FIATAL MŰSZAKIAK TUDOMÁNYOS ÜLÉSSZAKA

Kolozsvár, 1999. március 19-20.

Szimbolikus és szub-szimbolikus módszerek az analitikailag kezelhető problémák megoldásában

Stefán Péter

Abstract

This paper gives a brief overview of the importance of learning in industrial and business applications. Firstly the question why the learning itself is substantial in manufacturing tasks will be pondered and a general method, namely the reinforcement learning procedure will be discussed. Then the main concept of the reinforcement technique and the stability of a specified REINFORCE learning rule will be introduced through stochastic real-valued (SRV) units. The relationship of the reinforcement learning with back-propagation will also be presented.

1. Bevezetés

Napjainkra a gyártórendszerek tanulási képességgel történő felruházása egyre égetőbb feladattá vált. Egy tetszőleges rendszer tanulását a következőképp definiálhatjuk: a tanulás nem más, mint a rendszeren belül végrehajtott paraméter és/vagy struktúra változtatás, mely lehetővé teszi azt, hogy a rendszer növekvő teljesítményt nyújtson. Miért jelentős a tanulás? Erre a kérdésre a válasz az, hogy a modellezésnek egy új módját tárja fel. Egy rendszermodell megalkotható matematikai úton, mely egzakt és pontos megoldást ad, de az esetek zömében nehezen kivitelezhető, rengeteg differenciálegyenlet meghatározását igényli. A tapasztalatok, heurisztikák alkalmazása bár egyszerűsíti a komplexitást, de a környezet változásaira érzéketlen modelleket eredményezhetnek. A harmadik megoldás olyan modelleket alkotni, melyek képesek megtanulni tetszőleges rendszerek viselkedését. Ezek a modellek rendelkeznek egy tanulási koncepcióval, mely koncepció zártsága matematikai eszközökkel igazolható.

2. Megerősítő tanulás

A tanulási módszerek az alábbi kategóriákba sorolhatók:

- felügyelt,
- felügyelet nélküli,
- megerősítő.

A felügyelt tanulás jellemzője, hogy a tanuló-rendszeren kívül található egy külső referencia, egy tanár, aki minden pillanatban össze tudja hasonlítani a rendszer aktuális kimeneti értékeit a kívánt kimeneti értékekkel, és ezekből egy hibajelet, az esetek többségében egy hibavektort képes képezni. Ezen hibavektor hossza fordítottan arányos a rendszer „tudásával”. A felügyelet nélküli tanulási módszerek nem rendelkeznek külső referenciával, a működésük leggyakrabban önszervezésen alapul. Éppen emiatt alkalmazhatóságuk roppant korlátozott. A megerősítő tanulás átmenetet képez a felügyelt és a felügyelet nélküli módszerek között; itt is megtalálható a külső referencia, de ez egy vektor-jel helyett mindössze egy skaláris jelet ad vissza, melyben értékeli, osztályozza a rendszer teljesítményét. A megerősítő tanulás algoritmusának éppen az a lényege, hogy hogyan lehet ebből a skaláris információból a rendszerben változtatásokat végrehajtani. *Nuttin* [5] a disszertációjában a „zenész példáját” említi meg, mint jellegzetes példát megerősítő tanulásra (1. ábra).

1. ábra: a megerősítő tanulás koncepciója a „zenész példáján” keresztül

Az ábra bal oldala a zenészt, mint alrendszert ábrázolja, a jobb oldala pedig annak környezetét. A környezet kiértékeli a muzsikusi teljesítményt (nagy taps-kis taps) és neki ebből az információból konklúziókat kell levonnia a saját teljesítményét illetően. Ha a visszajelzett jel nagy értékű, akkor ez azt jelenti, hogy a közönség által kedvelt zenét játszik, és ez megerősíti abban, hogy helyes paraméter-beállításokkal játszik. Ha azonban ez a jel kis értékű, akkor a környezet nem honorálta a művészetét, és (feltételezve azt, hogy a zenész a közönség maximális tetszését törekszik elnyerni) emiatt paraméter és/vagy struktúra változtatásokat kénytelen végrehajtani. A lehetséges új paraméterekkel neki kell kísérletezni, nem a környezet mondja meg, emiatt szükség van a lehetséges paraméter-beállítások terében keresést, felderítést végrehajtani. A felderítés a megerősítő tanulás egy újabb jellemvonása.

3. A megerősítő tanulás matematikai modellje

A megerősítő tanulás koncepciója általános, alkalmazható többféle tanuló-rendszerre, úgymint tanuló-automatákra, és neurális hálókra. Itt a koncepció SRV (sztochasztikus valós-értékű) egységeken kerül bemutatásra. A sztochasztikus jelleg onnan adódik, hogy egy SRV egység kimenete egy véletlen-szám, melyet egy véletlen-szám generátor állít elő, normális eloszlásból. A normális eloszlás két paraméterrel rendelkezik: átlag (μ) és szórás (σ). Az SRV egységekkel történő implementáció a 2. ábrán látható.

2. ábra: a megerősítő tanulás koncepciója SRV egységek segítségével

Az eloszlás két paramétere viszont determinisztikus úton kerül meghatározásra, kiszámításukat két egyszerű neuron végzi az alábbi képletek felhasználásával:

$$\begin{aligned}\mu &= f(\mathbf{w}_\mu \mathbf{x}), \\ \sigma &= f(\mathbf{w}_\sigma \mathbf{x}),\end{aligned}\tag{1}$$

ahol a $\mathbf{w}_\mu$ az átlagot kiszámító neuron súlyvektora, a $\mathbf{w}_\sigma$ pedig a szórást kiszámító neuroné. Az $\mathbf{x}$ vektor a két neuron közös bemenete (az ábrán $I_1, I_2, \dots, I_N$ jelöli), f pedig a neuronok átviteli függvénye.

A tanulás itt egy maximumkeresési folyamat, mely során olyan paraméter beállításokat választunk, mely során a visszajelzett megerősítő jel várható értéke maximális. A tanulás SRV egységek esetén kétlépcsős (Williams [1]). Az első lépésben a megerősítő jelből a normális eloszlás paraméter-változásait vezetjük le, majd második lépésben a paraméter változásokból levezetjük a neuronok súlyainak megváltoztatását. Az első lépés végrehajtására az alábbi képletek szolgálnak:

$$\Delta\mu = \alpha_\mu (r - b_\mu) \frac{y - \mu}{\sigma^2}\tag{2}$$

$$\Delta\sigma = \alpha_\sigma (r - b_\sigma) \frac{(y - \mu)^2 - \sigma^2}{\sigma^3}\tag{3}$$

Ezen paraméter-változtatások az alábbiakat jelentik: a μ megváltoztatása azt jelenti, hogy a bemenethez a leginkább illeszkedő kimenetet keressük, a σ változtatása pedig azt jelenti, hogy az előbbi érték bizonytalanságát növeljük vagy csökkentjük, ami kisebb illetve nagyobb véletlenszerűséget, kisebb, illetve nagyobb felderítési képességet enged meg. A képletben a $\Delta\mu$, illetve a $\Delta\sigma$ az átlag, illetve a szórás változásai, az α_μ , illetve a α_σ a tanulási együtthatók, r a megerősítő jel, y a rendszer aktuális kimenete.

A tanulás második lépésében egy közönséges Back-Propagation eljárásen keresztül az előbbi paraméter-változtatásokat visszavezetjük súlyváltozásokra:

$$\Delta \mathbf{w}_\mu = -\eta \frac{\partial(\Delta\mu)}{\partial \mathbf{w}_\mu} + \gamma \mathbf{w}_\mu \quad (4)$$

$$\Delta \mathbf{w}_\sigma = -\eta \frac{\partial(\Delta\sigma)}{\partial \mathbf{w}_\sigma} + \gamma \mathbf{w}_\sigma \quad (5)$$

Itt a $\Delta \mathbf{w}_\mu$ illetve a $\Delta \mathbf{w}_\sigma$ a súlyváltoztatások, az η a tanulási együttható, γ az ún. tanulási momentum, a $\frac{\partial}{\partial \mathbf{w}}$ pedig strukturált deriválás operátor.

4. Felhasznált irodalom

- [1] Ronald J. Williams: Simple statistical gradient-following algorithms for connectionist reinforcement learning, Machine Learning, 1992, pp. 229-256
- [2] Vijaykumar Gullapalli: A stochastic reinforcement learning algorithm for learning real-valued functions, Neural Networks, 1990, pp. 671-692
- [3] L. Monostori, A. Márkus, H. Van Bussel, E. Westkämpfer: Machine learning approaches to manufacturing, Annals of the CIRP, 1996, pp. 675-712
- [4] R. Teti, S.R.T. Kumara: Intelligent computing methods for manufacturing systems, Annals of the CIRP, 1997, pp. 1-24
- [5] Marnix Nuttin: Learning Approaches to Robotics and Manufacturing
- [6] R.J. Kuo, K.C. Xue: An intelligent sales forecasting system through integration of artificial neural network and fuzzy neural network.

Név: Stefán Péter, doktorandusz hallgató

Munkahely: Alkalmazott Informatikai Tanszék, Miskolci-Egyetem,

3515 Miskolc-Egyetemváros, Magyarország

E-mail: stefan@iit.uni-miskolc.hu