LOGICKÁ ARGUMENTACE MEZI AGENTY[1]
LOGICAL ARGUMENTATION AMONG AGENTS

Svatopluk Nevrkla
Katedra logiky FF UK
svata@logici.cz

Michal Peliš
Oddělení logiky, FLÚ AV
pelis@ff.cuni.cz

Klíčová slova: Znalosti, argumentace, logika

Abstrakt:
Poznání jednotlivce je neustále dynamicky utvářeno nejen jeho kontaktem s 'neuvědomělou' přírodou, ale především jeho interakcí s ostatními poznávajícími subjekty. Tato interakce se odehrává v jazyce a často nabývá formy dialogu, jehož jedním druhem je argumentace (dalšími příbuznými druhy dialogu jsou např. vysvětlování, vyjednávání apod.). V argumentaci konfrontuje poznávající subjekt (agent) své znalosti a jejich opodstatnění se znalostmi a opodstatněními jiného agenta. Matematické modely pro zkoumání argumentace se pak zabývají pravidly, jakými je argumentace zpracovávána a vyhodnocována. Nás bude zajímat pouze argumentace založená na logických argumentech, jejichž závěr logicky vyplývá z premis.

Keywords: Knowledge, argumentation, logic, multiagent systems

Abstract:
Cognition of an individual is a dynamic process, shaped not only by his contact with 'inanimate' nature, but mainly by his interaction with other cognizing beings. This interaction takes place in language and often takes form of a dialogue, whose one particular form is argumentation (similar forms of dialogue being explanation, persuasion, negotiation etc.). Cognizing being (agent) confronts his beliefs and their justification with those of another agent in argumentation. There are several mathematical models for argumentation, describing rules of its evaluation and processing. We are mainly concerned with argumentation based on logical arguments, where claim logically follows from the premises.

[1] Práce na tomto příspěvku byla podpořena grantem GA ČR č. 401/09/H007 "Logické základy sémantiky" (první z autorů) a grantem Logical Models of Reasoning with Vague Information – EUROCORES programme LogiCCC of the European Science Foundation (druhý z autorů).

1. Úloha argumentace v lidském poznání

Poznání jednotlivce se v průběhu jeho života neustále vyvíjí na základě jeho konfrontace s novými poznatky, ale především v důsledku jeho komunikace s dalšími lidmi.

Komunikací a její úlohou při utváření poznání se pak zabývá mnoho rozličných vědních oborů od didaktiky, psychologie, sociologie, až po informatiku. Právě poslední jmenovaná disciplína se často opírá o použití matematických modelů. Jako jeden z formálních modelů k popisu informací a vztahu mezi nimi lze chápat i matematickou logiku, ve které jsou informace zachyceny pomocí vět formálního jazyka.
Předmětem zájmu logiky jsou sice především 'věčné' vztahy mezi těmito výroky, avšak nově vznikající odvětví logiky jako jsou např. dynamická logika [2] či nemonotónní usuzování a revize přesvědčení [1] se nově soustředí i na dynamiku mezi výroky a znalostmi či informacemi, které tyto výroky reprezentují.

Pro tyto nové přístupy je však typické, že se zabývají až důsledky přijetí nějaké nové informace na původní přesvědčení nositele, případně skupiny nositelů znalostí, a nezabývají se popisem a vyhodnocováním podmínek, za nichž tato nová informace může být oněmi nositeli oprávněně přijata, či odmítnuta.

Logika však také může být použita, a tradičně tak používána byla, i k tomuto účelu, a to konkrétně k analýze a popisu procedur, jakými jsou nám informace předávány v jistém specifickém druhu komunikace.

Tímto specifickým druhem komunikace je dialog. Dialog předpokládá interakci dvou (ale i více) agentů, která je obvykle vedena podle nějakých pravidel, a která má za účel revizi přesvědčení, přání, úmyslů či cílů účastníků dialogu. My se dále soustředíme pouze na popis těch druhů dialogů, jejichž účelem je pouze změna přesvědčení, či znalostí partnera v dialogu a ještě konkrétněji na argumentaci.

Argumentací budeme nadále rozumět druh dialogu, v němž se jeho účastníci (budeme předpokládat, že jsou alespoň dva), střídají v prezentaci argumentů, přičemž účelem argumentace je svými argumenty v posledku vyvrátit veškeré argumenty protivníka podle daných pravidel.

Oproti jiným druhům dialogu, ve kterém se také můžeme setkat s argumenty (např. argumenty jimiž účastník podkládá jiné své argumenty při zdůvodňování, či poukazuje na důsledky jiných argumentů při přesvědčování), pro argumentaci je právě typické, že argumenty jsou vyhodnocovány z hlediska jejich vztahů vzájemné výlučnosti.

2. Matematické modely argumentace

Jedním z převratných modelů pro formální zachycení argumentace jsou tzv. abstraktní argumentační rámce, které navrhl v článku [5] P.M. Dung. Dung ve svém modelu zachycuje pouze strukturu vzájemných vztahů mezi argumenty a zcela abstrahuje od povahy argumentů a jejich vnitřní struktury.

Definice 2.1. (Argumentační rámec)
Argumentačním rámcem budeme mínit libovolnou dvojici $\langle \boldsymbol{\Gamma}, \rightarrow \rangle$, kde $\boldsymbol{\Gamma}$ je libovolná množina reprezentující argumenty a $\rightarrow$ libovolná relace na množině $\boldsymbol{\Gamma}$, reprezentující relaci napadání, resp. vylučování.

Dále Dung definuje množiny argumentů, které hájí samy sebe proti všem možným protiargumentům zvenčí. Obzvláště významným příkladem takových množin jsou takzvané stabilní extenze.

Definice 2.2. (Stabilní extenze)
Nechť je dán argumentační rámec $\langle \boldsymbol{\Gamma}, \rightarrow \rangle$. Řekneme, že množina argumentů $\boldsymbol{\Delta} \subseteq \boldsymbol{\Gamma}$ je stabilní extenzí množiny $\boldsymbol{\Gamma}$, pokud pro každý argument $\Phi \in \boldsymbol{\Gamma}$ platí, že $\Phi \in \boldsymbol{\Delta}$ právě tehdy, když neexistuje argument $\Psi \in \boldsymbol{\Delta}$ takový, že $\Psi \rightarrow \Phi$.

Dung se navíc věnuje slabším typům extenzí, jejich vzájemným vztahům a podmínkám jejich existence a dále souvislosti argumentačních rámců s logickým programováním a nemonotónním usuzováním. Pro případné zájemce o hlubší poznání dané problematiky lze proto doporučit kromě původního Dungova článku [5] i vynikající přehled, který napsal M. Caminada [4]. V Caminadově úvodu jsou navíc uvedeny i pozoruhodné důkazy, které dávají do souvislosti některé sémantiky s existenci vítězné strategie pro proponenta v jisté dialogické hře. Tyto výsledky jsou z hlediska modelování poznání zajímavé především z toho důvodu, že aproximují neformální procedury, jimiž jsou argumenty v přirozené argumentaci obvykle vyhodnocovány.

Nás ale bude zajímat především vnitřní struktura argumentů, založených na logických vztazích mezi znalostmi či přesvědčeními nějakého agenta, případně skupiny agentů.

3. Logická argumentace

Pro specifikaci logických vztahů mezi znalostmi je třeba vzít do úvahy jazyk, v němž jsou dané znalosti reprezentovány a jeho strukturu. Známe-li strukturu jazyka, můžeme popsat relaci vyplývání mezi jednotlivými prvky a na základě znalosti relace vyplývání pak můžeme definovat pojem logického argumentu a následně i vztahy mezi argumenty.

Jazyk je proto nejen prostředkem k zachycování znalostí, ale i médiem aktivně se podílejícím na jejich utváření, obměně, aktualizaci a revizi, neboť jeho struktura je zásadní pro určení možných konstelací v procesu dialogu.

V našem příspěvku se zabýváme jen velmi jednoduchým jazykem výrokové logiky, který se skládá z (nekonečné) množiny výrokových atomů, které budeme značit malými písmeny latinské abecedy {p,q,r...}, a výrokových formulí, které budeme značit velkými písmeny latinské abecedy {A,B,C...} a jež jsou utvořeny kombinacemi z výrokových atomů za pomoci logických spojek negace $\neg$ a konjunkce $\wedge$ (a závorek).

Definice 3.1. (Výroková formule)
Množina výrokových formulí je nejmenší množina obsahující všechny výrokové atomy a obsahující s každou výrokovou formulí A i její negaci $\neg$A, a pro libovolné dvě formule A, B i jejich konjunkci (A $\wedge$ B).

Dále ještě definujeme pojmy pravdivostního ohodnocení, konzistence, vyplývání, argumentu a kanonického protiargumentu.

Definice 3.2. (Pravdivostní ohodnocení)
Pravdivostním ohodnocením výrokových formulí (pro klasickou logiku) budeme mínit libovolnou funkci z množiny všech výrokových formulí do množiny {0,1} splňující podmínky:

1) Hodnota výrokové formule $\neg$A je 1, pouze v případě, že hodnota formule A je 0, v opačném případě 0.
2) Hodnota výrokové formule (A $\wedge$ B) je 1, pouze v případě, že hodnota formulí A i B je 1, v opačném případě 0.

Definice 3.3. (Vyplývání a konzistence)
Řekneme, že množina výrokových formulí $\boldsymbol{A}$ je konzistentní, pokud existuje pravdivostní ohodnocení, které všem formulím z množiny $\boldsymbol{A}$ přiřadí hodnotu 1.
Výroková formule B vyplývá z množiny formulí $\boldsymbol{A}$ (píšeme $\boldsymbol{A} \models$ B), pokud každé pravdivostní ohodnocení, které přiřadí hodnotu 1 všem formulím z $\boldsymbol{A}$, přiřadí hodnotu 1 i formuli B.

Definice 3.4. (Argument)
Logickým argumentem budeme mínit uspořádanou dvojici $<\boldsymbol{A}, B>$, kde množina $\boldsymbol{A}$ je minimální konzistentní množina taková, že $\boldsymbol{A} \models$ B. Množinu $\boldsymbol{A}$ budeme nazývat množinou premis argumentu a formuli B jeho závěrem.

Z kompaktnosti klasické výrokové logiky víme, že tuto minimální konzistentní množinu premis $\boldsymbol{A}$ můžeme vždy volit jako konečnou pro libovolnou formuli B. To nám dovoluje každý logický argument psát ve tvaru $<\{A_1, \dots , A_n\}, B>$.

Definice 3.5. (Kanonický protiargument)
Kanonickým protiargumentem proti argument $<\{A_1, \dots , A_n\}, B>$ budeme mínit libovolný argument tvaru $<\{C_1, \dots , C_m\}, \neg(A_1 \wedge \dots \wedge A_n)>$.

Poslední dvě výše uvedené definice jsou převzaty z článku [3].
Nyní máme vše potřebné, abychom mohli definovat argumentační rámec založený na množinách výrokových formulí, představujících přesvědčení, či znalosti agentů.

Definice 3.6. (Množina argumentů generovaná množinou přesvědčení $\boldsymbol{A}$)
Nechť $\boldsymbol{A}$ je množina výrokových formulí, reprezentujících znalosti, či přesvědčení nějakého agenta.
Množinou argumentů generovaných množinou $\boldsymbol{A}$, budeme mínit množinu $\{<\boldsymbol{A}', B>;$ kde $<\boldsymbol{A}', B>$ je argument a $\boldsymbol{A}' \subseteq \boldsymbol{A}\}$, a budeme ji značit $[\boldsymbol{A}]$.

Povšimněme si, že narozdíl od přístupu běžného při reprezentaci znalostí pomocí epistemických či doxastických logik, nepředpokládáme předem žádné logické vlastnosti přesvědčení agenta, jako je např. bezespornost, či uzavřenost na logické vyplývání.

Tyto obvyklé charakteristiky racionálního přesvědčení chceme nahradit charakteristikami založenými na vlastnostech množin argumentů, které tato množina generuje v kontextu nějakého argumentačního rámce. Tím se vyhneme problémům jako je 'logická vševědoucnost' agentů.

Množina argumentů, které libovolný agent může generovat je sice vždy nekonečná a v jistém smyslu logicky uzavřená, což však nepředstavuje žádný problém, protože zde nereprezentuje množinu argumentů, které si agent explicitně uvědomuje, ale množinu potenciálních argumentů, které může případně vygenerovat a použít během nějaké argumentace, bude-li jich třeba.

Definujme nyní argumentační rámec reprezentující veškerou možnou logickou argumentaci mezi *n* agenty.

<u>Definice 3.7.</u>
Nechť $\boldsymbol{A}_1 \ldots \boldsymbol{A}_n$ jsou množiny formulí, reprezentujících znalosti, případně přesvědčení *n* agentů.
Logickým argumentačním rámcem pro *m*-agentů budeme rozumět dvojici $<\boldsymbol{\Phi}\ ;\ \Rightarrow>$, kde $\boldsymbol{\Phi} = [\boldsymbol{A}_1] \cup \ldots \cup [\boldsymbol{A}_n]$ a kde pro každé dva argumenty Φ, Ψ z množiny $\boldsymbol{\Phi}$ platí $\Phi \Rightarrow \Psi$, právě tehdy když argument Φ je kanonický protiargument proti argumentu Ψ.

4. Specifika a problémy logické argumentace

Použijeme-li pro definici argumentu a protiargumentu relaci vyplývání a spojky konjunkce a negace klasické výrokové logiky, které splňují jisté vlastnosti, vzniklý argumentační rámec bude mít poměrně jednoduchou strukturu a budou v něm platit následující tvrzení.

<u>Tvrzení 4.1.</u>
Pro libovolný argument $\Phi = <\{A_1, \ldots, A_n\}, B>$ bude existovat nějaký argument $X = <\{C_1, \ldots, C_m\}, D>$, který brání Φ a platí $\{C_1, \ldots, C_m\} \models A_i$ pro všechna $1 \leq i \leq n$.
<u>Důkaz:</u>
Buď $\Phi = <\{A_1, \ldots, A_n\}, B>$ libovolný argument a $\Psi = <\{C_1, \ldots, C_m\}, D>$ kanonický protiargument vůči Φ. Pak podle definice $D = \neg(A_1 \wedge \ldots \wedge A_n)$ a zároveň $C_1, \ldots, C_m \models D$.
V klasické logice odvodíme, že $A_1, \ldots, A_n \models \neg(C_1 \wedge \ldots \wedge C_m)$.

Vezměme nyní minimální konzistentní množinu $\{E_1, \dots , E_k\}$, takovou že platí $E_1, \dots , E_k \models A_i$ pro všechna $1 \leq i \leq n$. (Existuje, jelikož $\{A_1, \dots , A_n\}$ sama je konzistentní, neboť je množinou premis argumentu Φ.)
V klasické logice pak bude platit $E_1, \dots , E_k \models \neg(C_1 \wedge \dots \wedge C_m)$, dle našich definic je $X = <\{E_1, \dots , E_k \}, \neg(C_1 \wedge \dots \wedge C_m)>$ kanonický protiargument proti Ψ. Tudíž X brání Φ.

Nad klasickou logikou je tedy snadné vymyslet pro libovolný argument jiný, který ho brání, a navíc je silnější v tom smyslu, že z jeho premis vyplývají všechny premisy bráněného argumentu.

Tvrzení 4.2.
Nechť $<\boldsymbol{\Phi}; \Rightarrow>$ je logický argumentační rámec.
Pak platí, že pro každou $\boldsymbol{C}$ maximální konzistentní podmnožinu množiny $\underline{\boldsymbol{\Phi}} = \cup\{\{A_1, \dots , A_n\}; <\{A_1, \dots , A_n\}, B> \in \boldsymbol{\Phi}\}$, je $[\boldsymbol{C}]$ stabilní extenzí.
Důkaz:
Nechť $\boldsymbol{C}$ je maximální konzistentní podmnožina $\underline{\boldsymbol{\Phi}}$ a $\Phi \in \boldsymbol{\Phi}$.
Dokážeme, že $\Phi \in [\boldsymbol{C}]$, právě tehdy když neexistuje $\Psi \in [\boldsymbol{C}]$, že $\Psi \Rightarrow \Phi$.
Nechť nejprve $\Phi \in [\boldsymbol{C}]$, $\Phi = <\{A_1, \dots , A_n\}, B>$. Předpokládejme dále, že existuje $\Psi = <\{C_1, \dots , C_m\}, D> \in [\boldsymbol{C}]$, že $\Psi \Rightarrow \Phi$, čili $\{C_1, \dots , C_m\} \models D$ a $D = \neg(A_1 \wedge \dots \wedge A_n)$.
To znamená, že množina $\{C_1, \dots , C_m\} \cup \{A_1, \dots , A_n\}$ není konzistentní. Ale protože jsme brali $\Phi, \Psi \subseteq [\boldsymbol{C}]$, $\{C_1, \dots , C_m\} \cup \{A_1, \dots , A_n\} \subseteq \boldsymbol{C}$ a proto množina $\boldsymbol{C}$ není konzistentní, jak jsme předpokládali.
Nechť $\Phi \in [\boldsymbol{A}] \setminus \boldsymbol{\Phi}$, $\Phi = <\{A_1, \dots , A_n\}, B>$. Protože jsme brali $\Phi \notin [\boldsymbol{C}]$, pak existuje $1 \leq i \leq n$, že $A_i \in \boldsymbol{\Phi} \setminus \boldsymbol{C}$. Avšak $\boldsymbol{C}$ je maximálně konzistentní podmnožina $\boldsymbol{\Phi}$, tudíž $\boldsymbol{C} \cup \{A_i\}$ je nekonzistentní.
Z kompaktnosti klasické výrokové logiky nyní plyne, že existuje konečná $\boldsymbol{C'} \subseteq \boldsymbol{C}$ řekněme $\{C_1, \dots , C_m\}$, že už $\boldsymbol{C'} \cup \{A_i\}$ je nekonzistentní. Tudíž platí, že $C_1, \dots , C_m \models \neg A_i$. Avšak dále platí $\neg A_i \models \neg(A_1 \wedge \dots \wedge A_n)$, tudíž i $C_1, \dots , C_m \models \neg(A_1 \wedge \dots \wedge A_n)$. Zvolme nyní $\boldsymbol{C''}$ minimální podmnožinu $\boldsymbol{C'}$, mající uvedené vlastnosti.
Protože $\boldsymbol{C''} \subseteq \boldsymbol{C'} \subseteq \boldsymbol{C}$ a $\boldsymbol{C}$ je konzistentní, je $\boldsymbol{C''}$ konzistentní.
Tudíž $\Psi = <\boldsymbol{C''}, \neg(A_1 \wedge \dots \wedge A_n)>$ je z definice argument a z definice $[\boldsymbol{C}]$ plyne, že $\Psi \in [\boldsymbol{C}]$. Z definice $\Rightarrow$ plyne, že $\Psi \Rightarrow \Phi$.

Otevřenou otázkou nadále zůstává, zda každá stabilní extenze logického argumentačního rámce bude generována nějakou maximální konzistentní podmnožinou sjednocení znalostí všech agentů. Máme-li logický argumentační rámec $<\boldsymbol{\Phi}; \Rightarrow>$, pak je tedy otázkou, zda pro každou stabilní extenzi $\Sigma \subseteq \boldsymbol{\Phi}$ existuje maximální konzistentní $\boldsymbol{C} \subseteq \underline{\boldsymbol{\Phi}}$ taková, že $\Sigma = [\boldsymbol{C}]$.

Zdá se, že pro logické argumentační rámce pro jednoho agenta ještě toto platit bude, ale je jisté, že pro větší počet agentů již nikoliv. Např. v rámci $<[p] \cup [\neg(p \wedge q)] \cup [q]; \Rightarrow>$ je množina $\{p, q\}$ maximální konzistentní podmnožina sjednocených znalostí všech agentů, nicméně argument $<\{p,q\}, p \wedge q>$ generovaný touto množinou není ani v jedné z množin argumentů generovaných znalostmi jednotlivých agentů.

Dalším okruhem otázek, jimiž je možné se zabývat jsou otázky, zda by nebylo vhodné použít pro definici argumentů nějaké logiky, jejichž relace vyplývání nesplňuje všechny strukturní vlastnosti relace vyplývání logiky klasické (substrukturální logiky), nebo případně nějaké logiky se slabší negací (intuicionistická, fuzzy), kde by nemuselo nutně platit ani jedno z předchozích tvrzení a konstrukce extenzí by byla daleko méně triviální záležitostí.

Bibliografie

[1] ANTONIOU, G. *Nonmonotonic Reasoning*. Cambridge: MIT Press: London. 1996.
[2] BENTHEM, J. van. *Exploring logical dynamics.* Stanford: CSLI Publications: FOLLI. 1996.
[3] BESNARD, P. - HUNTER, A. A logic-based theory of deductive arguments. *Artificial Inteligence* 128. Amsterdam: Elsevier. 2001. s. 203–235
[4] CAMINADA, M. *A Gentle Introduction to Argumentation Semantics.* http://icr.uni.lu/~martinc/publications/Semantics_Introduction.pdf. Citováno 25.4.2008
[5] DUNG, P. M. On the acceptability of argumetns and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Inteligence* 77. Amsterdam: Elsevier. 1995, s. 321–357