

ANTECEDENS REDUNDANCIA FELTÁRÁSA A FUZZY SZABÁLY INTERPOLÁCIÓ ALAPÚ MEGERŐSÍTÉSES TANULÁSI MÓDSZERBEN

Tóth Alex

MSc hallgató, Miskolci Egyetem
Informatikai Intézet, Általános Informatikai Intézeti Tanszék
3515 Miskolc, Miskolc-Egyetemváros, e-mail: toth.alex@iit.uni-miskolc.hu

Vincze Dávid

egyetemi docens, Miskolci Egyetem
Informatikai Intézet, Általános Informatikai Intézeti Tanszék
3515 Miskolc, Miskolc-Egyetemváros, e-mail: vincze.david@iit.uni-miskolc.hu

Absztrakt

Jelen cikk olyan új módszereket mutat be, amelyekkel hatékonyabbá tehetők a FRIQ-learning (Fuzzy Rule Interpolation-based Q-learning – fuzzy szabály interpoláció alapú Q-tanulás) gépi tanulási módszert felhasználó automatikus tudáskinyerési eljárások, valamint új megoldások is bevezethetők a felhasználásukkal. A korábban kifejlesztett FRIQ-learning módszer egy olyan fuzzy szabálybázist képes felépíteni egy adott probléma megoldásához, melyben csak a kardinális szabályok szerepelnek. A fuzzy szabály interpolációnak (FRI) köszönhetően pedig a közvetlenül le nem fedett esetekre is számítható konklúzió. A konkrét szabályok kiválogatása, hogy valójában melyek a ténylegesen fontosak egy adott probléma megoldásához, melyek azok amelyek kiadódnak az interpoláció jóvoltából, nem triviális feladat. Teljes szabályok elhagyására már bemutatásra került néhány lehetséges stratégia, azonban olyan stratégia eddig még nem került kidolgozásra, amely szabályok antecedenseinek elhagyhatóságát vizsgálja. Ez a cikk egy ilyen megoldást vázol fel, lehetővé téve egy adott szabálybázis további egyszerűsítését, elősegítve olyan ritka fuzzy szabálybázisok létrehozását, amelyekből közvetlenül ember által is kiolvasható a feltárt tudás.

Kulcsszavak: mesterséges intelligencia, megerősítéses tanulás, fuzzy szabálybázis redukció, antecedens redundancia

Abstract

This paper introduces novel methods to improve the efficiency of the automated knowledge extraction methods used in the FRIQ-learning (Fuzzy Rule Interpolation-based Q-learning) machine learning method. For solving a given problem, the FRIQ-learning reinforcement learning method is capable of constructing a sparse fuzzy rule-base, which does not need to contain all the possible rules. Hence it is sufficient to keep only the most important rules due to the application of fuzzy rule interpolation (FRI). Finding the specific rules that are important to solve a problem is not a trivial task. Some possible strategies for removing unimportant rules from the rule-base have already been introduced, but no strategies addressing the antecedents of the rules have been developed yet. This paper introduces a solution allowing further reduction of rule-bases, thus facilitating the creation of a sparse fuzzy rule-base from which the knowledge can be directly extracted.

Keywords: artificial intelligence, reinforcement learning, fuzzy rule-base reduction, antecedent redundancy

1. Bevezetés

A FRIQ-learning (Fuzzy Rule Interpolation-based Q-learning – fuzzy szabály interpoláció alapú Q-tanulás, lásd: [25], [27], [28], [29] és [30]) egy olyan megerősítéses tanulási algoritmus, ami kibővíti a klasszikus megerősítéses tanulási módszerek halmazát, olyan módon, hogy a feltérképezett tudást fuzzy szabályok formájában tárolja. Ilyen módon a feltérképezett tudás ember által is közvetlenül értelmezhető fogalmakkal van leírva (pl. kicsi-nagy, sok-kevés, közel-távol). Továbbá a FRIQ-learning képes azon szabályok elhagyására, amelyek a rendszer működtetésében nem, vagy csak minimálisan vesznek részt. Ennek a módszernek az alkalmazásával olyan fuzzy szabálybázisok konstruálhatók, amelyek csak a fontos és tényleges működtető szabályokat fogják tartalmazni. Így a fuzzy szabálybázisok önmagyarazó mivolta miatt, a tudás közvetlenül kiolvasható.

Bár az ilyen módon létrehozott szabálybázisok jelentősen kisebb méretűek lehetnek, mint a teljes, fedő szabálybázisok, az adott problémától függően még így is viszonylag sok szabályt tartalmazhatnak, amely megnehezíti egy szabálybázis olvashatóságát és értelmezését. Ezért érdemes kutatni a lehetőségeket, hogy hogyan lehet ezeknek a szabálybázisoknak a méretét még tovább csökkenteni, úgy, hogy a rendszer továbbra is megfelelően működjön.

A cikkben bemutatásra kerülő új módszerek képesek a szabályok redundáns részeinek azonosítására, így a felesleges szabályok, szabályrészek eltávolíthatók, illetve összevonhatók más szabályokkal. Ennek eredményeként ember által is könnyebben értelmezhető szabálybázisok hozhatóak létre. A következőkben részletesen bemutatásra kerülnek a kifejlesztett módszerek, továbbá néhány mintaalkalmazáson [25] keresztül demonstrálja a cikk az új módszerek hatékonyságát, valamint összeveti azokat a FRIQ-learning gépi tanulási módszerben használt korábbi megoldásokkal.

2. Fuzzy szabály interpoláció alapú megerősítéses tanulás

A megerősítéses tanulás célja egy adott probléma megoldásának feltárása olyan módon, hogy a problémátér egyes állapotainak akciókhoz való megfelelő kapcsolásával egy, a megoldás minőségét jelző, numerikus jutalomértéket maximalizáljon. A módszer legnagyobb előnye, hogy nem kell megadni, hogy hogyan kell megoldani a problémát, csak azt, hogy mi a végső cél, azaz azon van a hangsúly, hogy mit kell megoldani, nem pedig azon, hogy hogyan. A tanuló programmal nincs közölve, hogy milyen szituációban milyen akciót kell választania, ehelyett próbálkozással kell rájönnie, adott helyzetben melyik akció végrehajtása a legcélravezetőbb. Ennek érdekessége, hogy egy akció kiválasztásának nem csak az aktuális állapotban van hatása a kapott jutalomértékre, hanem az azt követő, illetve minden további állapotban is. Ezek, a próba-hiba és késleltetett jutalomérték technikák, a megerősítéses tanulási módszerek legjellegzetesebb tulajdonságai [19].

A megerősítéses tanulás a dinamikus rendszerelméletből merített ötletekre, elsősorban a Markov-féle döntési folyamatokra épül [4] [35]. A módszer alapgondolata, hogy egy tanuló ágens időről-időre valamilyen módon interakcióba lép a környezetével valamilyen cél elérése érdekében. Az ágensnek érzékelnie kell környezet állapotát, valamint végre kell tudnia hajtani olyan akciókat, melyeknek hatására megváltozik a környezet állapota. Ezek mellett az ágensnek kell, hogy legyen egy, vagy több, a környezet állapotához kapcsolódó célja. A Markov-féle döntési folyamatokban, legegyszerűbb esetben, csupán erre a három tényezőre – a környezet állapota, a végrehajtható akciók, az elérendő cél(ok) – van szükség. Minden módszer, ami ilyen módon leírható problémák megoldására használható, megerősítéses tanulási módszernek tekinthető [19].

A megerősítéses tanulási módszerek alapelemei az állapotok, melyek a környezet és az ágens állapotváltozóinak lehetséges értékeit írják le, az akciók, melyeket az ágens végrehajthat, jutalomértékek,

melyek az egyes akciók végrehajtásának eredményét, hasznosságát jelzik, valamint azon összefüggések, amelyek az ágens viselkedését írják le állapotok és akciók közötti összerendelések formájában, azt határozzák meg, hogy adott állapotban melyik lehetséges akciót kell végrehajtani a legnagyobb jutalom megszerzése érdekében. A cél pedig a probléma megoldásához szükséges állapot- akció-érték függvény meghatározása.

Megerősítés tanulási módszerekkel megoldható problémák állapot-akció-érték függvényeinek meghatározására egy elterjedt algoritmus a TD-learning (Temporal Difference, időbeli különbség) [19], amely képes a végcél elérése nélkül megbecsülni a jutalomértékeket. Amikor rendelkezésre áll a végső jutalom, az értékeket visszaterjesztéses technikával frissíti. A TD-learning leggyakrabban használt változatai a Q-learning [34] és a SARSA [18] (módosított Q-learning).

2.1. FRIQ-learning

A Q-learning [34], az egyik legnépszerűbb megerősítéses tanulási módszer [17], célja megtalálni a Bellman-egyenlet fixpontos megoldását [4]. Az eredeti diszkrét Q-learning esetén az ehhez szükséges állapot-akció-érték függvény közelítését írja le [33]. Folytonos térben való alkalmazhatóságra számos különböző megoldás létezik, melyek fuzzy következtető rendszert használnak (lásd [1] [5] [6] [7]), így az állapot-akció-érték függvény közelítése fuzzy szabályok formájában írható le (Fuzzy Q-learning). Ennek megfelelően az eredeti diszkrét Q tábla helyett fuzzy szabálybázis tárolja a feltérképezett tudást. A FRIQ-learning [28] [29] (Fuzzy Rule Interpolation based Q-learning – fuzzy szabály interpoláció alapú Q-tanulás) a hagyományos Fuzzy Q-learning módszerek kibővítése olyan módon, hogy teljes fuzzy szabálybázisok helyett képes legyen ritka fuzzy szabálybázisokkal dolgozni. Ehhez a hagyományos Fuzzy Q-learning-ben használt nulladrendű Takagi-Sugeno fuzzy következtetési modell helyett egy fuzzy szabály interpolációs módszer (részletekért lásd [20]), a FIVE FRI (Fuzzy rule Interpolation in Vague Environment) [9] [10] [11] módszer került bevezetésre [25]. A FIVE fuzzy szabály interpolációs módszer bevezetése lehetővé teszi a szabálybázisból egyes, kisebb jelentőséggel bíró szabályok elhagyását és ezáltal az eredeti probléma megoldását egy csökkentett méretű szabálybázissal.

A FRIQ-learning gépi tanulási algoritmusban a szabálybázisban tárolt tudás olvashatóságának növelése érdekében [26] (azaz a szabálybázis méretének csökkentése érdekében) a tanulási folyamat közbeni szabálybázis létrehozása egy inkrementális szabálybázis építési módszerrel történik [28] [29]. A módszer lényege, hogy a kezdeti szabálybázishoz addig ad újabb és újabb szabályokat, amíg azokkal megoldhatóvá nem válik a probléma. Az állapottér egy n -dimenziós hiperkocka, ahol n az állapotleírók számát, azaz az állapottér dimenziószámát jelöli. Kezdetben, teljes szabálybázis létrehozása helyett csak minimális számú, 2^{n+1} darab (n darab állapotdimenzió és egy akció dimenzió van) fuzzy szabályt hoz létre a hiperkocka sarkaiban, amelyek az állapottér határait határozzák meg. Ez a kezdeti szabálybázis valójában a tárolt tudás szempontjából üresnek tekinthető. Az inkrementális szabálybázis építési módszer a tanulási folyamat minden iterációja során ellenőrzi az aktuális állapotban választott akcióért kapott jutalomérték alapján, hogy a meglévő szabályok elegendőek-e, vagy egy újabb szabály beiktatása szükséges. Ha a jutalomérték magasabb egy előre meghatározott határértéknél és a legközelebbi már meglévő szabály is egy megadott határértéknél távolabb van az aktuális állapottól, akkor az új szabályt beszúrja a legközelebbi lehetséges szabálypozícióba. A lehetséges szabály pozícióját egy új állapotnak a meglévők közé való illesztésével határozza meg [23] [28].

Ha a jutalomérték alacsony, vagy egy meglévő szabályhoz közel, annak a környezetében helyezkedik el, akkor új szabály beszúrása helyett csupán a környező szabályok Q értékét módosítja a FRIQ-learning. Az inkrementális szabálybázis építési folyamat akkor fejeződik be, ha a felépített szabálybázissal meg-

oldható a probléma anélkül, hogy új szabályt kellene beszúrni, vagy meglévő szabályokat kellene módosítani, illetve ha előre megadott számú iteráció alatt nem talált megoldást az algoritmus. Megoldás esetén, a folyamat eredménye egy ritka fuzzy szabálybázis. Az eredeti inkrementális szabálybázis építéssel kapott szabálybázissal elért eredményeket az 1. táblázat szemlélteti néhány alkalmazáspéldára.

1. táblázat. Inkrementális szabálybázis építéssel kapott szabálybázissal elért eredmények

Mintapélda	Szabályok száma	Lépések száma	Jutalomérték
<i>Acrobot</i>	367	57	440
<i>Cartpole</i>	182	1000	5010.5752
<i>Mountaincar</i>	110	472	-3710

Ez már önmagában véve is jelentős csökkenést jelenthet a szabályok számára nézve egy teljes szabálybázishoz viszonyítva, azonban a szabálybázis nagy valószínűséggel még tovább redukálható. Az inkrementális szabálybázis építési folyamat eredményeképp kapott szabálybázis tartalmazhat olyan szabályokat, amelyek redundánsak, vagy nincs jelentős hatásuk a végső szabálybázisban. Ennek oka, hogy az inkrementális szabálybázis építés során a szabálybázis dinamikusan változik, az egyes szabályok jelentősége minden egyes iterációban megváltozhat. Egy szabály lehet, hogy nagyon fontos a szabálybázis építés kezdeti fázisában, de akár teljesen jelentéktelenné válhat a végső szabálybázisban. A szabálybázis építés során az új, illetve megváltozott szabályok befolyásolhatják más szabályok jelentőségét a teljes rendszerre vonatkoztatva a fuzzy szabály interpolációból adódóan [30]. Számos stratégia került kidolgozásra [22] [30], melyek képesek ezeket a szabályokat azonosítani és eltávolítani a szabálybázisból, úgy, hogy az így eredményül kapott szabálybázis felhasználásával a probléma továbbra is megoldható marad. Azonban ezek a módszerek csak teljes szabályok eltávolítására képesek a szabályok antecedenseinek együttes vizsgálatával. A következőkben bemutatásra kerülnek olyan új módszerek, amelyek képesek a szabályokon belüli redundancia feltárására.

3. Antecedens redundancia feltárása

Az itt bemutatott módszer célja a redundáns antecedensek kiszűrése egyes szabálycsoportokból. A módszer az egyes szabályok közelségén, hasonlóságán alapszik. Ha egy szabálynak, mint pontnak a térben, vagyis az alkalmazott FIVE FRI módszer esetén a „bizonytalan környezetben” [11], egy másiktól mért távolsága megadott határon belül van, azt mondhatjuk, hogy a két szabály hasonló esetet fed le. Ezen az elven csoportokat képezhetünk a hasonló szabályokból. Ha egy szabálycsoport valamely szabályában nincs szükség egy antecedensre, feltehetőleg a csoport többi szabályában sem lesz rá szükség a hasonlóságuk miatt, mivel a fuzzy szabály interpoláció használata miatt ezek kiadódóak lehetnek. Első lépésként vesszük az első szabályt a szabálybázisból, majd megkeressük az ettől megadott maximális távolságra lévő szabályokat. A maximális távolság megválasztása problémafüggő (konkrét jutalomfüggvény, környezet leírása, lehetséges akciók stb.), így az alkalmazásonként jelentősen eltérő lehet.

Első lépésként az eredeti szabálybázis alapján létrehozunk egy ideiglenes szabálybázist, melyben a kapott szabálycsoport szabályaiból eltávolítjuk az első antecedensét. Ha az ideiglenes szabálybázissal továbbra is helyesen megoldható a probléma, akkor megtartjuk a módosításokat, egyébként visszavonjuk azokat. Ezt követően eltávolítjuk a szabálycsoport szabályainak következő antecedensét, majd így tovább az összes antecedensét és mindegyikre megvizsgáljuk, hogy az adott antecedens eltávolításával megoldható marad-e a probléma. Vesszük a következő szabályt és az előző lépéseket addig ismétljük, amíg a szabálybázis minden szabályát, valamint azok minden antecedensét meg nem vizsgáltuk. Az

eredményül kapott szabálybázisban a szabályok száma ilyen módon nem csökken, azonban egyes szabályoknak vagy szabálycsoportoknak esetlegesen törölve lett néhány antecedense, illetve előfordulhat, hogy valamely antecedens minden szabályból el lett távolítva, azaz valójában nincs is szükség az adott bemeneti tényezőre és a hozzá tartozó tagsági függvényre, állapotleírókra. A főbb lépéseket az 1. algoritmus mutatja be.

1. algoritmus. *Redundáns antecedensek feltárása.*

1. **Bemenet:**
2. R : Kiinduló szabálybázis
3. q_{min} : Elvárt minimális jutalomérték
4. ε : Jutalomérték változására vonatkozó tűréshatár
5. δ : Távolság, melyen belül hasonlónak tekintjük a szabályokat
6. n : Antecedensek száma
7. **Kimenet:**
8. $R_{új}$: Redukált szabálybázis
- 9.
10. $q, s \leftarrow FRIQ_episode(R)$
11. $i \leftarrow 1$
12. Amíg $i \leq |R|$, addig:
13. $j \leftarrow 1$
14. Amíg $j \leq n$, addig:
15. Ha $R_{ant_j} = \emptyset$, akkor:
16. $j \leftarrow j + 1$
17. Folytatás a következő iterációval
18. $R_{új} \leftarrow R$
19. $\forall r \in R_{új}, d(r, R_i) \leq \delta: r_{ant_j} = \emptyset$
20. $q_{új}, s_{új} \leftarrow FRIQ_episode(R_{új})$
21. $q_{diff} \leftarrow |q - q_{új}|$
22. Ha $q_{új} > q_{min} \wedge s_{új} \leq s \wedge q_{diff} \leq \varepsilon$, akkor:
23. $R \leftarrow R_{új}$
24. $q \leftarrow q_{új}$
25. $j \leftarrow j + 1$
26. $i \leftarrow i + 1$
27. $R_{új} = R$

A fenti módszer alkalmazását követően az eredményül kapott szabálybázisban esetlegesen a szabályok több antecedense is törölve lett, így kialakulhatott több olyan szabály is, melyek csupán a Q értékben (konzekvens) térnek el egymástól. Ezek a szabályok ugyanazokat az eseteket fedik le, így elég lehet csupán egyet megtartani közülük, így célszerű egy másodlagos redukciós stratégia alkalmazása is. Az egyforma eseteket lefedő szabályokat összegyűjtve valamilyen módon választanunk kell közülük,

hogy melyik legyen az az egy, amit megtartunk. Ezekre mutatnak be lehetséges megoldásokat a következő alfejezetek.

3.1. Egyforma szabályok közül az éppen vizsgált megtartása

Itt a kiinduló szabálybázis a redundáns antecedensek eltávolításának eredményeképp kapott szabálybázis. Ebből a szabálybázisból első lépésként vesszük az első szabályt, majd megkeressük azokat a szabályokat, amelyek ettől csak a konzekvensükben (Q értékükben) különböznek. Ezeket a szabályokat egyformának tekinthetjük, mivel pontosan ugyanazokat az eseteket fedik le, ugyanaz az értékük minden állapotdimenzióban és az akciódimenzióban is. Az eredeti szabálybázis alapján létrehozunk egy ideiglenes szabálybázist, melyből eltávolítjuk az éppen vizsgálttól csak Q értékükben különböző szabályokat, vagy, ha nincsenek ilyenek, magát a vizsgált a szabályt. Ha az ideiglenes szabálybázissal helyesen megoldható a probléma, megtartjuk a változtatásokat. Vesszük a következő szabályt és az előző lépéseket addig ismételjük, amíg a szabálybázis összes szabályát meg nem vizsgáltuk. A főbb lépéseket a 2. algoritmus írja le.

2. algoritmus. Az éppen vizsgált szabály megtartása redundánsak adódott szabályok esetén

1. **Bemenet:**
2. R : Kiinduló szabálybázis
3. q_{min} : Elvárt minimális jutalomérték
4. ε : Jutalomérték változására vonatkozó tűréshatár
5. n : Antecedensek száma
6. **Kimenet:**
7. R_{ij} : Redukált szabálybázis
- 8.
9. $q, s \leftarrow FRIQ_episode(R)$
10. $i \leftarrow 1$
11. Amíg $i \leq |R|$, addig:
12. $R_{egyformák} \leftarrow \{r \in R \mid \forall j \in \mathbb{N}, 1 \leq j \leq n: r_{ant_j} = R_{i_{ant_j}} \wedge r_{act} = R_{i_{act}}\}$
13. Ha $|R_{egyformák}| > 1$, akkor:
14. $R_{egyformák} \leftarrow R_{egyformák} \setminus R_i$
15. $R_{új} \leftarrow R \setminus R_{egyformák}$
16. $q_{új}, s_{új} \leftarrow FRIQ_episode(R_{új})$
17. $q_{diff} \leftarrow |q - q_{új}|$
18. Ha $q_{új} > q_{min} \wedge s_{új} \leq s \wedge q_{diff} \leq \varepsilon$, akkor:
19. $R \leftarrow R_{új}$
20. $q \leftarrow q_{új}$
21. $k \leftarrow |\{r \in R_{egyformák} \mid \forall j \in \mathbb{N}, 1 \leq j \leq i: r = R_j\}|$
22. $i \leftarrow i - k$
23. $i \leftarrow i + 1$
24. $R_{ij} = R$

A módszer alkalmazásával a mintapéldákon elért eredményeket a 2. táblázat tartalmazza.

2. táblázat. Egyforma szabályok közül az éppen vizsgált megtartásával elért eredmények

Mintapélda	Szabályok száma	Lépések száma	Jutalomérték
<i>Acrobot</i>	29	43	580
<i>Cartpole</i>	34	1000	7147.1689
<i>Mountaincar</i>	17	122	-210

3.2. Egyforma szabályok közül a legnagyobb vagy a legkisebb Q értékkel rendelkező megtartása

A megoldás annyiban különbözik a 3.1 fejezetben bemutatottól, hogy még mielőtt megvizsgáljuk a szabályokat, Q értékük abszolút értéke alapján csökkenő, vagy növekvő sorrendbe rendezzük azokat [4] [30]. A többi lépésben megegyezik a 3.1 fejezetben bemutatott módszerrel, így a 2. algoritmus bemenet és kimenet leírását követő sorai a következőképpen módosulnak és eredményezik a 3. algoritmust.

3. algoritmus. A legnagyobb/legkisebb Q értékű szabály megtartása egyforma szabályok esetén

8. Rendezzük R szabályait Q értékük abszolút értéke szerinti csökkenő vagy növekvő sorrendbe
9. $q, s \leftarrow \text{FRIQ_episode}(R)$
10. $i \leftarrow 1$

A módszer mintapéldákon való alkalmazásával kapott eredmények a 3. táblázatban, illetve a 4. táblázatban láthatók.

3. táblázat. Egyforma szabályok közül a legnagyobb Q értékű megtartásával elért eredmények

Mintapélda	Szabályok száma	Lépések száma	Jutalomérték
<i>Acrobot</i>	29	48	530
<i>Cartpole</i>	23	1000	7147.1689
<i>Mountaincar</i>	16	125	-240

4. táblázat. Egyforma szabályok közül a legkisebb Q értékű megtartásával elért eredmények

Mintapélda	Szabályok száma	Lépések száma	Jutalomérték
<i>Acrobot</i>	29	50	510
<i>Cartpole</i>	23	1000	7147.1689
<i>Mountaincar</i>	17	122	-210

3.3. Egyforma szabályokból új szabály létrehozása a Q értékek átlagolásával

Az előző fejezetekben bemutatott módszerekkel ellentétben ennél a megoldásnál az azonos esetet lefedő eredeti szabályok közül egyet sem tartunk meg, helyettük egy újat vezetünk be, amely ugyanazt az esetet fedí le, ugyanaz az értéke minden állapotdimenzióban és az akciódimenzióban, Q értéke pedig az eredeti szabályok Q értékeinek átlaga. Ilyen módon hasonló, vagy akár megegyező eredményt is kaphatunk, mint amit szabály interpoláció használata eleve adna az eredeti szabálybázissal, ellenben így a szabálybázis jelentősen egyszerűsödhet. Az algoritmus annyiban különbözik az előző fejezetekben bemutatottól, hogy ha vannak az éppen vizsgálttól csak konzekvensükben (Q érték) különböző szabályok, akkor

eltávolítjuk azokat az ideiglenes szabálybázisból és beszúrunk egy új szabályt, melynek az állapot- és akciódimenziókban megegyezik az értéke az eltávolított szabályokéval, Q értéke pedig azok Q értékének számtani közepe lesz. Ha az ideiglenes szabálybázissal helyesen megoldható a probléma, megtartjuk a változtatásokat. Az előző lépéseket addig ismételjük, amíg meg nem vizsgáltuk a szabálybázis összes szabályát. A 4. algoritmus mutatja be ezen módszer lépéseit.

4. algoritmus. *Redundánsnak adódott szabályok helyettesítése új szabállyal (Q értékek átlagával)*

1. **Bemenet:**
2. R : Kiinduló szabálybázis
3. q_{min} : Elvárt minimális jutalomérték
4. ε : Jutalomérték változására vonatkozó tűréshatár
5. n : Antecedensek száma
6. **Kimenet:**
7. R_{ij} : Redukált szabálybázis
- 8.
9. Rendezzük R szabályait Q érték szerinti csökkenő sorrendbe
10. $q, s \leftarrow FRIQ_episode(R)$
11. $i \leftarrow 1$
12. Amíg $i \leq |R|$, addig:
13. $R_{egyformák} \leftarrow \{r \in R \mid \forall j \in \mathbb{N}, 1 \leq j \leq n: r_{ant_j} = R_{i_{ant_j}} \wedge r_{act} = R_{i_{act}}\}$
14. $r_{új} \leftarrow R_i$
15. $r_{újconcl} \leftarrow \frac{\sum_j R_{egyformák}_{jconcl}}{|R_{egyformák}|}$
16. $R_{új} \leftarrow (R \setminus R_{egyformák}) \cup r_{új}$
17. $q_{új}, s_{új} \leftarrow FRIQ_episode(R_{új})$
18. $q_{diff} \leftarrow |q - q_{új}|$
19. Ha $q_{új} > q_{min} \wedge s_{új} \leq s \wedge q_{diff} \leq \varepsilon$, akkor:
20. $R \leftarrow R_{új}$
21. $q \leftarrow q_{új}$
22. $k \leftarrow |\{r \in R_{egyformák} \mid \forall j \in \mathbb{N}, 1 \leq j \leq i: r = R_j\}|$
23. $i \leftarrow i - k$
24. $i \leftarrow i + 1$
25. $R_{ij} = R$

5. táblázat. *Egyforma szabályokból a Q értékek átlagolásával történő szabálylétrehozással elért eredmények*

Mintapélda	Szabályok száma	Lépések száma	Jutalomérték
<i>Acrobot</i>	7	57	440
<i>Cartpole</i>	19	1000	7147.1689
<i>Mountaincar</i>	3	174	-730

3.4. Redundáns tagsági függvények eltávolítása

A redundáns antecedensek eltávolítását követően, az eredményül kapott szabálybázisban kialakulhattak olyan szabályok, melyeknek valamely antecedense törölve lett, illetve előfordulhat olyan eset is, hogy valamely antecedens minden szabályból törölve lett. Ilyen esetben, ha egyik szabályban sincs szükség egy adott antecedensre, akkor magára az antecedenshez tartozó tagsági függvényre sincs szükség, így az teljesen el is távolítható. Ennek megvalósítását mutatja be az 5. algoritmus.

5. algoritmus. *Redundánsnak adódott tagsági függvények eltávolítása a szabálybázisból*

1. **Bemenet:**
2. R : Kiinduló szabálybázis
3. n : Antecedensek száma
4. **Kimenet:**
5. R_{ij} : Redukált szabálybázis
- 6.
7. $i \leftarrow 1$
8. Amíg $i \leq n$, addig:
9. Ha $\forall r \in R: r_{ant_i} = \emptyset$, akkor:
10. Töröljük R i . oszlopát
11. Töröljük az i . antecedenshez tartozó állapotleírókat, tagsági függvényeket
12. $i \leftarrow i + 1$
13. $R_{ij} = R$

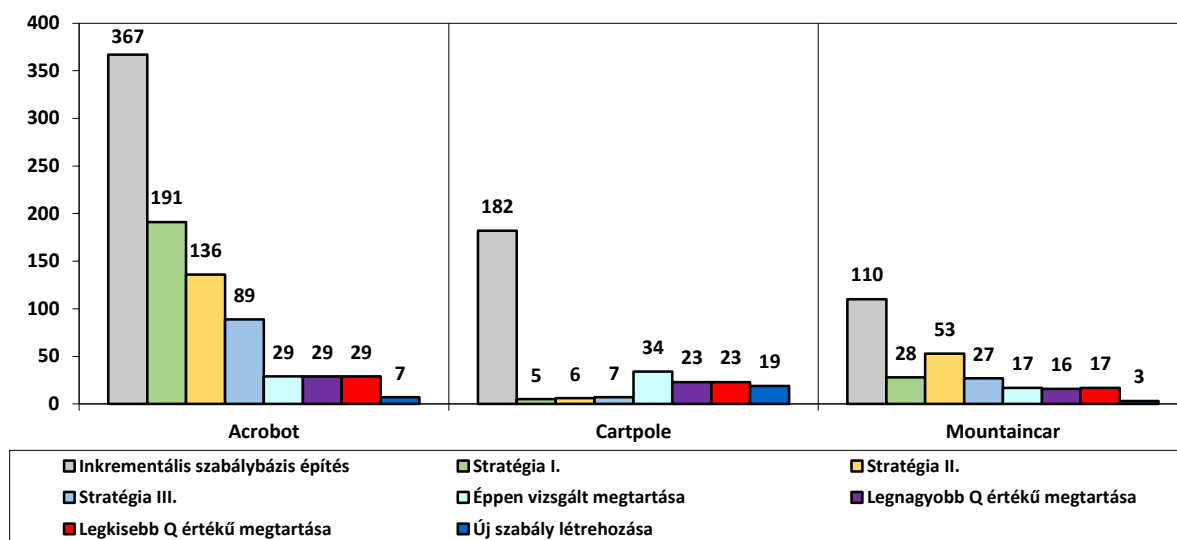
4. Összefoglalás

A bemutatott fuzzy szabálybázis redukciós módszerekkel hatékonyabbá tehető a FRIQ-learning megerősítéses tanulási módszerben használt automatikus tudáskinyerési eljárások [26]. A korábbi FRIQ-learning-ben használt redukciós módszerekkel ellentétben, amelyek csak teljes szabályokkal képesek dolgozni, a bemutatott új redukciós módszerek képesek a szabályok redundáns részeinek azonosítására, ezáltal a felesleges szabályrészek és szabályok eltávolíthatók a szabálybázisból. Bizonyos esetekben jelentősen egyszerűsíthetők a szabálybázisok az új módszerek felhasználásával. Az 1. ábrán láthatóak az eredeti, teljes szabályokkal dolgozó, stratégiák (Stratégia I-III. [28] [30]), illetve az új, antecedens redundanciát vizsgáló, módszerek által létrehozott szabálybázisok méretei (szabályok száma).

Továbbá olyan problémák megoldásánál hasznosulhatnak még ezen újonnan kifejlesztett eljárások, ahol a probléma leírásához használt bemenetek megfigyeléseken, illetve feltételezéseken alapulnak, például viselkedéseket leíró modellekben [8] [12] [13] [14] [24] [31]. Ilyen és hasonló problémákat előzetes szakértői tudás alapján lehet leírni [21], akár hiányos (ritka) fuzzy szabálybázisként is, illetve az erre a célra kifejlesztett leíró nyelven [15] [16]. A kifejlesztett új módszer az ilyen problémák megoldása után képes azonosítani azokat a bemeneti változókat, amikre, a szakértői tudás alapján, feltételezetten szükség volt, de valójában azok nélkül is megoldható a probléma. Ezzel rámutatva jelentős egyszerűsítési lehetőségekre a probléma leírásában, illetve korábbi összefüggések felülvizsgálatának szükségességére a szakértői tudásbázisban. Abban az esetben is hasznos a kifejlesztett algoritmus, amikor a szabálybázis nagy mérete miatt a fuzzy szabály interpolációs módszernek is jelentősen megnövekszik a számítási kapacitás igénye. Bár vannak kezdeményezések a FIVE FRI párhuzamosítására [2] [32], illetve hardver

specifikus implementációra [3], a szabálysám hatékony csökkentése jelentős javulást hozhat ezen a téren is.

A kifejlesztett módszer így várhatóan hatékonyabbá teszi a FRIQ-learning alkalmazhatóságát tudáskinyerésre [26] a szabálybázisok olvashatóságának könnyítésével.



1. ábra. Megoldáshoz szükséges szabályok száma különböző redukciós stratégiák (a bemutatott új stratégiákat is beleértve) alkalmazása után

5. Köszönetnyilvánítás

A cikkben ismertetett kutató munka az EFOP-3.6.1-16-2016-00011 jelű „Fiatalodó és Megújuló Egyetem – Innovatív Tudásváros – a Miskolci Egyetem intelligens szakosodást szolgáló intézményi fejlesztése” projekt részeként – a Széchenyi 2020 keretében – az Európai Unió támogatásával, az Európai Szociális Alap társfinanszírozásával valósul meg.

Irodalom

- [1] Appl, M.: *Model-based Reinforcement Learning in Continuous Environments*, München, Germany: Ph.D. thesis, Technical University of München, 2000.
- [2] Bartók, R., Vásárhelyi, J.: *Parallelization of FIVE method on multicore embedded system*, 2018 19th International Carpathian Control Conference (ICCC), Szilvasvarad, 2018, pp. 400-403. <https://doi.org/10.1109/CarpathianCC.2018.8399663>
- [3] Bartók, R., Vásárhelyi, J.: *A fuzzy rule interpolation base algorithm implementation on different platforms* Proc. of the 16th International Carpathian Control Conference (ICCC) pp. 37-40 May 27-30 2015. <https://doi.org/10.1109/CarpathianCC.2015.7145041>
- [4] Bellman, R. E.: *Dynamic Programming*, Princeton, NJ: Princeton University Press, 1957.
- [5] Berenji, H. R.: *Fuzzy Q-Learning for Generalization of Reinforcement Learning*, Proceedings of the 5th IEEE International Conference on Fuzzy Systems, pp. 2208-2214, 1996.
- [6] Bonarini, A.: *Delayed Reinforcement, Fuzzy Q-Learning and Fuzzy Logic Controllers*, Herrera, F., Verdegay, J. L. (Eds.), Genetic Algorithms and Soft Computing, pp. 447-466, 1996.

- [7] Horiuchi, T., Fujino, A., Katai, O., Sawaragi, T.: *Fuzzy Interpolation-Based Q-learning with Continuous States and Actions*, Proceedings of the 5th IEEE International Conference on Fuzzy Systems, pp. 594-600, 1996.
- [8] Johanyák, Zs. Cs.: *Fuzzy rule interpolation based model for student result prediction* in Journal of Intelligent and Fuzzy Systems. Vol. 36 (2). 2019, pp. 999-1008.
<https://doi.org/10.3233/JIFS-169875>
- [9] Kovács, Sz., Kóczy, L. T.: *Approximate Fuzzy Reasoning Based on Interpolation in the Vague Environment of the Fuzzy Rule base as a Practical Alternative of the Classical CRI*, Proc. of the 7th International Fuzzy Systems Association World Congress, Prague, Czech Republic, pp. 144-149, 1997.
- [10] Kovács, Sz.: *New Aspects of Interpolative Reasoning* in Proc. of the 6th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Granada, Spain, pp. 477-482, 1996.
- [11] Kovács, Sz., Kóczy, L. T.: *The use of the concept of vague environment in approximate fuzzy reasoning* in Fuzzy Set Theory and Applications, Tatra Mountains Mathematical Publications, Mathematical Institute Slovak Academy of Sciences, Bratislava, Slovak Republic, vol. 12, 1997, pp. 169-181.
- [12] Kovács, Sz., Vincze, D., Gácsi, M., Miklósi, Á., Korondi, P.: *Interpolation based Fuzzy Automaton for Human-Robot Interaction*, in Proc. of SYROCO2009, 9th IFAC Symposium on Robot Control, Gifu, Japan, Sep 9-12. 2009.
- [13] Kovács, Sz., Vincze, D., Gácsi, M., Miklósi, Á., Korondi, P.: *Ethologically Inspired Robot Behavior Implementation*, in Proceedings of the 4th International Conference on Human System Interaction (HSI 2011), Keio University, Yokohama, Japan, pp. 64-69, May 19-21, 2011.
<https://doi.org/10.1109/HSI.2011.5937344>
- [14] Kovács, Sz., Vincze, D., Gácsi, M., Miklósi, Á., Korondi, P.: *Fuzzy automaton based Human-Robot Interaction*, in Proc. of the 8th IEEE International Symposium on Applied Machine Intelligence and Informatics (SAMI), 28-30 Jan 2010, pp. 165-169.
<https://doi.org/10.1109/SAMI.2010.5423746>
- [15] Piller, I., Vincze, D., Kovács, Sz.: *Declarative language for behaviour description*. in Emergent trends in robotics and intelligent systems. Springer, 2015, pp. 103-112.
https://doi.org/10.1007/978-3-319-10783-7_11
- [16] Piller, I., Kovács, Sz.: *Fuzzy Behavior Description Language: A Declarative Language for Interpolative Behavior Modeling* in Acta Polytechnica Hungarica, 16 (9), pp. 48-72, 2019.
- [17] Precup, D., Sutton, R. S., Dasgupta, S.: *Off-policy temporal-difference learning with function approximation*, Proceedings of the 18th International Conference on Machine Learning, pp. 417-424, 2001.
- [18] Rummery, G. A., Niranjan, M.: *On-line Q-learning using connectionist systems*, Cambridge University, UK: CUED/F-INFENG/TR 166, 1994.
- [19] Sutton, R. S., Barto, A. G.: *Reinforcement Learning: An Introduction*, Cambridge, Massachusetts: The MIT Press, 2018.
- [20] Tikk, D., Johanyák, Zs. Cs., Kovács, Sz., Wong, K. W.: *Fuzzy rule interpolation and extrapolation techniques: Criteria and evaluation guidelines* in Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 15 (3), 2011, pp. 254-263.
- [21] Tompa, T., Kovács, Sz. (2018): *Szakértői tudás alapú FRIQ-learning*. SzámOkt 2018 XXVIII. Nemzetközi Számítástechnika és Oktatás Konferencia, Tusnádfürdő.
- [22] Tompa, T., Kovács, Sz.: *Clustering-based fuzzy knowledgebase reduction in the FRIQ-learning*

- in Proc. of the 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI2017), Herl'any, Slovakia, pp. 197-200, Jan 26-28, 2017.
<https://doi.org/10.1109/SAMI.2017.7880302>
- [23] Tompa, T., Kovács, Sz.: *Determining the minimally allowed rule-distance for the incremental rule-base construction phase of the FRIQ-learning*, in Proc. of the 19th International Carpathian Control Conference (ICCC2018), pp. 480-483, May, 2018.
<https://doi.org/10.1109/CarpathianCC.2018.8399677>
- [24] Tóth, A., Vincze, D.: *Futball szimuláció megvalósítása fuzzy szabály interpoláció alapú fuzzy automatával*, Multidiszciplináris Tudományok, Évf. 9 szám 1 (2019), pp. 12-22.
<https://doi.org/10.35925/j.multi.2019.1.2>
- [25] Vincze, D.: *Fuzzy Rule Interpolation and Reinforcement Learning* in Proc. of the 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI 2017), Herl'any, Slovakia, pp. 173-178, January 26-28, 2017. <https://doi.org/10.1109/SAMI.2017.7880298>
- [26] Vincze, D.: *Automatikus tudáskinyerés fuzzy szabály interpoláció alapú Q-tanulással*, Multidiszciplináris Tudományok, 4 (1), pp. 135-146.
- [27] Vincze, D., Kovács, Sz.: *Fuzzy Rule Interpolation-based Q-learning*, Proceeding of the 5th International Symposium on Applied Computational Intelligence and Informatics (SACI 2009), Timisoara, Romania, pp. 55-59, 2009. <https://doi.org/10.1109/SACI.2009.5136311>
- [28] Vincze, D., Kovács, Sz.: *Reduced Rule Base in Fuzzy Rule Interpolation-based Q-learning*, Proceedings of the 10th International Symposium of Hungarian Researchers on Computational Intelligence and Informatics, CINTI 2009, Budapest Tech, Budapest, pp. 533-544, 2009. <https://doi.org/10.1109/SACI.2009.5136311>
- [29] Vincze, D., Kovács, Sz.: *Incremental Rule Base Creation with Fuzzy Rule Interpolation-Based Q-Learning*, Judas, I. J. et al. (Eds.), Computational Intelligence in Engineering, Studies in Computational Intelligence, Vol. 313/2010, Springer-Verlag, Berlin Heidelberg, pp. 191-203, 2010. https://doi.org/10.1007/978-3-642-15220-7_16
- [30] Vincze, D., Kovács, Sz.: *Rule-Base Reduction in Fuzzy Rule Interpolation-Based Q-Learning*, Recent Innovations in Mechatronics (RIIM) Vol. 2. (2015) No. 1-2., 2015. <https://doi.org/10.17667/riim.2015.1-2/10>
- [31] Vincze, D., Kovács, Sz., Niitsuma, M., Hashimoto, H., Korondi, P., Gácsi, M., Miklósi, Á.: *Ethologically inspired human-robot interaction interfaces* In Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments, 2012, Hamamatsu, Japan, pp. 51-57. <https://doi.org/10.1145/2160749.2160761>
- [32] Vincze, D.: *Parallelization by Vectorization in Fuzzy Rule Interpolation Adapted to FRIQ-Learning* 2018 World Symposium on Digital Intelligence for Systems and Machines (DISA) pp. 131-136 2018. <https://doi.org/10.1109/DISA.2018.8490614>
- [33] Watkins, C. J. C. H.: *Learning from Delayed Rewards*, Cambridge, England: Ph.D. thesis, Cambridge University, 1989.
- [34] Watkins, C. J. C. H., Dayan, P.: *Q-learning*, Machine Learning, vol. 8. no. 3, pp. 279-292, 1992. <https://doi.org/10.1023/A:1022676722315>
- [35] White, C. C.: *Markov decision processes*, Encyclopedia of Operations Research and Management Science, Boston, MA, Springer, 2001.