

A Magyar Közös Tudás projekt első eredményei

Kusper Gábor, Tajti Tibor, Szabó Máté
{gkusper,tajti,tenor}@aries.ektf.hu
Eszterházy Károly Főiskola

Összefoglaló

A közös tudás az a tudás, ami mindennapjainkban segít eligazodni. Ennek segítségével tudjuk hirtelen eldönteni, hogy egy helyzet meglepő-e vagy sem. Például, ha azt látjuk, hogy egy kutya ugat, akkor ezen nem lepődünk meg, hiszen tudjuk, hogy a kutyák szoktak ugatni. Ha viszont egy kutya megszólalna, akkor igencsak meglepődnénk, hiszen tudjuk, hogy a kutyák nem tudnak beszélni. A mesterséges intelligencia területén a közös tudás olyan adatbázist jelent, amelyben az emberek által igaznak vélt mondatok találhatóak. Itt ne az emberiség matematika vagy egyéb tudomány eredményeire gondoljunk, hanem az emberek közös tapasztalataira. A Magyar Közös Tudás projekt célkitűzése magyar nyelvű mondatok formájában gyűjteni a közös tudást. Elkészült a weblapunk, amely a <http://commonsense.ektf.hu/> címen érhető el. Három hónap alatt közel 10ezer mondatot gyűjtöttünk. A mondatok felvételét sablonok segítik, de van egy szabad szöveges beviteli mező is. A mondatokat a Self-Organizing Map (SOM) típusú neuronháló segítségével vizsgáljuk. Az első SOM képes eldönteni egy új mondatról, hogy az eddigi mondatok alapján beletartozik-e a közös tudásba vagy sem. Közeli terveink között szerepel a mondatláncok kialakítása, amikor is kiindulunk egy mondatból és a „De miért?” kérdést feltéve további mondatokkal bővítjük a láncot. Távoli célunk, hogy a közös tudás alap problémáit felsoroló Common Sense Problem Page oldalon felsorolt problémákat megoldó programot hozzunk létre az adatbázis alapján.

The first results of the Hungarian Common Sense project

Abstract

The common sense is the knowledge which helps us to know our way about everyday problems. With the help of common sense we can decide whether a situation is strange or not. For example, if we see that a dog is barking, then this is not strange for us, because we know that dogs can bark. But if a dog starts to talk to us, then we would be astonished, since we know that dogs cannot talk! In the field of artificial intelligence common sense means a database which contains such sentences which are commonly thought to be true. One should not think on mathematical knowledge or on other sciences, but on the common experiences of the mankind. The Hungarian Common Sense project aims to collect the common sense in Hungarian language. Our website can be found here: <http://commonsense.ektf.hu/>. We could collect almost 10 thousand sentences in 3 months. We have templates to help to add new sentences and a free text input field too. We analyze the sentences by Self-Organizing Map (SOM) neural networks. The first one is able to tell about a new sentence whether it belongs to the common sense or not based on the sentences which we already have. Our short term goal is to collect sentence chains. We start a chain from a sentence from our database and ask the question „But why?”. The answer is inserted into the chain. Our long term goal is to develop an application which can solve problems from the Common Sense Problem Page. This program should be based on our database.

1. Bevezetés

A közös tudás az a tudás, ami mindennapjainkban segít eligazodni. Ennek segítségével tudjuk hirtelen eldönteni, hogy egy helyzet meglepő-e vagy sem. Például, ha azt látjuk, hogy egy kutya ugat, akkor ezen nem lepődünk meg, hiszen tudjuk, hogy a kutyák szoktak ugatni. Ha viszont egy kutya megszólalna, akkor igencsak meglepődnénk, hiszen tudjuk, hogy a kutyák nem tudnak beszélni. A mesterséges intelligencia területén a közös tudás olyan adatbázist jelent, amelyben az emberek által igaznak vélt mondatok találhatóak. Itt ne az emberiség matematika vagy egyéb tudomány eredményeire gondoljunk, hanem az emberek közös tapasztalataira. A Magyar Közös Tudás projekt célkitűzése magyar nyelvű mondatok formájában gyűjteni a közös tudást. Elkészült a weblapunk, amely a <http://commonsense.ektf.hu/> címen érhető el. Három hónap alatt közel 10ezer mondatot gyűjtöttünk. A mondatok felvételét sablonok segítik, de van egy szabad szöveges beviteli mező is. A mondatokat a Self-Organizing Map (SOM) típusú neuronháló segítségével vizsgáljuk. Az első SOM képes eldönteni egy új mondatról, hogy az eddigi mondatok alapján beletartozik-e a közös tudásba vagy sem. Közeli terveink között szerepel a mondatláncok kialakítása, amikor is kiindulunk egy mondatból és a „De miért?” kérdést feltéve további mondatokkal bővítjük a láncot. Távoli célunk, hogy a közös tudás alap problémáit felsoroló Common Sense Problem Page oldalon felsorolt problémákat megoldó programot hozzunk létre az adatbázis alapján.

2. Hasonló rendszerek

A közös tudás (common sense) mesterséges intelligencia belüli használatát Marvin Minsky vezette be és fejte ki két könyvében: The Society of Mind [Minsky1988], The Emotion Machine [Minsky2006].

2.1. Open Mind Common Sense projekt

Több közös tudás projekt is ismert. Ezek közül az egyik leginkább hasonlatos a mi kezdeményezésünkhöz az Open Mind Common Sense projekt [OpenMind, OpenMind2002], amely az MIT gondozása alatt működik. Ez a projekt 1999-ben indult és az eltelt idő alatt több mint egy millió angol nyelvű mondatot gyűjtöttek. A projekt érdekessége, hogy önkéntesek újabb nyelvekkel is bővíthetik az adatbázist, így például magyar nyelvű mondatokat is gyűjtenek. A cikk írásakor 2158 magyar nyelvű mondatot tartalmaz az adatbázisuk.

Az Open Mind Common Sense projekt keretében gyűjtött mondatokat használja fel a ConceptNet szemantikus háló és az AnalogySpace következtető rendszer, amely dimenzió redukción alapszik.

2.2. OpenCyc 2.0 rendszer

Az OpenCyc 2.0 rendszer [Cyc, Cyc1995] 2009 júliusától elérhető el. Az adatbázisba 47 000 fogalom és 306 000 állítás található. Az OpenCyc a fizetős Cyc technológia szabadon elérhető verziója. A mögöttes ontológia OWL formátumban is letölthető.

Egy példa: „A tree is a plant.” Ennek a belső reprezentációja: (#\$isa #\$Tree-ThePlant #\$Plant).

2.3. Wolfram Alpha rendszer

A Wolfram Alpha [Alpha] rendszer egy kérdés megválaszoló rendszer, amelyet a Wolfram Research cég fejlesztett ki, amely a jól ismert Mathematica rendszert is fejleszti. A rendszer viszonylag kevés szabályon viszont több mint 1000 adatbázison alapszik (pl. élelmiszer összetevők, aranykőpések, térképek, ...).

Példa kérdés: Two things are infinite?

Válasz: The universe and human stupidity; and I'm not sure about the universe. (according to Albert Einstein)

De ha átfoglazzuk a kérdést: „Two things which are infinite?” , akkor már nem kapunk rá választ.

2.4. YAGO2

A YAGO2 [YAGO2, YAGO2007] egy szemantikus tudás bázis (semantic knowledge base), amely a benne tárolt tudást a Wikipedia, a WordNet és a GeoNames szolgáltatásokból nyeri. Tehát amíg a Wolfram Alpha mögött nagyon sok specifikus adatbázis van, e mögött három eléggé általános, bár a GeoNames egyértelműen földrajzi adatokat tartalmaz. Az összegyűjtött adatbázis 10 millió fogalmat és 460 millió tényt tartalmaz.

Példa kérdés: Scientists who were born nearby Ulm and won a Nobel Prize (in any discipline)

Válasz: Albert Einstein

2.5. DBpedia

A DBpedia [DBpedia, DBpedia??] rendszer is a Wikipedia szolgáltatásból tölt le strukturált információt. A DBpedia legnagyobb hozzájárulása a területhez a SPARQL lekérdező nyelv kifejlesztése, amivel információt lehet kinyerni az adatbázisból.

Egy példa SPARQL lekérdezés:

```
PREFIX dbo: <http://dbpedia.org/ontology/>
```

```
SELECT ?manufacturer ?name ?car
```

```
WHERE {
```

```
    ?car <http://purl.org/dc/terms/subject> <http://dbpedia.org/resource/Category:Luxury_vehicles>
```

```
    ?car foaf:name ?name .
```

```
    ?car dbo:manufacturer ?man .
```

```
    ?man foaf:name ?manufacturer
```

```
}
```

```
ORDER by ?manufacturer ?name
```

A PREFIX részben adjuk meg a felhasználni kívánt ontológiát. A SELECT kulcsszó után nevek következnek. Ezeket kérdőjellel kell kezdeni. A WHERE záradékban adhatók meg a keresési feltételek, az ORDER BY záradék a rendezésre vonatkozik. Látható, a SPARQL nagyon hasonlít az SQL lekérdező nyelvhez a szintaxis szintjén.

3. A Magyar Közös Tudás projekt

A Magyar Közös Tudás projekt 2011-ben kezdődött az Eszterházy Károly Főiskola Matematikai és Informatikai Intézetében. Ez az első cikk, amelyben a projekt első eredményeit tesszük közvé. A közös tudás fogalmának meghatározása során a következő értelmezések hatottak ránk:

- A közös tudás a mesterséges intelligencia azon része, amely arra hivatott, hogy az emberek közös tudását rögzítse. Itt ne az emberiség matematika vagy egyéb tudomány eredményeire gondoljunk, hanem az emberek közös tapasztalataira.
- A közös tudás az a tudás, ami mindennapjainkban segít eligazodni. Ennek segítségével tudjuk hirtelen eldönteni, hogy egy helyzet meglepő-e vagy sem. Például, ha azt látjuk, hogy egy kutya ugat, akkor ezen nem lepődünk meg, hiszen tudjuk, hogy a kutyák szoktak ugatni. Ha viszont ez a kutya hozzánk szólna, hogy „Mikor megyünk már sétálni?“, akkor igencsak meglepődnénk, hiszen tudjuk, hogy a kutyák nem tudnak beszélni.
- Másként megfogalmazva, az hogy egy kutya beszél, az már túlesik az elvári horizontunkon, azaz nem felel meg annak az elvárásnak, amit magunkban előre megfogalmazunk, amikor felkészülünk egy helyzet értelmezésére. Habár minden embernél más és más az elvárési horizont, azt lehet mondani, hogy az emberek elvárési horizontja tart egy határértékhez. Ami ez alatt van, az a közös tudás része.
- Más felfogás szerint a közös tudás maga a józan ész, vagy még inkább a józan paraszti ész. A józan paraszti eszünkkel próbáljuk a felmerülő problémákat józanul, a lehető legegyszerűbben megoldani, úgy hogy ne menjünk fejfel a falnak. Például nyaralni nyáron menjünk, ne télen. Vagy, ha már zsibbad a lábad egy testhelyzetben, akkor válts testhelyzetet.
- A közös tudást tekinthetjük a közmondásokban, szólásmondásokban, anekdotákban, viccekben megtestesülő bölcsességeknek is. Például nem tanácsos hitelt felvenni, mert: Csak addig nyújtózkodj, amíg a takaród ér!
- Megint más felfogás szerint a közös tudás segítségével tudjuk eldönteni, hogy mi lenne a társadalmilag elvárt cselekvés saját magunktól, azaz mit csinálna helyemben a velem szembe sétáló ismeretlen. Például, ha valaki teherbe ejt egy lányt, akkor feleségül kell vennie.
- A közös tudást szemlélhetjük úgy is, mint egy szintjét a saját cselekedeteink ellenőrzésének. Az ember tudatosan vagy félig tudatosan próbálja értelmezni saját cselekedeteit, megpróbálja ellenőrizni, hogy azt csinálta-e, amit akart. Ha egy nem sokkal ezelőtti cselekedetünk nem felel meg a közös tudásnak, akkor egy kicsit rosszul érezzük magunkat. Például, ha egy férfi nem enged előre egy hölgyet az ajtónál, akkor egy kis lámpa felvillan az agyában, hogy ez nem volt illedelmes.
- A fenti példából látszik, hogy a közös tudáshoz tartoznak az illemszabályok, a tradíciók és a hagyományok is.
- Megint más megfogalmazásban a közös tudáshoz tartozik egy tudás darabka, mondjuk egy mondat, ha ezt a mondatot 10 ismerősödből 10 elfogadja a közös tudáshoz tartozónak. Például, ha egy baráti körben mindenki egyetért azzal, hogy „Judit jó csaj“, akkor ez a közös tudás segíti, hogy megértsék egymást fél szavakból.
- A fent példából látható, hogy minden emberi közösségnek van saját csoport tudása, amely a közösség hiedelmire, szertartásaira vonatkozó ismereteket tartalmazza. Ebben az értelemben a közös tudás ezen csoport tudások közös része.

Tehát a közös tudást sokféleképpen lehet definiálni. Mi ezek közül leginkább azt fogadjuk el, hogy a közös tudás az a része az intelligenciának, ami alapján eldöntjük, hogy egy helyzet meglepő-e vagy sem.

A közös tudást esetünkben egy mondat gyűjteménnyel reprezentáljuk. Ez egy olyan mondat gyűjtemény, amely olyan mondatokat tartalmaz, amit mindenki igaznak fogad el. Gyakran magát a mondat gyűjteményt is közös tudásnak hívjuk. A közös tudáshoz tartozó néhány mondat:

- Minden kutya tud ugatni.
- A fű zöld.

Nagyon fontos megjegyezni, hogy a közös tudáshoz nagyon sok olyan mondat is tartozik, amely matematikai logikai értelemben hamis. Például: „Az ember tud beszélni.” Könnyen belátható, hogy ez a mondat nem igaz, hiszen a csecsemők még nem tudnak beszélni. Ugyanakkor ez a mondat része a közös tudásnak, hiszen a természetes nyelv pontatlanságával azt a logikai értelemben is igaz mondatot fejezi ki, hogy: „Az emberi faj egyedei képesek beszéd útján kommunikálni, hacsak nem csecsemők, vagy némák, vagy egyéb, a beszéd képességét akadályozó módon hátráltatottak.”

Azt is fontos megjegyezni, hogy mondatok a közös tudás szempontjából lehetnek azonos jelentésűek. Ilyenkor ugyanazt a tudást fejezik ki a rövid, frappáns mondatok, amiket esetleg félre is lehet érteni, és a hosszabb, több részletet tartalmazó, nehezebben félreérthető mondatok. Például:

- A kutya ugat.
- A kutyák tudnak ugatni.
- Minden kutya tud ugatni.
- A kutyák szoktak ugatni.

A közös tudás szempontjából azonos jelentésű mondatok automatikus megtalálására nincs a szerzők által ismert algoritmus.

További problémát jelent, hogy a magyar mondatokban nincs kötött szórend és az ugyanazokból a szavakból álló, hasonló jelentésű mondatok mégse jelentik teljesen ugyanazt, hiszen az első szó a hangsúlyos. Példa:

- A fű zöld. (Milyen a fű?)
- Zöld a fű. (Mi zöld?)

3.1. A rendszer bemutatása

A közös tudás összegyűjtésére létrehoztunk egy web felületet, amely jelenleg a következő címen érhető el: <http://commonsense.ektf.hu/>

A közös tudás összegyűjtését önkéntesek bevonásával tervezzük. A munka segítése érdekében sablonokat adunk a feltöltőknek. Néhány sablon minta:

- A(z) PI.: A fű zöld.
- A(z) tud-ni. PI.: A kutya tud ugatni.

A sablonok több szempontból hasznosak. Egyrészt segítik a feltöltők munkáját, másrészt a mondatok értelmezését is segítik. Ez utóbbi sajnos csak korlátozottan igaz, hiszen az első mintával többfajta

nyelvtani szerkezetű, és így másképp értelmezhető (ontológiává lefordítható), mondatot adhatunk meg. Például:

- A fű zöld. (Milyen a fű? Zöld.)
- A bálna gerinces. (Milyen fajta (állat) a bálna? Gerinces.)
- A kutya ugat. (Mit csinál a kutya? Ugat.)

Ugyanakkor van egy szabad szöveges mező is, ahol a sablonoknak nem megfelelő mondatokat lehet felvinni. Ennek előnye, hogy itt a helytelenül beírt szavakra piros aláhúzással figyelmeztet a weboldal. A másik előny, hogy ha esetleg mégis belefér valamelyik sablonba az általunk beírt mondat, akkor erre figyelmeztet a rendszer és felajánlja, hogy a sablonhoz tartozó mondatként kerüljön be az adatbázisba az új mondat.

A közös tudás adatbázis töltése közben figyeltük fel arra, hogy mindenkinek sokkal könnyebb úgy felvinnie a mondatokat, hogy azoknak van valami köze egymáshoz, például egy fogalomról alkotott közös tudáshoz tartoznak. Egy szóhoz tartozó eddig bevitt mondatokat a rendszer képes listázni, ezzel is segítve a feltöltő munkáját, ha épp nem jut már eszébe semmi az adott fogalomról. Például a sálról a következő mondatok találhatóak jelenleg az adatbázisban:

- A sál egy ruhadarab.
- A sál meleg.
- A sál egy téli ruhadarab.
- A sál hosszú.
- A sál puha.
- A sál védi a nyakat a hidegtől.
- A sálat a nyakunk köré tekerve hordjuk.
- Sálat télen hordunk.
- A női sálak színesek.
- A sál készülhet pamutból.
- A sál készülhet selyemből.

4. Az oldal technikai megvalósítása

A Common Sense weboldal elkészítésekor figyelembe kellett venni, hogy rengeteg adatot kell eltárolni és megjeleníteni, és nem árt, ha mindez gyorsan történik. Egy MySQL adatbázis lett ebben segítségemre. A lehető legkisebb számú SQL lekérdezés fut egy-egy lap letöltésnél a gyorsaság érdekében. Az oldal motorja PHP nyelven íródott, mivel gyorsan és megbízhatóan tud kommunikálni az adatbázissal. Ügyelve a redundáns kód elkerülésére, rengeteg függvényt tartalmaz a forráskód, amelyeket így, elég csak meghívni egy szimpla `'call_user_func()'`-al. Erre azért is van szükség, mert a mondatok megjelenítése mellett, választható opció a regisztrálás is, melynek köszönhetően a felhasználókat jobban számon tudjuk tartani és kiszűrni a helytelen ismereteket. Keretrendszer ennél a résznél nem használtam.

A weboldal több különálló fájlból áll, amelyek kommunikálnak egymással. Például: Amikor új kifejezést viszünk fel, azt az Upload.php végzi. De amikor elkezdünk gépelni, automatikusan meghívja a CheckForSentence.php fájlt, amely viszont azt ellenőrzi, hogy milyen hasonló mondatok lettek már eltárolva. Ehhez egy másik fontos funkciót is implementálni kellett. Az AJAX segítségével bizonyos részei az oldalnak interaktívak, azaz a lap újratöltése nélkül is új információkat kapunk. Az előbbi példához hasonlóan, amikor új mondatot szeretnénk felvinni, de nem találjuk az alkalmas mondatszerkezetet, mint például „Minden” vagy „A(z) tud”, akkor lehetőség adódik, hogy csak egyszerűen egy textbox-ra beleírjuk, és ha az oldal egyezést talál egy már létező mondatszerkezet és az általunk megadott mondat között, kiírja azt, az oldal újratöltése nélkül.

Hasonló hiba akadt a mondatok között való keresésnél is. A hiba: ha rákeresünk arra a szóra, hogy 'zöld', azokat a mondatokat is adja ki, amelyek tartalmazzák a 'zöldebb' vagy 'zöldülnek' szavakat is. Ennek kiküszöbölésére, a weboldal megvizsgálja a szavak szótövét, és az alapján dönti el, hogy jó-e vagy sem. Az adatbázisban tárolt rengeteg szó lehetőséget adott egy Rímkeresési funkcióra is, ahol a kereső, a megadott szó rímshavait listázza ki.

Másik fontos probléma az oldal megjelenítése volt. Az oldal szerkezete CSS stíluslapokkal van tagolva, ügyelve a böngészőfüggetlenítésre is. Rengeteg design elemet köszönhet az oldal a style.css fájlban, mint például a lekerekítések, árnyékolás hatása és minden egyes elem, amit az oldalon látunk, itt van megformázva. Az egyes megjelenített elemek az oldalon egyszerű HTML tag-ek.

A kliens oldali funkciók és animálások a JavaScriptnek, azon belül is a jQuery keretrendszernek köszönhetőek. Például: A regisztrált és belépett felhasználóknak lehetőségük van szavazni az egyes ismeretekre, továbbá meg is változtathatják azokat, ha elírás található benne. Ehhez egy ikonra kell kattintani, aminek hatására egy JavaScript kód eltünteti a mondatot, és helyette beilleszt egy textbox-ot, az ismerettel együtt. Ha készen vagyunk, folytathatjuk a böngészést.

Implementálásra került egy a jsPlumb nevű Javascript keretrendszer is, annak érdekében, hogy a szavak és mondatok közötti kapcsolatokat grafikusán is, még szemléltethetőbben kiemeljük. Pl.: Egy az 'ember' és a 'írni' szavak között a kapcsolat: A(z) tud Ez a keretrendszer a HTML5-el bejövő canvas-t használja alapul, melynek másik nagy előnye a grafikai megjelenítés mellett, hogy az okostelefonok is támogatják.

5. Az adatbázis feldolgozása

A common-sense adatbázisban tárolt mondatokat különböző algoritmusokkal végzett vizsgálatoknak vetjük alá. Ezen vizsgálatok célja egyrészt a mondatokból való információkinyerés lehetséges

eszközeivel való tapasztalatszerzés, másrészt a közben kinyert információk értelmezése, feldolgozása (pl. publikációk formájában).

A vizsgálatokhoz használt első eszközünk Kohonen Self Organized Map [Kohonen1990] algoritmuson alapszik. Ez a típusú neurális hálózat gyakran használatos adatbányászathoz. Igen hasznos olyan esetekben, amikor egy adathalmaz megismerése a cél, az adathalmaz elemeinek csoportosíthatóságára, osztályozhatóságára vagyunk kíváncsiak.

Az algoritmus alkalmazásához az adatokat - esetünkben a mondatokat - többdimenziós térben kezeljük. Ebben a többdimenziós térben mérünk távolságokat az adatok (mondatok) között.

A koordináták megválasztása lényeges az eredményességhez. A mondatok elemzéséhez többféle koordináta-rendszer használata is lehetséges.

Az egyik legegyszerűbb ezek közül valahány jellemző szó (pl. a leggyakoribb szavak) kiválasztásával történik. A mondatokat ezen szavakhoz mérjük aszerint, hogy az adott szó szerepel-e a mondatban.

Finomabb használata szinonima-szótár alkalmazásával lehetséges. Ebben az esetben a távolság nem csak 0 vagy 1 lehet, hanem attól függ, hogy hány szinonima keresztül található kapcsolat a mondat és az adott szó között.

Másik lehetőség, hogy fogalmakat nevezünk meg, amelyekhez számított távolságát mérjük a mondatoknak. Ez már természetesen nehezebb feladat.

Újabb viszonylag egyszerű mérce az adatbázisból nyert nem nyelvi adatoknak a használata a méréshez. Ilyenek lehetnek pl. egy adott tudás elfogadottsága, a felvivő által felvitt mondatok száma, a felvitel mely napszakban történt, a felvivő lakhelye, stb.

A vizsgálatok következő eszköze a leggyakrabban használatos mesterséges neurális háló algoritmus, a backpropagation hálózat. Ez a fajta neurális hálózat képes megtanulni valamilyen típusú bemeneti adatokra megfelelő válaszok adását, amennyiben egy tanuló adatbázison a tanítást el tudjuk végezni.

A common-sense adatbázison való kísérleti alkalmazásának célja annak vizsgálata, hogy a mondatok fontossága, meglepőse előrejelezhető-e.

Ehhez a mondatok egy részéhez szükséges az elvárt válasz megadása, azaz az adatbázisban rögzítjük az adott mondat fontosságát, meglepőségét.

A Kohonen hálózathoz hasonlóan itt is szükséges a mondatokból input adatok előállítását, amellyel most a backpropagation neurális hálózatot tápláljuk. Ez többé-kevésbé az előzőekben említett módokon történhet.

6. Irodalomjegyzék

[Alpha] <http://www.wolframalpha.com/>

[DBpedia] <http://www.dbpedia.org/>

[DBpedia2007] Sören Auer, Jens Lehmann: What have Innsbruck and Leipzig in common? Extracting Semantics from Wiki Content. In Proceedings of European Semantic Web Conference (ESWC'07), LNCS 4519, pp. 503–517, Springer, 2007.

[Cyc] <http://www.opencyc.org/>

[Cyc1995] Douglas Lenat, CYC: A Large-Scale Investment in Knowledge Infrastructure, Communications of the ACM, Volume 38, pages 33—38, 1995.

[Kohonen1990] T. Kohonen: The self-organizing map, Proceedings of the IEEE, Volume: 78 Issue: 9, pages 1464 – 1480, 1990.

[Minsky1988] Marvin Minsky: The Society of Mind, ISBN 0671657135, published by Simon & Schuster, 1988.

[Minsky2006] Marvin Minsky: The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind, ISBN 0743276647, published by Simon & Schuster, 2006.

[OpenMind] <http://openmind.media.mit.edu/>

[OpenMind2002] Push Singh, Thomas Lin, Erik T. Mueller, Grace Lim, Travell Perkins and Wan Li Zhu: Open Mind Common Sense: Knowledge Acquisition from the General Public, Lecture Notes in Computer Science, 2002, Volume 2519, On the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE, pages 1223-1237, 2002.

[YAGO2] <http://www.mpi-inf.mpg.de/yago-naga/yago/>

[YAGO2007] Fabian M. Suchanek, Gjergji Kasneci and Gerhard Weikum: Yago: A Core of Semantic Knowledge, 16th international World Wide Web conference (WWW 2007), 2007.