



KLASZTEREZÉSI MÓDSZEREN ALAPULÓ FUZZY SZABÁLYBÁZIS-REDUKÁLÁS A FRIQ-LEARNING-RENDSZERBEN

TOMPA TAMÁS

Miskolci Egyetem

Informatikai Intézet

Általános Informatikai Intézeti Tanszék

tamas.tompai@uni-miskolc.hu

KOVÁCS SZILVESZTER

Miskolci Egyetem

Informatikai Intézet

Általános Informatikai Intézeti Tanszék

szilveszter.kovacs@uni-miskolc.hu

Absztrakt. A Fuzzy-logikán alapuló megerősítéses tanulási módszerek tudásábrázolási formája fuzzy szabályok alkotta fuzzy szabályrendszer által leírt. A tanulási folyamat után előállt szabálybázis szabályainak száma, azaz a tudásbázis mérete meghatározza a rendszer komplexitását, számítási igényét. Ennek következtében a szabálybázis méretének optimalizálása, az esetleges elhagyható szabályok kiszűrése kulcsfontosságú ezen tanulási módszerek hatékonyságának növelése érdekében. A túlzottan nagy méretű szabálybázis növelheti a számítási igényt és csökkentheti a döntéshozatali folyamat átláthatóságát. A cikk célja egy klaszterezési eljáráson alapuló fuzzy szabálybázis-redukációs módszer bemutatása és alkalmazása a FRIQ-learning megerősítéses tanulási rendszerben, amely a releváns szabályok kiemelésével lehetővé teszi a szabálybázis méretének csökkentését, miközben a rendszer döntéshozatali képessége változatlan marad.

Kulcsszavak: megerősítéses tanulás, *Q-learning*, Fuzzy *Q-learning*, tudásbázis-redukálás, klaszterezés

1. Bevezetés

A klaszterezés [11] egy olyan adatbányászati módszer [10], amely célja az adatok csoportosítása olyan módon, hogy a hasonló tulajdonságokkal rendelkező adatelemek egy klaszterbe kerüljenek, míg a jelentősen eltérőek külön csoportokat alkossanak. Mivel a klaszterezés nem igényel előre meghatározott címkéket (mint az osztályozás [5]), a nem felügyelt tanulási algoritmusok [3] közé tartozik, és kizárólag az adatok közötti hasonlóság alapján végzi el a csoportosítást. Az adatok közötti ha-

sonlóság mértéke jellemzően távolságfüggvényekkel [9] mérhető, a közel álló adatpontok hasonló tulajdonságokat mutatnak, míg a távol esők jelentősen különböznek. Az adatok közötti távolság mérésére többféle metrika létezik, például az Euklideszi-, Manhattan-, Minkowski- és Hamming-távolság [9]. A klaszterezési technikák számos területen alkalmazhatók, többek között a képfeldolgozásban, a piackutatásban, a bioinformatikában és az egészségtudományban [4] [8]. A különböző klaszterezési algoritmusok, például a K-means, a hierarchikus klaszterezés vagy a DBSCAN eltérő módszereket kínálnak az adatok szerkezetének feltárására és a rejtett mintázatok azonosítására [11].

A FRIQ-learning (Fuzzy Rule-Interpolation based Q-learning) [19] [21] és a HFRIQ-learning (Heuristically Accelerated Fuzzy Rule Interpolation based Q-learning) [15] [17] olyan fuzzy szabályinterpoláción [7] [13] alapuló megerősítéses tanulási (reinforcement learning) [12] módszerek, amelyek tudásbázisa fuzzy szabálybázis által leírt. A probléma megoldását leíró fuzzy szabálybázis a tanulási folyamat végeztével áll elő és „ha-akkor” típusú fuzzy szabályokat tartalmaz, ahol a „ha” rész a szabály antecedense (állapot-akció érték), az „akkor” rész pedig a szabály konzekvense (Q-értéke). Ezen szabálybázis tartalmazhat olyan szabályokat, amelyek csak a tanulási folyamat során voltak relevánsak, így ezen szabályok törlésével adott esetben csökkenthető a végső szabálybázis mérete.

A cikk célja, egy olyan klaszterezési módszeren alapuló fuzzy szabálybázis-redukálási algoritmus bemutatása, amely alkalmas a tanulási folyamat végeztével előállt szabálybázis méretének (azaz a szabályok számának) csökkentésére a lényegi szabályok kiemelésével és az elhagyható szabályok törlésével. A javasolt megközelítés koncepcionálisan hasonlít a Multi-Vantage Point Tree (MVP-tree) módszerhez [1], amely több referenciapont (vantage point – VP) alapján végzi az elemek csoportosítását és tér rekurzív felosztását, ezáltal csökkentve a redundanciát és növelve a hasonlóság keresés hatékonyságát. A fejlesztett módszer hatékonysága különböző megerősítéses tanulási mintapéldák (Cart-Pole, Mountain Car) alkalmazásán keresztül kerül bemutatásra.

2. A FRIQ-learning és a HFRIQ-learning megerősítéses tanulási módszerek

A FRIQ-learning (Fuzzy Rule-Interpolation based Q-learning) [19] [21] és ennek a heurisztikusan gyorsított (szakértői tudásbázissal kiegészített) verziója (Heuristically Accelerated Fuzzy Rule-Interpolation based Q-learning – HFRIQ-learning) [15] [17] olyan fuzzy szabályinterpoláció alapú Q-tanulási módszerek, amelyek a „FIVE” (Fuzzy Interpolation in the Vague Environment) [7] fuzzy szabályinterpolációs eljárást alkalmazzák a Q-függvény leírására. Ezen FRI (Fuzzy Rule Interpolation) módszer alkalmazása következtében a rendszer tudásbázisa egy ritka szabálybázis által leírt, amelyben egy r_i ($i \in [1, m]$) szabály formátuma az m méretű R szabálybázisban a következő [19] [21]:

$$r_i: \text{If } s_1 \text{ is } S_1^i \text{ And } s_2 \text{ is } S_2^i \text{ And ... And } s_n \text{ is } S_n^i \text{ And } a \text{ is } A^i \text{ Then } \tilde{Q}(s, a) = q^i \quad (1)$$

ahol $s \in \mathcal{S}$ az n -dimenziós állapot megfigyelés, A^i az i -edik szabály egydimenziós akcióuniverzumának (U) fuzzy halmaza, $a \in U$ az akció, S_j^i az i -edik ($i \in [1, m]$) szabály j -edik ($j \in [1, n]$) állapotdimenziójának fuzzy halmaza az n -dimenziós \mathcal{S} állapottérben, s_j a j -edik dimenziója az s állapot megfigyelésnek, $\tilde{Q}(s, a)$ a FIVE FRI [7] által becsült Q-függvény. Az i -edik szabály antecedense az $s_1 \text{ is } S_1^i \text{ And } s_2 \text{ is } S_2^i \text{ And ... And } s_n \text{ is } S_n^i \text{ And } a \text{ is } A^i$, konzekvensé (Q-értéke) pedig a q^i .

A FIVE FRI modellel közelített $\tilde{Q}(s, a)$ függvény i -edik fuzzy szabályának konzekvensé a $(k + 1)$ -edik iterációban a következő [19] [21]:

$$q_i^{k+1} = \begin{cases} q_i^k + \Delta\tilde{Q}^{k+1}(s, a) & \text{ha } (s, a) = (s^i, a^i) \\ & \text{valamennyi } i\text{-re,} \\ q_i^k + \Delta\tilde{Q}^{k+1}(s, a) * (1/\delta_{v,i}^\lambda) / \left(\sum_{i=1}^m 1/\delta_{v,i}^\lambda \right) & \text{egyébként} \end{cases} \quad (2)$$

Ahol q_i^k -edik az i -edik fuzzy szabály konzekvensé a k -edik iterációban, $\delta_{v,i}^\lambda$ a skálázott távolság az (s, a) állapot-akció megfigyelés és az i -edik szabály (s^i, a^i) állapot-akció antecedense között, λ a Shepard-paraméter, $\Delta\tilde{Q}^{k+1}(s, a)$ a Q-függvény $(k + 1)$ -edik iterációbeli frissítési értéke az (s, a) állapot-akció párban, amely a következő módon határozható meg [19] [21]:

$$\tilde{Q}^{k+1}(s, a) = \tilde{Q}^k(s, a) + \Delta\tilde{Q}^{k+1}(s, a) \quad (3)$$

$$\Delta\tilde{Q}^{k+1}(s, a) = \alpha * \left(g(s, a, s') + \gamma * \max_{a' \in U} \tilde{Q}^k(s', a') - \tilde{Q}^k(s, a) \right) \quad (4)$$

Ahol γ a leszámítási tényező, $\alpha \in [0, 1]$ a tanulási ráta, q_i^{k+1} az i -edik szabály singleton konklúziója a $(k + 1)$ -edik iterációban, a a végrehajtott akció s -ben, s' az új állapotmegfigyelés, $g(s, a, s')$ a jutalom értéke az $s \rightarrow s'$ állapotátmenetre, \tilde{Q}^k a k -edik, \tilde{Q}^{k+1} pedig a $k + 1$ -edik iteráció becsült konklúziója (Q-értéke) a FIVE FRI által.

A FRIQ-learning tanulási fázisa kezdetben 2^{n+1} darabszámú, 0 konzekvensértékkel rendelkező fuzzy szabállyal indul, amit az inkrementális szabálybázis-építési módszer [19] [21] iterációról iterációra bővít, illetve hangol. Ezek a kezdeti vagy úgynevezett sarokponti szabályok az $(n + 1)$ -dimenziós hiperkocka sarkaiban, azaz az univerzumok határain helyezkednek el. A továbbiakban a módszer ezt a kezdeti szabálybázist bővíti új szabályokkal vagy azok konzekvensét (Q-értékét) hangolja a (2) formula által, attól függően, hogy új szabály beillesztésére, vagy csak a meglévő szabálybázis Q-értékének a frissítésére van-e szükség.

Új szabály szabálybázisba történő beszúrása az ágens környezetéből érkező megerősítési információk és a Q-függvény frissítési értékei ($\Delta\bar{Q}$) alapján történik. Ha $\Delta\bar{Q}$ értéke magasabb, mint egy előre meghatározott Q-frissítési limit ($\Delta\bar{Q} > \varepsilon_Q$) és a létező legközelebbi szabály is távol van az éppen beszúrandó szabály pozíciójához képest, akkor új szabály felvétele történik az adott lehetséges szabálypozícióba. Abban az esetben, ha $\Delta\bar{Q}$ értéke kisebb, mint az előre meghatározott Q-frissítési limit ($\Delta\bar{Q} < \varepsilon_Q$), akkor nem történik új szabály beszúrása a szabálybázisba. Ebben az esetben a teljes szabálybázis konzekvensének, azaz Q-értékének a frissítése (hangolása) valósul meg. Az említett lépések minden egyes iterációban végrehajtásra kerülnek addig, amíg a tanulási fázis (azaz az inkrementális szabálybázis-építési fázis) véget nem ér. Akkor áll elő a rendszert működtető, végleges tudásbázis (fuzzy szabályrendszer) és ér véget a tanulási folyamat, ha már nem kerül új szabály beszúrásra a szabálybázisba és a $\Delta\bar{Q}$ frissítési értékek relatívan kicsik maradnak.

A HFRIQ-learning [15] [17] ezen FRIQ-learning-módszer kiterjesztése, amely alkalmas szakértő által megadott tudásbázis injektálására [14] [16] a tanulási folyamatba, majd a megadott szakértői szabályrendszer hangolására [17] és redukálására [17] [18] a tanulási folyamat során. Azáltal, hogy ezen módszer esetében már a tanulási folyamat során megtörténik a szabálybázis-redukálás egy szabálytávolság-alapú módszer által [17] [18], így ebben az esetben nem feltétlenül van szükség további szabálybázis-redukálásra.

2.1. FRIQ-learning szabálybázis-redukálási módszerek

A FRIQ-learning inkrementális szabálybázis-építési fázisában létrejött szabálybázis tartalmazhat redundáns szabályokat vagy olyan szabályokat melyeknek kiadódhatnak más, már létező szabályokból. Ezen szabályok a szabálybázisból való végleges törlésével a szabálybázis mérete csökkenthető [20], csökkentve ez által a szabálybázis komplexitását és a működtető végleges tudásbázis méretét. Az elhagyható szabályok megállapítására eredetileg 3 dekrementális tudásbázis-redukálási módszerrel (I., II., III.) [20] [21] rendelkezik a FRIQ-learning-rendszer, amelyek az inkrementális szabálybázis-építési fázis után alkalmazhatók opcionálisan. Ezen módszerek mindegyike a tanulási folyamat végén előállt teljes szabályrendszer szabályait vizsgálja, hogy az egyes szabályok lényegiek (kardinális), vagy kiadódók (redundáns). A szabálybázis-redukálási módszerek eltávolítják a redundáns szabályokat a szabályrendszerből, így az eredetivel közel azonos az információt hordozó szabályrendszert alkotnak a lényegi szabályokból. A redukációs módszerek közös jellemzője, hogy a szabályok konzekvens értékét, azaz a Q-értéket vizsgálja. Az I–III. redukálási módszerek dekrementálisak, azaz a végső redukált szabálybázis a tanulási fázis végén kapott teljes szabálybázis egyes szabályainak elhagyásával jön létre, fokozatosan csökkentve annak méretét. Az egyes szabálybázis-redukálási módszerekkel kapott csökkentett méretű szabálybázisok közel ugyanazt a Q-függvényt (irányítási felületet) írják le, mint a redukálás előtti esetben, de kevesebb szabállyal (azaz interpolációs tartóponttal).

Az I. jelölésű szabálybázis-redukálási stratégia [20] [21] azon szabályokat törli a teljes szabálybázisból, amelyeknek abszolút értékben alacsony a Q-értékük (konzekvens-értékük). Minden egyes szabály törlése után megvizsgálja, hogy az adott szabályt elhagyva a probléma még megoldható-e és ha igen, akkor folytatja a folyamatot. Ezt addig teszi, amíg az adott szabály törlése után kapott eredmény nem tér el lényegesen az azt megelőzőtől. Ha lényegesen eltér, azaz a feladat már nem oldható meg, akkor a törölt szabályt visszahelyezi a szabálybázisba és fontos szabályként jelöli meg. Ellenkező esetben azonban véglegesen törli azt a szabálybázisból.

A II. jelölésű tudásbázis-redukálási módszer [20] [21] hasonló, mint az I., de azzal a különbséggel, hogy ebben az esetben a legnagyobb Q-értékkel rendelkező szabályok kerülnek vizsgálatra, feltételezve, hogy a nagyobb Q-értékkel rendelkező szabályok jelentősebb befolyással bírnak.

A III. jelölésű szabálybázis-redukációs módszer [20] [21] nem egyesével vizsgálja a szabályokat, hanem szabálycsoportokat alakít ki, majd ezeket távolítja el. A szabálycsoportok kialakítása szintén Q-érték alapján történik, a módszer meghatározza a Q-értékek teljes tartományát (legkisebb és legnagyobb érték közötti értérest), majd ezen tartomány alapján hoz létre két szabálycsoportot, úgy hogy a tartomány fele lesz a tűrészatar. Ezt követően a nagyobb Q-értékkel rendelkező szabálycsoport kerül kiértékelésre. Ha ezzel a probléma még sikeresen megoldható, akkor az ebből a megmaradt szabálycsoportból kiindulva ismétlődik az eljárás. Ha nem oldható meg sikeresen, akkor a törölt szabályok visszakerülnek a szabályrendszerbe, de a vizsgált Q-érték tartomány újra megfelezésre kerül. Ez addig ismétlődik, amíg a tűrészatar értéke olyan nem lesz, hogy az adott szabálycsoport már eltávolítható. Abban az esetben, ha a tűrészatar alapján csak egyetlen szabály marad a csoportban és a probléma így sem oldható meg, akkor ez a szabály fontos (állandó) jelölést kap, majd a továbbiakban ezen állandónak jelölt szabályokat már nem vizsgálja.

3. Klaszterezésen alapuló fuzzy szabálybázis-redukálási módszer

A fejlesztett módszer egy hierarchikus klaszterezési eljárásan alapszik, mely a tanulási folyamat végeztével előállt szabálybázis szabályait klaszterekbe (és alklaszterekbe) rendezi, majd az egyes klaszterekből az adott klasztert jellemző lényegi (kardinális) szabályokként kiemeli a legnagyobb és a legkisebb Q-értékű szabályokat. A módszer úgy csökkenti a teljes szabálybázis méretét, hogy az egyes klaszterek (és alklaszterek) lényegi szabályait tartja csak meg, a többi szabályt pedig elhagyja a szabálybázisból. A szabálybázis szabályai ebben az értelemben, mint objektumok, azaz az adathalmaz adatpontjaiként tekinthetők.

Az algoritmus első lépésként egy D távolságmátrix kerül meghatározásra, mely a szabálybázis minden egyes szabálya közötti távolságot tartalmazza. Az egyes adatpontok (szabályok) közötti távolságok meghatározása a „FIVE” FRI-módszer által

alkalmazott többdimenziós Euklideszi-távolság-alapú távolságszámítási eljárás [6] (jelöljük *FIVE_dist-el*) alapszik. A számított távolságmátrix egy négyzetes ($m * m$) mátrix, melynek mérete a szabálybázis szabályainak a számától (m) függ. A mátrix főátlójában csupa nullák helyezkednek el, annak következtében, hogy egy szabálypont saját magától vett távolsága mindig 0. A szabálybázis két i -edik (r_i) és j -edik (r_j) indexű szabálya (azaz adatpontja) közötti távolságot d_{ij} jelöli, ahol $i, j = [1, \dots, m]$:

$$D_{[ij]} = d_{ij} = FIVE_dist(r_i, r_j) \quad (5)$$

A következő fázisban a távolságmátrix alapján p_1, p_2 pivot objektumok (a klaszterezési módszer speciális pontjai, jelen esetben fuzzy szabályok) kerülnek meghatározásra, amelyek a két egymástól legtávolabbi szabályok lesznek. Ezen p_1 és p_2 pivot pontok megfelelnek az MVP-tree módszer vantage point-jainak. Tulajdonképpen a D távolságmátrix d_{ij} távolságértékei közül a legkisebb és a legnagyobb távolságértékkel rendelkező i, j indexű szabálypárok, amelyek között a távolság maximális:

$$(p_1, p_2) = arg \max_{d_{ij} \in D} (d_{ij}) \quad (6)$$

A következő iterációban a szabálybázis minden egyes szabályának a távolsága kerül kiszámításra a p_1 és p_2 adatpontoktól. Ezen távolságszámítás szintén a távolságmátrix meghatározásánál is alkalmazott „FIVE” FRI távolságszámítási módszerének [6] az alkalmazásával történik:

$$d(r_i, p_1) = FIVE_dist(r_i, p_1) \quad (7)$$

$$d(r_i, p_2) = FIVE_dist(r_i, p_2) \quad (8)$$

A szabálybázis szabályai a pivot objektumoktól való távolságuk alapján kerülnek besorolásra klaszterbe. A klaszterhez való hozzárendelés alapja egy ε távolságküszöb, amely a következőképpen számítható:

$$\varepsilon = \frac{\max(d(r_i, p_1))/2 + \max(d(r_i, p_2))/2}{2} = \frac{\max(d(r_i, p_1)) + \max(d(r_i, p_2))}{4} \quad (9)$$

Az ε távolságküszöb értéke tehát, amely alapján az egyes klaszterek (majd alklaszterek) kerülnek kialakításra, a p_1 és p_2 adatpontoktól vett legtávolabbi távolságértékek felének az összegei, majd az ezek alapján számított átlag, azaz a két klaszter „sugár” átlagának a fele. Ezen ε küszöbérték alapján a klaszterekhez való hozzárendelés a következő:

$$Cluster_{\varepsilon}(r_i) = \begin{cases} C_{left}, & \text{ha } d(r_i, p_1) \leq \varepsilon \\ C_{right}, & \text{egyébként} \end{cases} \quad (10)$$

Kezdetben a teljes szabálybázis minden egyes szabálya egyetlen klasztert alkot, majd ez a kezdeti klaszter kerül felosztásra további két alklaszterre minden egyes lépésben (felosztó, top-down klaszterezés). Az alklaszterekre (right branch, left branch) való bontás alapja a (9) formula alapján számított ε távolságküszöb. Minden egyes szabály távolsága meghatározásra kerül a p_1 és p_2 adatpontoktól, majd ha az éppen vizsgált szabály távolsága az adott p_1 vagy p_2 pivot elemnél kisebb, mint a számított ε távolságküszöb értéke, akkor az adott szabály a bal alklaszterhez (left branch), ellenkező esetben pedig a jobb alklaszterhez (right branch) kerül hozzáadásra. Az egyes iterációkban kialakult alklaszterek szabályai közül a legnagyobb és a legkisebb Q-értékkel (konzekvenssel) rendelkező szabályok kerülnek fontos szabályként megjelölésre annak következtében, hogy a rendszer működésének hatékonyságát a legrosszabb Q-érték (illetve megerősítés) is befolyásolja [2]. Tehát, az adott iterációban kialakult alklaszterek szabályai közül mindig a legkisebb és a legnagyobb Q-értékkel rendelkező szabályok kerülnek lényegi szabályként megjelölésre (és a teljes szabálybázisból eltávolításra), majd egy átmeneti, kezdetben üres szabálybázishoz történő hozzáadásra, amely mérete ezért inkrementálisan növekszik. Így a fontos szabályként megjelölt szabályok kikerülnek az egyes klaszterekből és bekerülnek a fontos szabályok halmazába. A folyamat a kialakult alklaszterekre rekurzívan ismétlődik, az alklaszterek újabb alklaszterekre lesznek felosztva, majd a lényegi szabályok kiemelve. Ez a folyamat addig ismétlődik, amíg a lényegi szabályokat tartalmazó (átmeneti) redukált szabálybázis meg nem oldja az adott problémát, azaz az így kapott megerősítés értéke nagyobb, mint egy előre definiált megerősítési küszöbérték.

A fejlesztett fuzzy szabálybázis redukálási algoritmus pszeudokódja az alábbi:

Algoritmus: clusteringBasedRBreduce(R)

Bemenet: szabálybázis

Kimenet: redukált szabálybázis

D távolságmátrix számítása

p_1 és p_2 pivot objektumok meghatározása a számított távolságok alapján

szabályok távolságának számítása p_1 és p_2 objektumoktól

ε távolságküszöb számítása

ε távolságküszöb alapján szabálybázis felosztása bal és jobb alklaszterekre legnagyobb és legkisebb Q-értékű (kardinális) szabályok kiválasztása az egyes alklaszterekből

a 6. lépésben meghatározott kardinális szabályok hozzáadása az átmeneti redukált szabálybázishoz

megerősítés számítása a 7. lépésben létrejött átmeneti redukált szabálybázisra (szabálybázis kiértékelése)

If (számított megerősítés > előre definiált megerősítés)

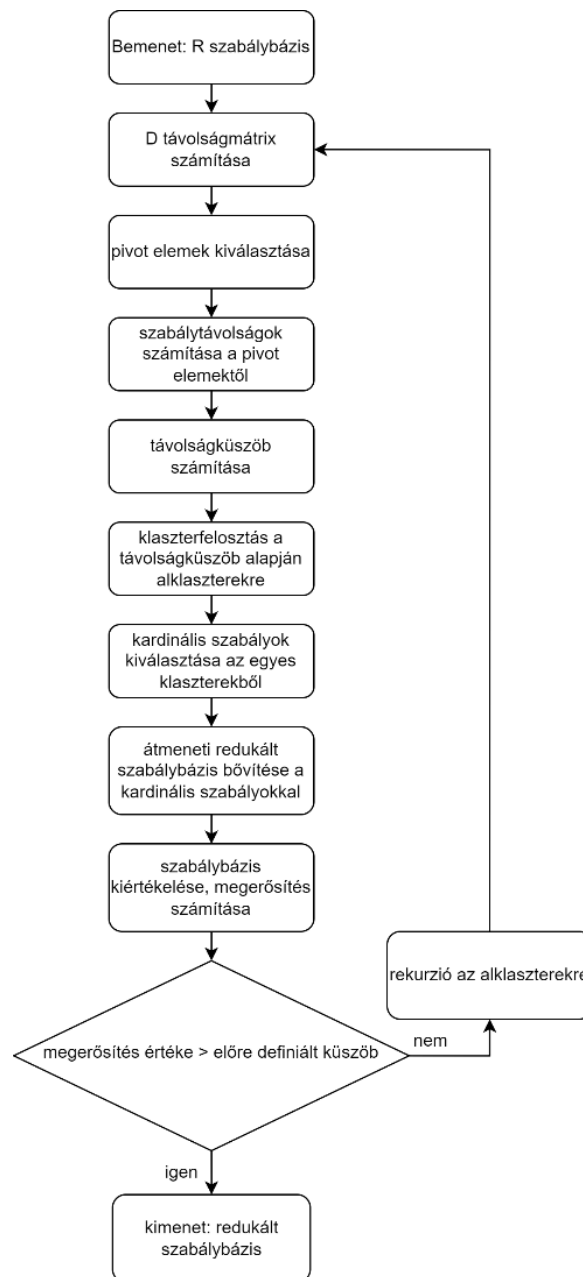
vége, kardinális szabályok megtalálva, return redukált szabálybázis

else

következő rekurzív lépés a bal és jobb alklaszterekre
end

1. algoritmus: A fejlesztett, klaszterezési eljárás alapuló szabálybázis-redukálási módszer algoritmus, amely a tanulási folyamat után alkalmazható

A következő ábra az 1. algoritmus folyamatábráját szemlélteti:



3.1. Alkalmazáspéldák

A fejlesztett algoritmus hatékonysága a „Cart-Pole” és a „Mountain Car” mintapéldák alkalmazásával kerül bemutatásra, összevetve az előzőleg bemutatott I., II. és III. jelölésű szabálybázis-redukálási módszerek által kapott eredményekkel. A klaszterezési eljárás alapján alapuló szabálybázis-redukálási módszer (IV. jelölésű) alkalmazásával az egyes futási esetekben kapott szabálybázisméret (szabályszaámok) nem feltétlenül lettek kisebbek, mint az I–III. jelöléssel ellátott redukálási módszerek alkalmazása esetében, de a futási idejében nagy eltérés mutatkozik.

A Cart-Pole mintapélda esetében kapott futási eredményeket a következő 1. táblázat foglalja össze:

1. táblázat

A „Cart-Pole” mintapélda esetében kapott futási eredmények

Redukciós módszer	Szabályszaám	Futási idő (sec)	Sebesség
I.	7	3546,70	1×
II.	6	3530,50	1×
III.	8	261,32	13,6×
IV.	24	17,60	201,5×

A Cart-Pole mintapélda esetében a fejlesztett klaszterezési eljárás alapján alapuló szabálybázis-redukálási módszer (IV. jelölésű) által kapott szabálybázis 24 darab szabályt tartalmaz, amely több, mint az I–III. redukálási módszerek esetében kapott szabálybázis méretek, de a futási idejében jelentős javulás áll elő, amely így 201,5-ed része, mint az I. módszer esetében.

A Mountain Car mintapélda esetében kapott futási eredményeket a következő 2. táblázat foglalja össze:

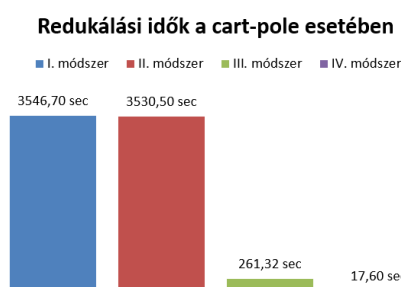
2. táblázat

A „Mountain Car” mintapélda esetében kapott futási eredmények

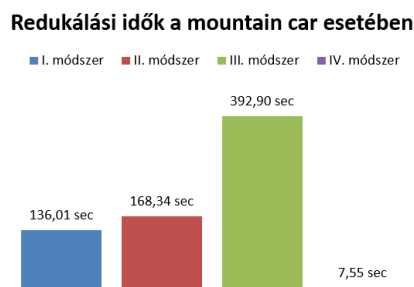
Redukciós módszer	Szabályszaám	Futási idő (sec)	Sebesség
I.	27	136,01	1×
II.	53	168,34	0,8×
III.	26	392,90	0,3×
IV.	32	7,55	18×

A Mountain Car mintapélda esetében a javasolt szabálybázis-redukálási módszer (IV. jelölésű) alkalmazásával a redukált szabálybázis 32 darab szabályt tartalmaz, amely a II. redukálási módszer esetében létrejött 53 darab szabályhoz viszonyítva kevesebb. Jelentős eltérés szintén a futási időben mutatkozik meg, amely így 18-ad része, mint az I. redukálási módszer esetében.

Az 1. és 2. ábrák a tanulási fázis után alkalmazott egyes szabálybázis-redukálási módszerek által kapott futási időket szemléltetik (másodpercben):



1. ábra. Szabálybázis redukálási idők a Cart-Pole mintapélda esetében



2. ábra: Szabálybázis redukálási idők a Mountain Car mintapélda esetében

A fejlesztett szabálybázis-redukálási módszer teljes futási idejének nagy részét nem a hierarchikus klaszterezési algoritmus (illetve a „FIVE” FRI-távolságszámítás) lépései teszik ki, hanem az adott iterációban előálló, redukált szabálybázis tesztelése, azaz, hogy az helyesen megoldja-e már az adott problémát vagy szükséges egy újabb rekurzív iteráció futtatása. A teljes szabálybázis-redukálási időt és annak a szabálybázisok tesztelésére fordított idő mértékét a 3. táblázat foglalja össze:

3. táblázat

Szabálybázis tesztelési ideje a teljes redukálási időből

Mintapélda	Teljes redukálási idő (sec)	Szabálybázis tesztelési idő (sec)	%
Cart-Pole	17,6	11,13	63
Mountain Car	7,55	6,49	86

4. Összefoglalás

A cikkben egy klaszterezési eljárás alapuló fuzzy szabálybázis-csökkentési módszer került bemutatásra, amely a FRIQ-learning tanulási folyamatot követően lehetőséget biztosít a szabálybázis méretének optimalizálására. Az alkalmazott hierarchikus klaszterezési algoritmus a szabályok közötti távolságokat figyelembe véve alakítja ki a klasztereket, majd minden klaszteren belül a legkisebb és legnagyobb Q-értékű szabályokat kiválasztva építi fel a redukált (csökkentett méretű) szabálybázist. A kiválasztási folyamat után a módszer ellenőrzi, hogy a redukált szabályrendszerrel a probléma megfelelően megoldható-e és ezt a folyamatot mindaddig ismétli, amíg a rendszer az így létrejött szabálybázissal meg nem oldja a problémát. A Cart-Pole és a Mountain Car mintapélda alkalmazása során kapott eredmények

alapján elmondható, hogy a bemutatott szabálybázis-redukációs eljárás alkalmas a fuzzy szabálybázis méretének csökkentésére illetve végrehajtási ideje jelentősen rövidebb, mint a bemutatott I., II., és III. FRIQ-learning szabálybázis redukálási módszerek esetében.

Irodalomjegyzék

- [1] Bozkaya, Tolga, and Meral Ozsoyoglu (1997). Distance-based indexing for high-dimensional metric spaces. *Proceedings of the 1997 ACM SIGMOD International Conference on Management of Data*.
- [2] Fuchida, Takayasu, Kathy Thi Aung, and Atsushi Sakuragi (2010). A study of Q-learning considering negative rewards. *Artificial Life and Robotics*, 15, 3, 351–354. <https://doi.org/10.1007/s10015-010-0822-7>
- [3] Ghahramani, Zoubin (2003). Unsupervised learning. Summer school on machine learning. Berlin, Heidelberg: Springer Berlin Heidelberg, 72–112.
- [4] Ghosal, Attri et al. (2020). A short review on different clustering techniques and their applications. *Emerging Technology in Modelling and Graphics: Proceedings of IEM Graph 2018*, 69–83. https://doi.org/10.1007/978-981-13-7403-6_9
- [5] Gordon, Allan David (1999). *Classification*. CRC Press.
- [6] Kovács, Sz., Kóczy, L. T. (1997). The use of the concept of vague environment in approximate fuzzy reasoning. *Fuzzy Set Theory and Applications*. Tatra Mountains Mathematical Publications, Mathematical Institute Slovak Academy of Sciences, Bratislava, Slovak Republic, vol. 12, 169–181.
- [7] Kovács, Szilveszter (2006). Extending the fuzzy rule interpolation “five” by fuzzy observation. *Computational Intelligence, Theory and Applications: International Conference 9th Fuzzy Days in Dortmund, Germany, Sept. 18–20, 2006 Proceedings*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. https://doi.org/10.1007/3-540-34783-6_48
- [8] Oyewole, Gbeminiyi John, and George Alex Thopil (2023). Data clustering: application and trends. *Artificial intelligence review*, 56, 7 6439–6475. <https://doi.org/10.1007/s10462-022-10325-y>
- [9] Pandit, Shraddha, and Suchita Gupta (2011). A comparative study on distance measuring approaches for clustering. *International journal of research in computer science*, 2, 1, 29.
- [10] Pujari, Arun K (2001). *Data mining techniques*. Universities press.
- [11] Rokach, Lior, and Oded Maimon (2005). *Clustering methods*. Data mining and knowledge discovery handbook, 321–352.
- [12] Sutton, Richard S., and Andrew G. Barto. Reinforcement learning: An introduction. MIT press, 2018. <https://doi.org/10.1017/s0263574799211174>
- [13] Tikk, D., Johanyák, Z. C., Kovács, S., & Wong, K. W. (2011). Fuzzy rule interpolation and extrapolation techniques: Criteria and evaluation guidelines. *Journal of advanced computational intelligence and intelligent informatics*, 15 (3), 254–

263. <https://doi.org/10.20965/jaciii.2011.p0254>
- [14] Tomba, Tamás, and Szilveszter Kovács (2020). Applying Expert Heuristic as an a Priori Knowledge for FRIQ-Learning. *Acta Polytechnica Hungarica*, 17, 4. <https://doi.org/10.12700/aph.17.4.2020.4.2>
- [15] Tomba, Tamás, and Szilveszter Kovács (2022). Heuristically accelerated FRIQ-learning. *20th Jubilee International Symposium on Intelligent Systems and Informatics (SISY 2022)*, IEEE. <https://doi.org/10.1109/iccc57093.2023.10178919>
- [16] Tomba, Tamás, and Szilveszter Kovács (2024). Integrating Expert Knowledge into Fuzzy Reinforcement Learning. *2024 IEEE 18th International Symposium on Applied Computational Intelligence and Informatics (SACI)*, IEEE. <https://doi.org/10.1109/saci60582.2024.10619888>
- [17] Tomba, Tamás, and Szilveszter Kovács (2024). Knowledge Base Optimization of the HFRIQ-Learning. *Acta Polytechnica Hungarica*, 21, 10. <https://doi.org/10.12700/aph.21.10.2024.10.6>
- [18] Tomba, Tamás, and Szilveszter Kovács (2023). Tudásbázis redukálás a heurisztikusan gyorsított FRIQ-learning rendszerben. *Production Systems and Information Engineering*, 11, 2, 1–12. <https://doi.org/10.32968/psaie.2022.4.4>
- [19] Vincze, Dávid, and Szilveszter Kovács (2009). Fuzzy rule interpolation-based Q-learning. *2009 5th International Symposium on Applied Computational Intelligence and Informatics*, IEEE. <https://doi.org/10.1109/saci.2009.5136311>
- [20] Vincze, Dávid, and Szilveszter Kovács (2015). Rule-base reduction in Fuzzy Rule Interpolation-based Q-learning. *Recent Innovations in Mechatronics*, 2, 1–2, 1-6. <https://doi.org/10.17667/riim.2015.1-2/10>
- [21] Vincze, Dávid (2013). *Fuzzy Rule Interpolation-based Q-learning*. PhD dissertation.