



Logici pentru argumentare

Teza prezintă metode prin care agenții software pot rezolva conflicte și lua decizii utilizând argumentare. Am investigat expresivitatea diferitelor logici pentru a formaliza argumentarea între agenți software. Logicile utilizate sunt: logica defezabilă (capitolul 2), logica de descriere (capitolul 3), logica de descriere fuzzy (capitolul 4), logica subiectivă (capitolul 5), logica justificativă (capitolul 5), logici hibride (capitolul 6) sau logica deontică (capitolul 8).

Capitolul 2 prezintă o soluție de rezolvare a conflictelor în ansambluri de clasificatori. Soluția combină învățarea computațională cu argumentarea în logica defezabilă. Pentru învățare s-a utilizat un ansamblu de trei clasificatori: un arbore de decizie, o rețea neurală și o mașină cu suport vectorial. Dacă cei trei clasificatori nu indică o decizie de clasificare unică, opiniile clasificatorilor sunt date unui proces de argumentare care încearcă rezolvarea conflictelor utilizând logica defezabilă. Pentru aceasta e nevoie ca deciziile de clasificare să fie exprimate prin reguli extrase din cei trei clasificatori. Sistemul argumentativ beneficiază și de reguli expert pentru clasificare. Metoda propusă este superioară metodei de rezolvare a conflictelor bazată pe votare. Experimentele pe un scenariu de clasificare a culturilor agricole din imagini satelitare au indicat importanța cunoștințelor expertului (formalizate ca reguli) în performanța algoritmului. Am dezvoltat o metodă de extragere de reguli defezabile din clasificatori, pentru a facilita integrarea cunoștințelor agentului uman. Avantajele unei mașinării argumentative peste ansambluri de clasificatori sunt: i) argumentele permit adăugare de cunoștințe expert în procesul de clasificare; ii) experimentele indică că argumentarea poate înlocui metodele bazate pe votare pentru rezolvarea conflictelor; iii) argumentarea crește transparența în luarea deciziei. Metoda propusă este utilă la luarea deciziilor în situații caracterizate prin disponibilitatea datelor de antrenament, existența unor cunoștințe expert și mai ales când e necesară justificarea deciziei luate.

Problema abordată în capitolul 3 este în ce măsură se poate distinge între argumente și explicații în dialog. Fiind dată întreteserea de argumente și explicații în limbaj natural, contribuțiile abordării din acest capitol sunt: (i) identificarea indicatorilor lingvistici care semnalizează argumente sau explicații; (ii) modelarea argumentelor și explicațiilor sub aceeași umbrelă a ontologiei *ArgExp*; (iii) modelarea perspectivelor subiective ale agenților cu privire la tipul de dialog în care se afla: argumentativ sau explicativ; (iv) formalizarea actelor comunicative specifice pentru transmiterea de argumente sau explicații. Prin utilizarea raționării în logica de descriere, sistemul poate

clasifica argumente sau explicații pe baza adnotărilor parțiale ale utilizatorilor. Principalul câștig este ca agenții identifică mai rapid situațiile de consens sau cele conflictuale în timpul dialogului.

Capitolul 4 propune Logica de Descriere Fuzzy ca mecanism pentru acoperirea distanței dintre argumentele agentului uman și cele ale agentului software. Argumentele în limbaj natural sunt o mixtură de variabile și cunoștințe fuzzy. Capitolul se focalizează pe modelarea argumentelor imprecise pentru o mai bună cooperare a agenților software cu cei umani în scenarii argumentative. Scopul este de a extinde scenariile practice ale argumentării prin includerea agentului uman. Notăm legăturile dintre rationarea fuzzy (semantica Lukasiewicz sau Godel) și aspecte din teoria argumentării (e.g., principiul verigii celei mai slabe, agregarea argumentelor). Scenariul exemplifică și scheme de argumentare plauzibile. Abordarea fuzzy facilitează calculul puterii unui argument sau contraargument în timpul dezbaterii. De asemenea, se pot semnaliza situațiile în care cunoștințele fuzzy sunt inconsistente cu bugetul de conflict tolerat de către părți.

Capitolul 5 analizează argumente postate în arena publică. Pornim de la logica subiectivă pentru a răspunde la o serie de întrebări: Sunt persoanele dintr-o anumită comunitate apriori inclinate să accepte sau să respingă un subiect controversat? Care subiecte de dezbatere sunt cele mai acceptate, respinse, populare într-o comunitate? Este utilizat un vocabular diferit între argumente pozitive și contra argumente? Sunt argumentele pozitive sau cele negative mai ușor de înțeles de către audiență? Există corelație între lungimea topicului dezbaterii și popularitatea lui? În ce măsură formularea dezbaterii ca interogare, respectiv afirmație influențează gradul de interes în dezbatere? Pentru aceste întrebări s-a dezvoltat unealta ARGSENSE care analizează argumentele publice legate de încălzirea globală. Cele patru contribuții sunt: 1) un corpus de argumente legate de încălzirea globală etichetat; 2) unealta ARGSENSE care facilitează înțelegerea opiniilor agenților cu privire la schimbările climatice; 3) o metodă computațională pentru agregarea și consolidarea argumentelor, bazată pe deducere textuală; 4) o analiza lexicală a corpusului de argumente. În primul rând, corpusul este (conform cunoștințelor noastre) cea mai mare colecție de argumente etichetate legate de încălzirea globală. În al doilea rând, ARGSENSE furnizează rapoarte pentru realizatorii de politici legate de modalitatea de comunicare în domeniul schimbărilor climatice. În al treilea rând, ARGSENSE este util în științe sociale, acestea având nevoie de instrumente care pot măsura automat tendințele sociale. Metoda propusă ajută la măsurarea suportului, dezacordului sau ignoranței unor comunități cu privire la subiecte de pe agenda publică. Pentru aceasta folosim deducere textuală pentru identificarea similarității, contradicției sau subsumarea între argumente în limbaj

natural. În al patrulea rând, cei interesați în informarea efectivă a publicului trebuie să țină cont de anumite aspecte lingvistice în comunicarea informațiilor științifice. De exemplu, analiza lexicală oferită de ARGSENSE în domeniul încălzirii globale a descoperit următoarele: 1) cei care recunosc încălzirea globală sunt interesați de *emisii*, *gaze de seră* sau *apa*, spre deosebire de oponenții ideii de schimbare climatică care sunt interesați de *gheață* de exemplu; 2) subiectele de dezbatere formulate ca afirmații au față de cele formulate ca interogări de 3 ori mai multe șanse de a fi câștigătoare ($p < 0.0001$); 3) argumentele pozitive și cele negative au aceiași indici de complexitate a textului; 4) lungimea unui subiect de dezbatere nu influențează numărul de argumente pro și contra pe care le atrage; 5) subiectele la forma interogativă au mai multe șanse de a atrage argumente decât cele la forma afirmativă; Astfel de analize lexicale pot fi utilizate de experți pentru a comunica mai eficace subiecte științifice către public.

În capitolul 6 extind logica de justificare cu argumente și explicații, rezultând un nou formalism: Logica Argumentării și Explicației (\mathcal{AEL}). \mathcal{AEL} este aplicată pentru etichetarea cooperativă a argumentelor de către agenți cu perspective și cunoștințe diferite. Contribuțiile acestui capitol se referă la formalizarea \mathcal{AEL} pentru a distinge între argumente și explicații, respectiv la dezvoltarea unui model computațional pentru etichetarea grafurilor de argumente. \mathcal{AEL} permite formalizarea și raționarea pe meta-argumente, scopul fiind cel de a facilita înțelegerea de către agentul software a situației curente, precum și creșterea eficienței în comunicare.

Capitolul 7 prezintă un sistem de decizie pentru vehicule aeriene autonome. Am combinat argumentarea cu verificarea modelului în evaluarea deciziilor care trebuie luate în cazul informațiilor incomplete și inconsistente. Comportamentul dronei este modelat cu structuri Kripke. Deciziile sunt luate folosind verificare formală cu logici hibride și logica defezabilă. În cazul schimbărilor legislative cu privire la siguranța zborului, modelul curent poate eșua în a verifica noile constrângeri. În acest caz, agentul își inițiază un proces argumentativ prin care încearcă să își modifice modelul comportamental. Modelul rezultat în urma argumentării este supus verificării formale pentru a se asigura că respectă noile norme. Ipoteza de cercetare a fost că argumentarea poate fi utilizată în siguranța sistemelor complexe, oferind un mijloc de a ajunge la decizii raționale justificate.

Pentru a verifica conformitatea unor procese de business cu normele în vigoare, capitolul 8 propune o logica normativă temporală peste logicile de descriere. Scopul este de facilita verificarea normelor abstracte pe procese de business specifice. Demonstrarea conceptului s-a făcut pe standardul Hazard Analysis at Critical Control Points (HACCP) care urmărește prevenirea hazardurilor în industria alimentară.