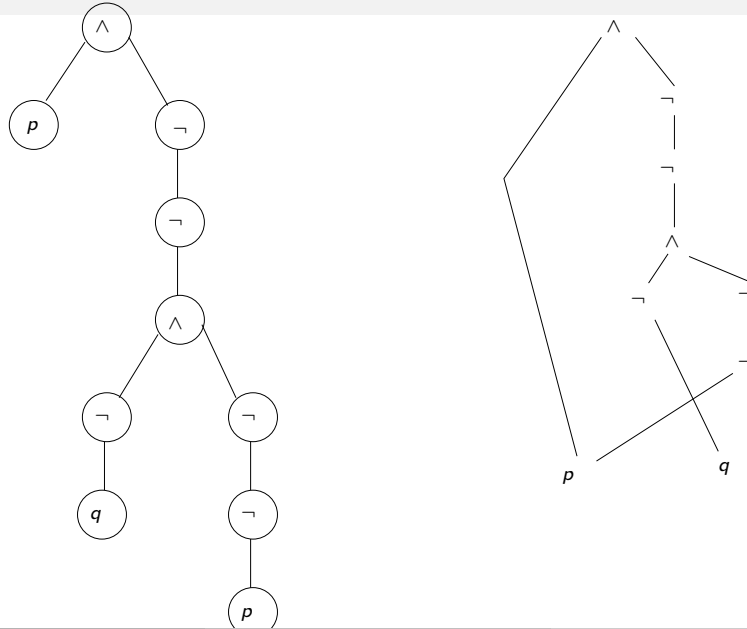
- Każda formuła może być przetłumaczona do równoważnej semantycznie formuły z tej klasy.
- Tłumaczenie:

$$\begin{aligned} T(p) &= p \\ T(\neg\varphi) &= \neg(T(\varphi)) \\ T(\varphi_1 \wedge \varphi_2) &= T(\varphi_1) \wedge T(\varphi_2) \\ T(\varphi_1 \vee \varphi_2) &= \neg(\neg T(\varphi_1) \wedge \neg T(\varphi_2)) \quad (\text{deMorgan}) \\ T(\varphi_1 \rightarrow \varphi_2) &= \neg(T(\varphi_1) \wedge \neg T(\varphi_2)) \end{aligned}$$

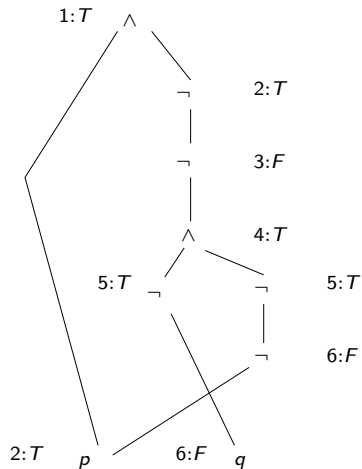
- Formuły φ i $T(\varphi)$ mają równe wartości w tych samych modelach.
- Tłumaczenie T zamienia jeden węzeł oryginalny na co najwyżej cztery węzły nowe.

$$T(p \wedge \neg(q \vee \neg p)) = p \wedge \neg(\neg q \wedge \neg p)$$

Reprezentacja jako DAG



Szukanie modelu



Wartościowanie w grafie

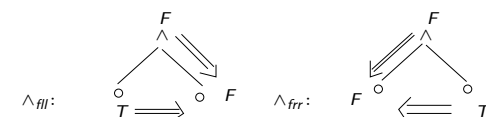
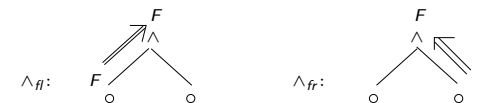
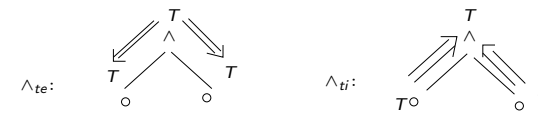
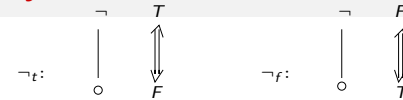
- Jeśli formuła dana grafem ma być spełnialna, to w korzeniu musi być przypisana wartość **T**.
- W dalszych krokach być może uda się wpisać konieczne wartości dla innych węzłów.
- Jeśli otrzymamy wpisy przy formułach atomowych oznacza to, że model jest jedynym możliwym, w którym formuła jest spełniona. Sprawdzamy, czy rzeczywiście i mamy dowód spełnialności.

Definicja

SAT solver – algorytm sprawdzający spełnialność formuły.

- Przykład dla wcześniejszej formuły na następnym slajdzie

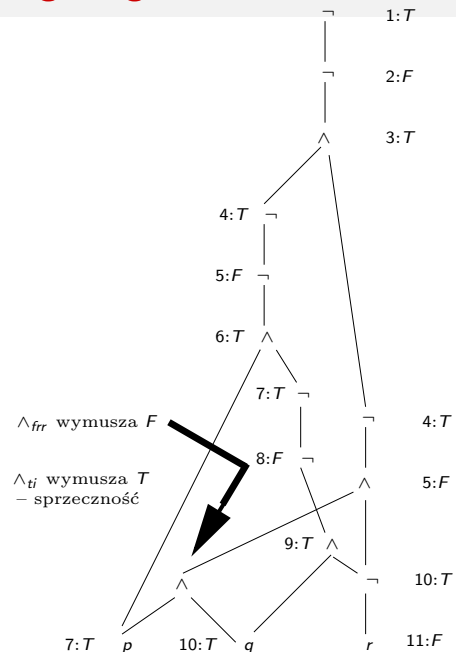
Reguły wpisywania wartości



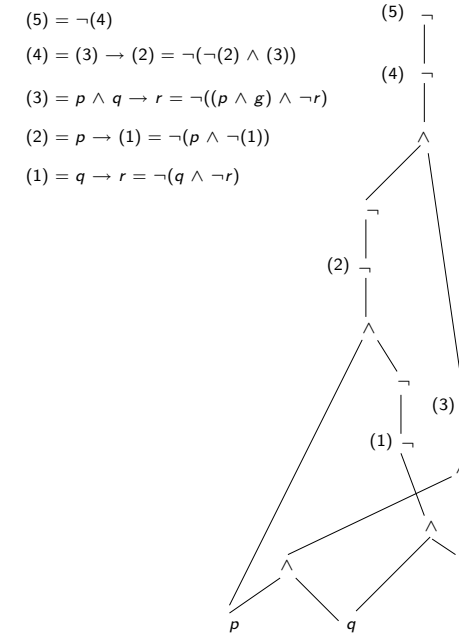
Większy przykład

- Wiemy, że $p \wedge q \rightarrow r \vdash p \rightarrow q \rightarrow r$ (zakładamy najpierw p , potem q , tworzymy koniunkcję i odrywamy r).
- Chcemy to sprawdzić automatycznie. Oczywiście nie testując wszystkich wartościowań, jest ich wykładniczo wiele.
- Formuła $(p \wedge q \rightarrow r) \rightarrow p \rightarrow q \rightarrow r$ jest prawdziwa w.t.w. $\neg((p \wedge q \rightarrow r) \rightarrow p \rightarrow q \rightarrow r)$ nie jest spełnialna.
- Okaże się, że próba wstawienia znaczników tak, by w korzeniu było **T** nie powiedzie się.

Brak wartościowania zgodnego



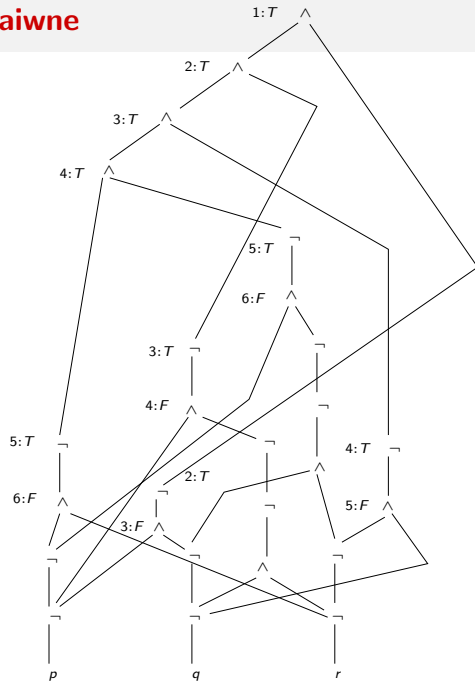
Graf dla $\neg((p \wedge q \rightarrow r) \rightarrow p \rightarrow q \rightarrow r)$



Złożoność (i prostota) tego rozwiązania

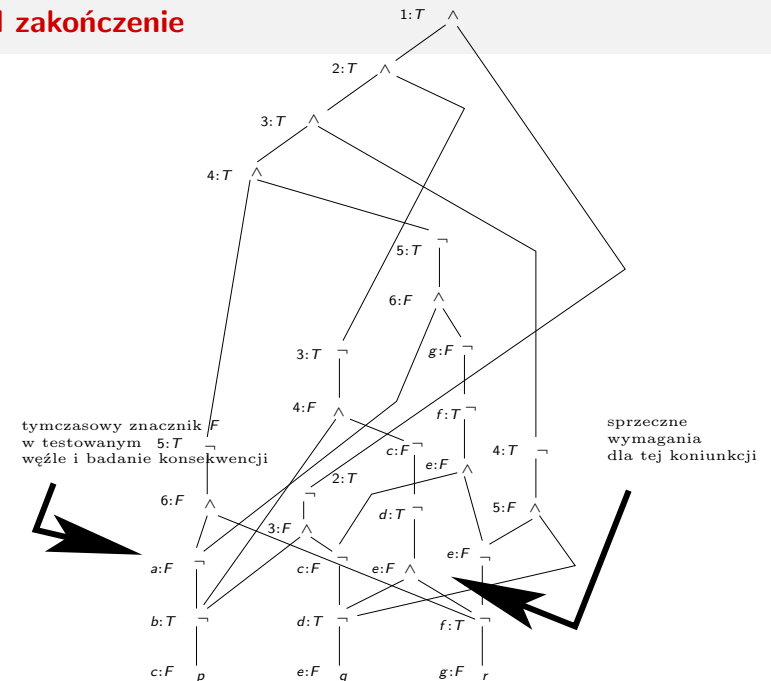
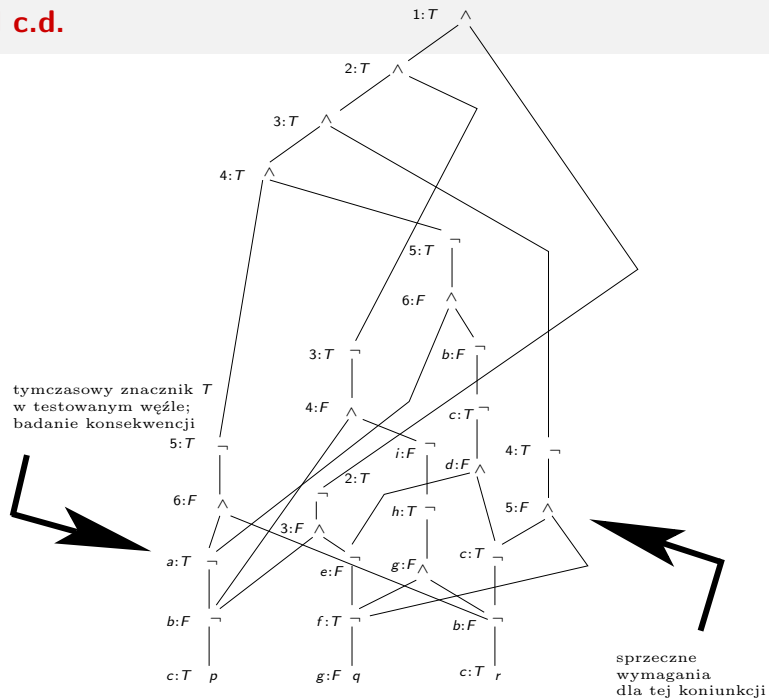
- Podany SAT solver ma złożoność liniową
- dowód: graf wypełnia się stopniowo znacznikami
- Wady: ten SAT solver nie może zbadać spełnialności formuły $\neg(\varphi_1 \wedge \varphi_2)$
- Przykład: formuła $(p \vee q \vee r) \wedge (p \vee \neg q) \wedge (q \vee \neg r) \wedge (r \vee \neg p) \wedge (\neg p \vee \neg q \vee \neg r)$
- w środku mamy $q \rightarrow p \rightarrow r \rightarrow q$ tzn. wszystkie wartości są równe
- i jednocześnie co najmniej jeden atom jest prawdziwy i jeden fałszywy: formuła nie jest spełnialna
- w pięciu węzłach postaci $\neg(\varphi_1 \wedge \varphi_2)$ próba obliczenia wartości zatrzymuje się, nie mamy żadnych informacji, czy istnieje zgodne wartościowanie węzłów

SAT solver nieliniowy



- ➊ Wartościmy węzły grafu zaczynając od **T** w kierunku
 - ➋ dla każdego węzła niezaznaczonego próbujemy wpisać tymczasowo **T** (i konsekwentnie dla innych węzłów)
 - ➌ analogicznie wpisujemy tymczasowo **F**
 - ➍ jeśli obie próby dla pewnego węzła doprowadziły do sprzeczności, oznacza to niespełnialność formuły
 - ➎ w.p.p. można wpisać *na stałe* wartości wspólne dla obu prób
- złożoność jest sześcienna: przebiegamy wszystkie niezaznaczone węzły, dla każdego próbujemy rozwiązania, dla każdego innego węzła sprawdzamy, czy ma stałe znaczniki w obu próbach

Przykład zakończenie



Udoskonalenia

- j.w.
- ale, jeśli jedna z prób doprowadza do sprzeczności, a druga nie, to przyjmujemy to zaznaczenie na stałe
- Nadal nie każda formuła może być przetworzona do końca
- są formuły, dla których nie potrafimy w ten sposób potwierdzić lub zaprzeczyć spełnialności
- przykład na następnym slajdzie

Problemy to nie przypadek

Twierdzenie (Cook 1971)

Problem SAT jest NP-trudny (a ściślej NP-zupełny)

- tzn. łatwo sprawdzić podane rozwiązanie,
- ale nie ma łatwego algorytmu znajdującego rozwiązanie
- oczywiście zawsze można wypróbować wszystkie możliwe wartościowania formuł atomowych

Konsekwencje NP-zupełności

- żaden problem NP-zupełny nie posiada znanego algorytmu o wielomianowej złożoności ($P \neq NP?$)
- nawet najlepsze algorytmy weryfikacji SAT mają wykładniczą złożoność pesymistyczną
- w praktyce pesymistyczne przypadki nie zdarzają się same, są one celowo budowane
- współczesne algorytmy weryfikacji SAT rozwiązują problemy wielkości milionów klauzul i setek tysięcy zmiennych zdaniowych w sekundy

Kolejna trudna formuła

