1 Formalizarea afirmațiilor

Traducem în logică următorul exercițiu, luat din `http://www.cs.utexas.edu/users/novak/reso.html` .

1. Fiecare investitor a cumpărat acțiuni sau obligațiuni.
2. Dacă indicele Dow Jones scade, toate acțiunile mai puțin aurul scad.
3. Dacă trezoreria crește dobânda, toate obligațiunile scad.
4. Orice investitor care a cumpărat ceva care scade nu e bucuros.
5. Dacă indicele Dow Jones scade și trezoreria crește dobânda, toți investitorii bucuroși au cumpărat ceva acțiuni de aur.

1. Fiecare investitor a cumpărat acțiuni sau obligațiuni.

În logica predicatelor, variabilele reprezintă elemente *arbitrare*, de orice fel, din univers. Nu putem spune "în această formulă, am ales X să fie un investitor", fiindcă X poate fi orice. Pentru a reprezenta categorii (tipuri) de entități, folosim *predicate*. Introducem predicatul $inv(X)$ (X e investitor).

Cuvintele "fiecare", "orice", "toți", etc. introduc o variabilă cuantificată universal: $\forall X \ldots$ Oricum alegem X *din univers*, formula cuantificată e adevărată. X poate fi de orice fel (investitor, elev, casă, număr, etc.), dar fraza e despre investitori, deci spune ceva despre X doar *dacă* X ales e investitor. Din acest motiv, cuantificatorul universal apare de regulă cu implicația:

Pentru *orice* X, *dacă* X e investitor, a făcut ceva

$$\forall X.inv(X) \rightarrow \boxed{ce\ se\ spune\ despre\ X}$$

Ce știm despre investitor? A cumpărat *ceva*, deci *există* ceva ce investitorul a cumpărat. Când definim un predicat, folosim argumentele în ordinea uzuală din propoziție (subiect, apoi complement). Deci $cump(X, Y)$: X a cumpărat Y.

$$\forall X.inv(X) \rightarrow \exists Y.cump(X, Y) \wedge \boxed{ce\ știm\ despre\ Y}$$

A_1: $\forall X.inv(X) \rightarrow \exists Y.cump(X, Y) \wedge (act(Y) \vee oblig(Y))$

2. Dacă indicele Dow Jones scade, toate acțiunile mai puțin aurul scad.

Indicele Dow Jones e o noțiune unică, îl reprezentăm deci printr-o *constantă dj*. Rafinăm și aici succesiv:

$$scade(dj) \rightarrow \boxed{\text{ce se întâmplă}}$$

$$scade(dj) \rightarrow \forall X. \boxed{condiții\ pentru\ X} \rightarrow scade(X)$$

A_2: $scade(dj) \rightarrow \forall X.act(X) \wedge \neg aur(X) \rightarrow scade(X)$

3. Dacă trezoreria crește dobânda, toate obligațiunile scad.

"Trezoreria crește dobânda" e o propoziție cu subiect, predicat și complement, deci sugerează introducerea unui predicat. Dar propoziția mai apare identic în fraza 5, iar verbul "crește" apare doar aici, deci nu mai sunt alte entități care pot "crește". Deci putem reprezenta "trezoreria crește dobânda" ca o *propoziție*: ea e fie adevărată, fie falsă, nu se referă la vreo altă variabilă.

A_3: $crdob \rightarrow \forall X.oblig(X) \rightarrow scade(X)$

4. Orice investitor care a cumpărat ceva care scade nu e bucuros.

$$\forall X.inv(X) \rightarrow \boxed{ce\ știm\ despre\ X}$$

$$\forall X.inv(X) \rightarrow (\boxed{condiție\ pentru\ X} \rightarrow \neg bucuros(X))$$

$$\forall X.inv(X) \rightarrow (\exists Y.cump(X, Y) \wedge scade(Y)) \rightarrow \neg bucuros(X)$$

Avem o structură cu două implicații, de forma $A \rightarrow (B \rightarrow C)$. O vom întâlni scrisă și fără paranteze, deoarece convențional, implicația e asociativă la dreapta.

Alternativ, știm că $p \to (q \to r) = \neg p \vee (\neg q \vee r) = (\neg p \vee \neg q) \vee r = \neg(p \wedge q) \vee r = (p \wedge q) \to r$, deci putem rescrie cu o conjuncție. Reamintim că e prioritară față de $\to$, deci $A \wedge B \to C = (A \wedge B) \to C$.

A_4: $\forall X.inv(X) \wedge (\exists Y.cump(X,Y) \wedge scade(Y)) \to \neg bucuros(X)$

Tautologia $A \to (B \to C) = (A \wedge B) \to C$ are o analogie și în programare: `if (A) if (B) C;` e echivalent cu `if (A && B) C;`

5. Dacă indicele Dow Jones scade și trezoreria crește dobânda, toți investitorii bucuroși au cumpărat ceva acțiuni de aur.

$scade(dj) \wedge crdob \to \boxed{\textit{ce se întâmplă}}$

$scade(dj) \wedge crdob \to \forall X.inv(X) \wedge bucuros(X) \to \boxed{\textit{ce știm despre } X}$

C: $scade(dj) \wedge crdob \to \forall X.inv(X) \wedge bucuros(X) \to \exists Y.cump(X,Y) \wedge act(Y) \wedge aur(Y)$

2 Aducerea la formă clauzală (formă normală conjunctivă)

Am rescris formulele cu paranteze în loc de punct pentru a evita neînțelegeri și greșeli la cuantificatori.

A_1: $\forall X(inv(X) \to \exists Y(cump(X,Y) \wedge (act(Y) \vee oblig(Y))))$

A_2: $scade(dj) \to \forall X(act(X) \wedge \neg aur(X) \to scade(X))$

A_3: $crdob \to \forall X(oblig(X) \to scade(X))$

A_4: $\forall X(inv(X) \wedge \exists Y(cump(X,Y) \wedge scade(Y)) \to \neg bucuros(X))$

C: $scade(dj) \wedge crdob \to \forall X(inv(X) \wedge bucur(X) \to \exists Y(cump(X,Y) \wedge act(Y) \wedge aur(Y)))$

Pentru a demonstra prin reducere la absurd, arătăm că ipotezele A_1–A_4 împreună cu *negația* concluziei $\neg C$ duc la o contradicție. $\boxed{\text{Negăm concluzia } \textit{la început}\text{, înainte de a transforma cuantificatorii!}}$

$\neg C$: $\neg(scade(dj) \wedge crdob \to \forall X(inv(X) \wedge bucur(X) \to \exists Y(cump(X,Y) \wedge act(Y) \wedge aur(Y))))$

Parcurgem aceiași pași ca și pentru formulele propoziționale, cu pași suplimentari specifici predicatelor.

1a. Eliminăm implicația $A \to B = \neg A \vee B$, $\neg(A \to B) = A \wedge \neg B$
1b. Ducem negația înăuntru până la predicate $\neg\forall x P(x) = \exists x \neg P(x)$ $\neg\exists x P(x) = \forall x \neg P(x)$

Cei doi pași se pot aplica împreună, parcurgând formula din *exterior* (ca la formulele propoziționale), adică începând cu conectorii din afara parantezelor spre interior.

Evident, orice transformare într-o *subformulă* nu afectează ce e în afara ei. În particular, cuantificatorii se schimbă *numai* când ducem negația înăuntru (când negația trece peste cuantificator, îl schimbă), nu și când facem alte transformări (implicație, negație, etc.) *înăuntrul* formulei cuantificate.

A_1: $\forall X(inv(X) \to \exists Y(cump(X,Y) \wedge (act(Y) \vee oblig(Y))))$
$\forall X(\neg inv(X) \vee \exists Y(cump(X,Y) \wedge (act(Y) \vee oblig(Y))))$

A_2: $scade(dj) \to \forall X(act(X) \wedge \neg aur(X) \to scade(X))$
$\neg scade(dj) \vee \forall X(\neg act(X) \vee aur(X) \vee scade(X))$

A_3: $crdob \to \forall X(oblig(X) \to scade(X))$
$\neg crdob \vee \forall X(\neg oblig(X) \vee scade(X))$

A_4: $\forall X(inv(X) \wedge \exists Y(cump(X,Y) \wedge scade(Y)) \to \neg bucuros(X))$
$\forall X(\neg(inv(X) \wedge \exists Y(cump(X,Y) \wedge scade(Y))) \vee \neg bucuros(X))$
$\forall X(\neg inv(X) \vee \neg\exists Y(cump(X,Y) \wedge scade(Y)) \vee \neg bucuros(X))$
$\forall X(\neg inv(X) \vee \forall Y \neg(cump(X,Y) \wedge scade(Y)) \vee \neg bucuros(X))$
$\forall X(\neg inv(X) \vee \forall Y(\neg cump(X,Y) \vee \neg scade(Y)) \vee \neg bucuros(X))$

$\neg C$: $\neg\,(scade(dj) \wedge crdob \rightarrow \forall X(inv(X) \wedge bucur(X) \rightarrow \exists Y(cump(X,Y) \wedge act(Y) \wedge aur(Y))))$
$scade(dj) \wedge crdob \ \wedge\ \neg \forall X(inv(X) \wedge bucur(X) \rightarrow \exists Y(cump(X,Y) \wedge act(Y) \wedge aur(Y)))$
$scade(dj) \wedge crdob \ \wedge \exists X \neg\,(inv(X) \wedge bucur(X) \rightarrow \exists Y(cump(X,Y) \wedge act(Y) \wedge aur(Y)))$
$scade(dj) \wedge crdob \ \wedge \exists X(inv(X) \wedge bucur(X) \wedge\ \neg \exists Y(cump(X,Y) \wedge act(Y) \wedge aur(Y)))$
$scade(dj) \wedge crdob \ \wedge \exists X(inv(X) \wedge bucur(X) \wedge \forall Y \neg\,(cump(X,Y) \wedge act(Y) \wedge aur(Y)))$
$scade(dj) \wedge crdob \ \wedge \exists X(inv(X) \wedge bucur(X) \wedge \forall Y(\neg cump(X,Y) \vee \neg act(Y) \vee \neg aur(Y)))$

2. Dăm variabilelor cuantificate nume unice în fiecare formulă

O variabilă cuantificată are înțeles *doar* în interiorul subforumulei cuantificate (așa cum în program numele de parametri au sens doar în interiorul funcției). Deci putem scrie $\forall x P(x) \vee \forall x Q(x)$ fără ambiguitate. Vom elimina însă cuantificatorii, și atunci ne trebuie nume unice. E suficient să dăm nume unice variabilelor cuantificate cu $\forall$; cele cuantificate cu $\exists$ vor fi înlocuite prin skolemizare.

Astfel, în $\forall x\, P(x) \vee \forall x\, \exists y\, Q(x,y)$ am înlocui al doilea x cu un nou nume, z: $\forall x\, P(x) \vee \forall z\, \exists y\, Q(z,y)$. Nu avem asemenea cazuri în problema noastră.

A_1: $\forall X(\neg inv(X) \vee \exists Y(cump(X,Y) \wedge (act(Y) \vee oblig(Y))))$

A_2: $\neg scade(dj) \vee \forall X(\neg act(X) \vee aur(X) \vee scade(X))$

A_3: $\neg crdob \vee \forall X(\neg oblig(X) \vee scade(X))$

A_4: $\forall X(\neg inv(X) \vee \forall Y(\neg cump(X,Y) \vee \neg scade(Y)) \vee \neg bucuros(X))$

$\neg C$: $scade(dj) \wedge crdob \ \wedge \exists X(inv(X) \wedge bucur(X) \wedge \forall Y(\neg cump(X,Y) \vee \neg act(Y) \vee \neg aur(Y)))$

3. Eliminăm cuantificatorii existențiali prin skolemizare

Într-o formulă $\forall x_1 \ldots \forall x_n \exists y\, A$, oricum alegem $x_1, \ldots, x_n$ *există* un y care face A adevărat, dar alegerea lui y *depinde* de $x_1, \ldots, x_n$ aleși înainte. Deci y este o *funcție* f de $x_1, \ldots, x_n$, chiar dacă afirmația nu ne spune care e funcția. Cuantificatorul $\exists y$ dispare: pentru orice $x_1, \ldots, x_n$, formula A e adevărată înlocuind y cu $f(x_1, \ldots, x_n)$. Introducem câte o astfel de *funcție Skolem*, cu nume nou, ales diferit pentru fiecare cuantificator existențial (regulile după care depind variabilele fiind diferite).

Ca un caz particular, dacă $\exists$ apare *în afara* oricărui $\forall$, variabila cuantificată existențial *nu depinde* de nimic altceva. Funcția Skolem nu are parametri, deci devine o nouă *constantă*.

În problema noastră, avem cuantificatori existențiali în A_1 și $\neg C$:

A_1: $\forall X(\neg inv(X) \vee \boldsymbol{\exists Y}\, (cump(X,Y) \wedge (act(Y) \vee oblig(Y))))$

Y din $\exists$ depinde de X. Alegem o nouă funcție Skolem f. Înlocuim Y cu $f(X)$, iar $\exists Y$ dispare:

A_1: $\forall X(\neg inv(X) \vee (cump(X,f(X)) \wedge (act(f(X)) \vee oblig(f(X)))))$

În $\neg C$, $\exists X$ e în afara oricărui $\forall$. Înlocuim X cu o nouă constantă Skolem b, iar $\exists X$ dispare:

$\neg C$: $scade(dj) \wedge crdob \wedge \boldsymbol{\exists X}\, (inv(X) \wedge bucur(X) \wedge \forall Y(\neg cump(X,Y) \vee \neg act(Y) \vee \neg aur(Y)))$

$\neg C$: $scade(dj) \wedge crdob \wedge inv(b) \wedge bucur(b) \wedge \forall Y(\neg cump(b,Y) \vee \neg act(Y) \vee \neg aur(Y))$

4. Aducem cuantificatorii universali în față și îi eliminăm

Formulele obținute conțin doar $\forall$, $\wedge$, $\vee$, și negația $\neg$ aplicată doar direct la predicate. Având variabile cu nume unic, cuantificatorii universali pot fi mutați în față (intuitiv, putem “alege” valoarea variabilei de la început, nu contează dacă mutăm $\forall x$ și peste o subformulă care nu depinde de x). Aceasta se numește *forma normală prenex*. Apoi eliminăm cuantificatorii universali, considerându-i impliciți. Practic, pornind de la pasul anterior, putem șterge direct toți cuantificatorii universali din formule:

A_1: $\neg inv(X) \vee (cump(X,f(X)) \wedge (act(f(X)) \vee oblig(f(X))))$

A_2: $\neg scade(dj) \vee \neg act(X) \vee aur(X) \vee scade(X)$

A_3: $\neg crdob \vee \neg oblig(X) \vee scade(X)$

A_4: $\neg inv(X) \vee \neg cump(X,Y) \vee \neg scade(Y) \vee \neg bucuros(X)$

$\neg C$: $scade(dj) \wedge crdob \wedge inv(b) \wedge bucur(b) \wedge (\neg cump(b,Y) \vee \neg act(Y) \vee \neg aur(Y))$

5. Aducem la forma clauzală aplicând distribuitatea

Ducem *conjuncția* în *exteriorul disjuncției* prin distributivitate: $A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$. Scriem fiecare clauză (disjuncție de literali) separat. Obținem *forma clauzală* (forma normală conjunctivă).

A_1: $\neg inv(X) \vee (cump(X, f(X)) \wedge (act(f(X)) \vee oblig(f(X))))$
$(\neg inv(X) \vee cump(X, f(X))) \wedge (\neg inv(X) \vee act(f(X)) \vee oblig(f(X)))$

În exemplul dat, doar A_1 trebuie transformat. Am indicat la fiecare clauză formula din care provine.

(1) $\neg inv(X) \vee cump(X, f(X))$ — A_1

(2) $\neg inv(X) \vee act(f(X)) \vee oblig(f(X))$ — A_1

(3) $\neg scade(dj) \vee \neg act(X) \vee aur(X) \vee scade(X)$ — A_2

(4) $\neg crdob \vee \neg oblig(X) \vee scade(X)$ — A_3

(5) $\neg inv(X) \vee \neg cump(X, Y) \vee \neg scade(Y) \vee \neg bucuros(X)$ — A_4

(6) $scade(dj)$ — $\neg C$

(7) $crdob$ — $\neg C$

(8) $inv(b)$ — $\neg C$

(9) $bucuros(b)$ — $\neg C$

(10) $\neg cump(b, Y) \vee \neg act(Y) \vee \neg aur(Y)$ — $\neg C$

3 Demonstrația prin rezoluție

Generăm rezolvenți până la clauza vidă. Căutăm clauze cu predicate opuse, P și $\neg P$, unificăm argumentele și adăugăm *rezolventul* obținut (restul clauzelor, substituit și reunit cu $\vee$). Euristici posibile sunt să folosim clauze cu un singur literal, și să eliminăm pe rând predicatele, ca în calculul propozițional.

(11) $\neg act(X) \vee aur(X) \vee scade(X)$ — eliminăm $scade(dj)$ din (3, 6)

(12) $\neg oblig(X) \vee scade(X)$ — eliminăm $crdob$ din (4, 7)

(13) $\neg inv(b) \vee \neg cump(b, Y) \vee \neg scade(Y)$ — eliminăm $bucuros$ cu $X = b$ din (5, 9)

Variabilele au semnificație proprie în fiecare clauză. Înainte de a unifica în clauze cu variabile numite la fel, le redenumim într-una din clauze, de exemplu X cu X_{12} în (12) pentru a face rezolventul cu (2).

(14) $\neg inv(X) \vee act(f(X)) \vee scade(f(X))$ — eliminăm $oblig$ cu $X_{12} = f(X)$ din (2, 12)

În rezolvent putem obține de două ori același predicat. Putem simplifica, $P(X) \vee P(X) = P(X)$, doar dacă argumentele sunt *aceleași*, de ex. $\neg act(X)$ în rezolventul (15) sau $scade(f(X))$ în (16) mai jos:

(15) $\neg cump(b, X) \vee \neg act(X) \vee scade(X)$ — eliminăm aur cu $Y = X$ din (10, 11)

(16) $\neg inv(X) \vee scade(f(X)) \vee \neg cump(b, f(X))$ — eliminăm act cu $X_{15} = f(X)$ din (14, 15)

(17) $\neg inv(b) \vee \neg cump(b, f(b))$ — eliminăm $scade$ cu $X = b$, $Y = f(b)$ din (13, 16)

(18) $\neg inv(b)$ — eliminăm $cump$ cu $X = b$ din (1, 17)

(19) $\emptyset$ — eliminăm $inv(b)$ din (8, 18)

Am obținut clauza vidă, deci $A_1 \wedge A_2 \wedge A_3 \wedge A_4 \wedge \neg C$ nu e realizabilă, și am demonstrat afirmația pe care ne-am propus-o inițial: $A_1 \wedge A_2 \wedge A_3 \wedge A_4 \rightarrow C$.