

# The case of BabelNet 2.0

Roberto Navigli

DIPARTIMENTO  
DI INFORMATICA



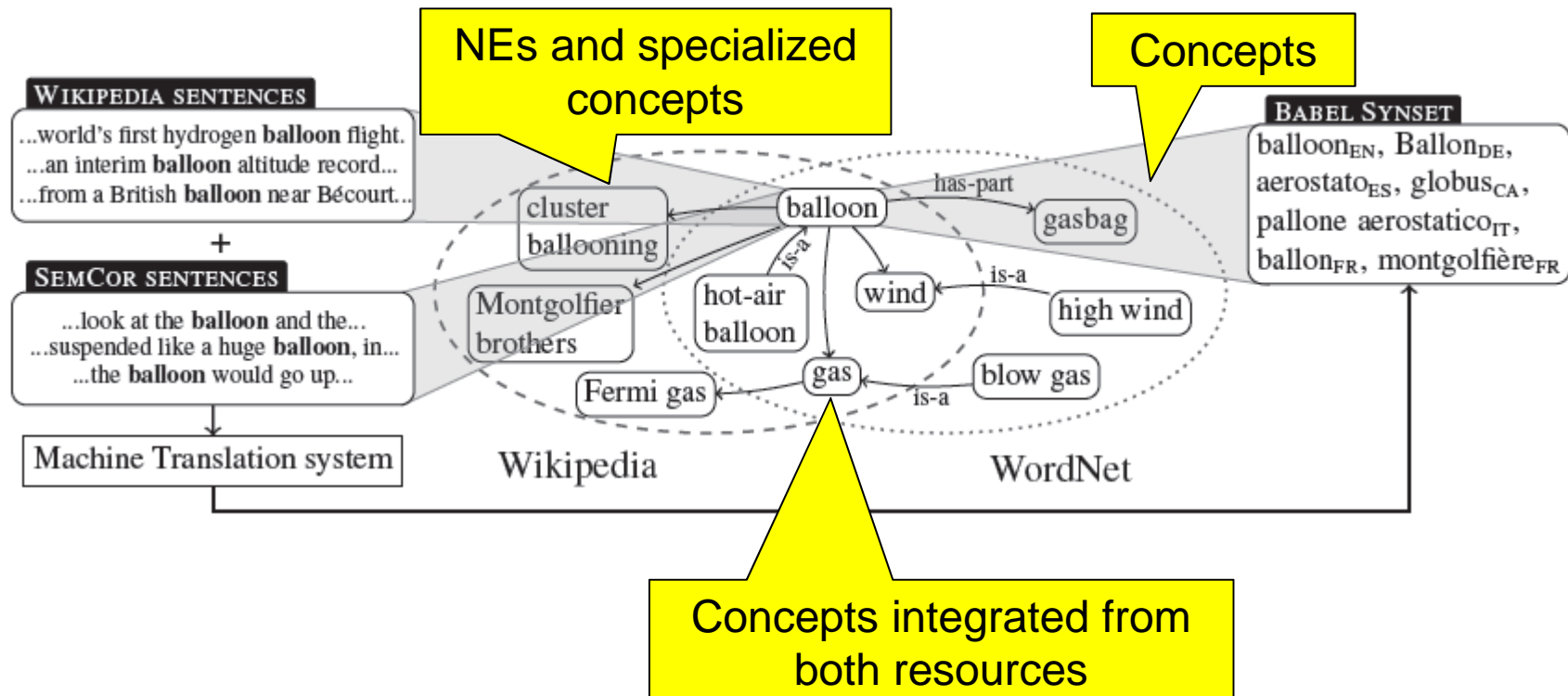
SAPIENZA  
UNIVERSITÀ DI ROMA

Linguistic Computing Laboratory  
<http://lcl.uniroma1.it>

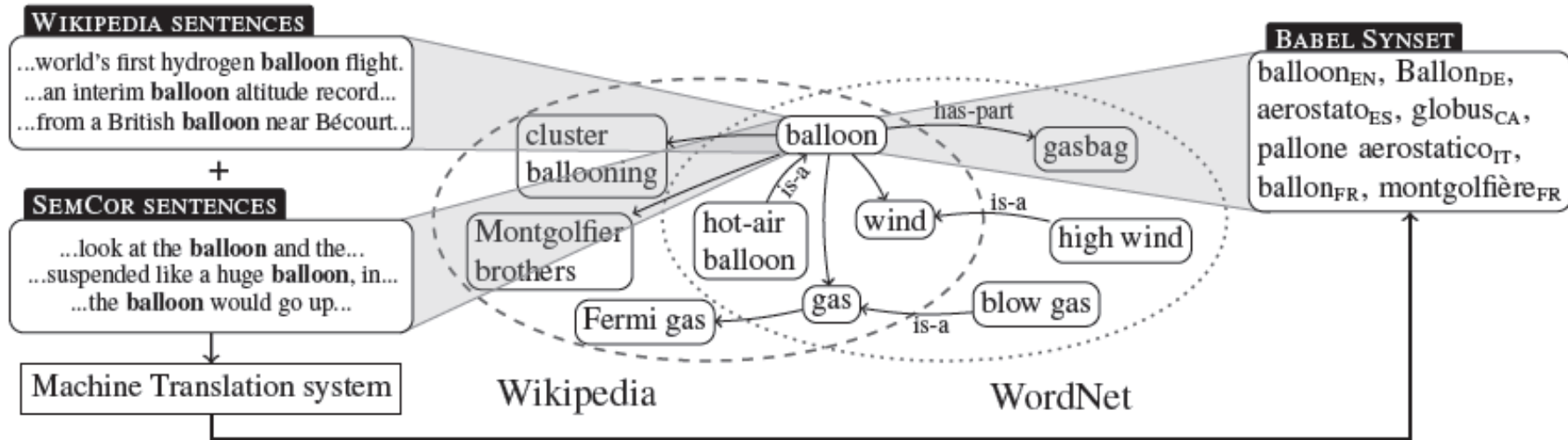
Athens, Greece, March 21, 2013

# BabelNet (<http://babelnet.org>)

- A wide-coverage multilingual semantic network including both **encyclopedic** (from Wikipedia) and **lexicographic** (from WordNet and OmegaWiki) entries



# BabelNet: a multilingual encyclopedic dictionary!



- Available online: <http://babelnet.org>

# BabelNet 2.0 is online: <http://babelnet.org>

 about •  statistics •  preferences

## BabelNet<sup>2.0</sup>

A very large multilingual **encyclopedic dictionary** and **ontology**

**BabelNet** is both a **multilingual encyclopedic dictionary**, with lexicographic and encyclopedic coverage of terms in **50 languages**, and an **ontology** which connects concepts and named entities in a **very large network of semantic relations**, made up of **more than 9 millions of entries**. [Read more...](#)



## Anatomy of BabelNet 2.0

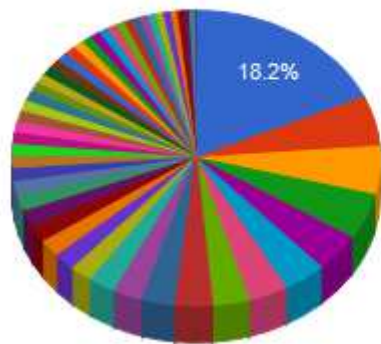
- **50** languages covered (including Latin!)
- Full list at <http://babelnet.org/stats.jsp>



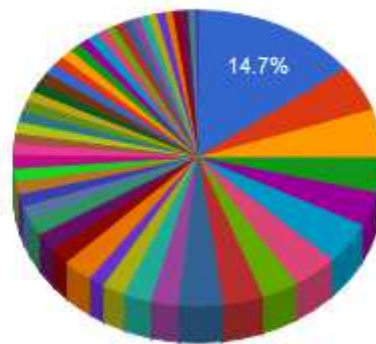
## Anatomy of BabelNet 2.0

- **50** languages covered (including Latin!)
- **9.3M** Babel synsets (concepts and named entities)
- **50M** word senses
- **262M** semantic relations (28 relations per synset on avg.)
- **7.7M** synset-associated images
- **18M** textual definitions

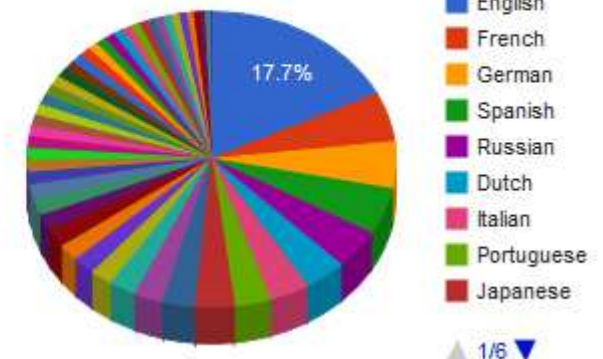
Lemmas by Language



Synsets by Language



Senses by Language






## Anatomy of BabelNet 2.0










































- **50 languages covered** (including Latin!)
- Integrates:
  - WordNet 3.0
  - Wikipedia (October 2012 dumps)
  - **OmegaWiki**: a collaborative multilingual dictionary
  - **Open Multilingual WordNet** [Bond and Foster, 2013]
- Translations for **all open-class parts of speech**















































# WordNet+Open Multilingual WordNet+Wikipedia+...













































































Meaning: airplane<sup>1</sup> • ID: bn:00001697n • Type: Concept [\[explore\]](#)

**Senses:**

  airplane<sup>1</sup>,  aeroplane<sup>1</sup>,  plane<sup>1</sup>

  aereo,  aeroplano,  apparecchio,  areoplano,  aeroplano,  avión,  aeroplà,  avió,  samolot,  aeroplan,  エアプレーン,  エアプレーン,  機,  航空機,  銀翼,  飛行機,  鵬翼,  אָוִיּוֹן,  אוֹיּוֹן,  طياره,  fly,  flygar,  flyver,  lentokone,  kapal terbang,  pesawat,  pesawat terbang,  aeroplan,  fly,  flyvemaskine,  flyver,  kapal terbang,  pesawat,  pesawat terbang,  avión,  aeroplan,  aeroplano,  avión,  hegazkin,  aireplano,  aireko,  abioi

  Fixed-wing aircraft,  Aeroplano,  Avión,  Flugzeug,  Avion,  Avió,  Samolot,  Avião,  Vliegtuig,  Самолёт,  Avion,  Uçak,  飛行機,  固定翼飛機,  Máy bay,  Самолет,  طائرة,  비행기,  οιοση,  هواينة,  वायुयान,  Літак,  Fly,  Lentokone,  Letoun,  Repülőgép,  비행기,  Pesawat bersayap tetap,  Flyvemaskine,  Авион,  Lietadlo,  Lėktuvas,  Pesawat kepak kaku,  Avion,  Αεροπλάνο,  Vliegtuig,  Letalo,  Aeroplani,  Lennuk,  Eitleán,  Lidmašina,  Awyren,  Flugvél,  Ndege (uanahewa),  Eroplano,  Hegazkin,  Aviadilo,  Avión (medio de transporte),  Aëroplanum

  Air-plane,  Fixed wing,  Airplane,  Airoplane,  Aëroplane,  ✈,  Aeroplane,  Fixed-winged aircraft,  Aerplane,  Fixed-wing,  Airplanes,  Aeroplanes,  Plane (aircraft),  Aereoplano,  Aeroplani,  Aereo,  Aeroplano,  Avión de alas fijas,  Aviones,  Aeronave de alas fijas,  Aeronave de ala fija,  Avion de alas fijas,  ✈,  Viajes aereos,  Viajes aéreos,  Aëroplan,  Flugmaschine,  Verkehrsflugzeug,  Starrflügel,  Avions,  Fonctionnement d'un avion,  Aeronau d'ala fixa,  Aeroplà,  Avions,  Aparat latający,  Aeroplan,  Samoloty,  Cechy konstrukcyjne samolotów,  Aviões,  Aeronave de asa fixa,  Aeroplano,  Avião a Jato,  Jatos,  Jato,  Avioneta,  Indústria aeronáutica,  Aviao,  Vliegtuigen,  Vliegmaschine,  Аэроплан,  Самолёты,  Самолет,  Самолеты,  ✈,  Avioane,  Teyyare,  Sabit kanatlı uçak,  Uçak kalkış-iniş,  ✈,  Tayyare,  大型飛行機,  大型旅客機,  蒸気飛行機,  電気動力式飛行機,  固定机翼飞机,  飞机,  固定翼航空器,  固定翼機,  固定翼,  固定翼机,  飛機,  固定翼飞机,  定翼機,  Tàu bay,  Máy bay cánh cố định,  Phi cơ,  Самолети,  Аэроплан,  ✈,  Aircraft,  الطائرة,  الطائرة





























## +OmegaWiki+automatic translations...

  aeroplane, plane, fixed-wing aircraft, bird,  aereo, aeroplano,  avión, aeroplano,  Flugzeug,  avion, aéroplane,  samolot,  aeroplano, avião,  toestel, vliegtuig, vliegmaschine, kist,  аэроплан, самолёт,  avion,  uçak,  飛行機,  飛機, 飛行機, 飞龙机, 飞行机, 飞机, 飛龍機,  máy bay,  аероплан, самолет,  طائرة,  flygplan,  repülőgép,  הַמְּטוֹרָן,  هوايما,  वायु-यान, हवाई जहाज़,  аероплан, літак,  fly,  lentokone,  lietadlo,  repülőgép,  비행기,  fly, flyvemaskine,  авион,  lietadlo,  lėktuvas, aeroplanas,  kapal terbang,  avion, zrakoplov,  αεροπλάνο,  vliegtuig,  letalo,  avion,  aeroplan,  lennuk, aeroplaan,  eitleán,  awyren,  flugvél,  eroplano,  hegazkin,  aviadilo, flugmašino, avio,  aeroplanum

  velivoli ad ala fissa, ala fissa, aereo, aerei,  aeronaves de ala fija, ala fija, avión, aviones,  starrflügler, flugzeug, flugzeuge, flugzeugen,  aéronefs à voilure fixe, voilure fixe, avion, avions,  aeronaus d'ala fixa, ala fixa, avió, avions,  stałej skrzydło, samolotów,  aeronaves de asa fixa, de asa fixa, asa fixa, avião, aviões,  vliegtuigen met vaste vleugels, vliegtuig, vaste vleugels, vliegtuigen,  самолетов, самолеты,  avion, cu aripă fixă, avioane,  sabit kanatlı uçak, sabit kanatlı, uçak,  固定翼, 飛行機,  固定翼飞机, 固定翼, 飞机,  máy bay cánh cố định, cánh cố định, bay, máy bay,  самолет, с неподвижно крило, неподвижно крило, самолети,  الطائرات,  flygplan med fasta vingar, fasta vingar, flygplan,  מִטוֹסֵי כָּנָה קְבועָה, מִטוֹסֵי כָּנָה קְבועָה, מִטוֹסֵי כָּנָה קְבועָה,  מִטוֹסֵי כָּנָה קְבועָה,  हायिमा,  بال ثابت,  फिक्स्ड विंग विमान, तय विंग, हवाई जहाज, फिक्स्ड विंग,  літаків, літаки,  fast ving fly, faste vinger, fly, fast ving,  lentokone,  fixovaný-letadla křídla, pevné křídlo, letadel, letadla,  merevszárnyú repülőgépek, merevszárnyú, repülőgép, repülőgépek,  고정 날개,  pesawat sayap tetap, sayap tetap, pesawat, pesawat terbang,  fastvingede fly, fastvingede, fly,  фиксне крила авиона, фиксне крила, авиона,  fixovaný-lietadla křídla, lietadiel, lietadlo,  fiksuo tujy sparny orlaivis, fiksuo tujy sparno, fiksuo tujy sparny, léktuvo, léktuvai,  pesawat sayap tetap, sayap tetap, kapal terbang,  nepokretnim krilima zrakoplova, zrakoplova,  αεροσκάφη σταθερής πτέρυγας, αεροπλάνο, σταθερής πτέρυγας, αεροπλάνα,  vaste- vlerk vliegtuie, vastevlerk, vliegtuig, vaste- vlerk, vliegtuie,  nepremičnimi krili letala, fiksno krilo, letala,  aeroplan, aeroplanë,  fikseeritud hävituslennuk, jäigatiivaliste, lennuk, lennukid,  eitleán, sciatháin dhobhogtha,

## +textual definitions

<b>Glosses:</b>	  A powered heavier-than-air aircraft with fixed wings that obtains lift by the Bernoulli effect and is used for transportation.
	  an aircraft that has a fixed wing and is powered by propellers or jets; "the flight was delayed due to trouble with the airplane"
	  A fixed wing aircraft is an aircraft capable of flight using wings that generate lift due to the vehicle's forward airspeed and the shape of the wings.
	  Veicolo dotato di ali che grazie all'effetto Bernoulli riesce, sollevandosi da terra, a rimanere in aria e viene utilizzato per il trasporto.
	  L'aeroplano, o aereo, è un mezzo di trasporto.
	  Artefacto más pesado que el aire con alas fijas que puede levantarse del suelo debido al efecto de Bernoulli y es utilizado para transporte.
	  Avión, también denominado aeroplano, es un aerodino de ala fija, o aeronave con mayor densidad que el aire, provisto de alas y un espacio de carga capaz de volar, impulsado por uno o más motores.
	  Ein Flugzeug ist ein Luftfahrzeug, das schwerer als Luft ist und den zum Flug nötigen dynamischen Auftrieb mit starren Tragflächen erzeugt.
	  Un engin aérien motorisé plus lourd que l'air pourvu d'ailes fixes lui assurant une portance en vertu de l'effet de Bernoulli, qui est utilisé à des fins de transport.
	  Un avion, selon la définition officielle de l'Organisation de l'aviation civile internationale, est un aéronef plus lourd que l'air, entraîné par un organe moteur, dont la sustentation en vol est obtenue principalement par des réactions aérodynamiques sur des surfaces qui restent fixes dans des conditions données de vol.
	  Un avió, és una aeronau, més pesant que l'aire, capaç de volar per l'atmosfera.
	  Um avião ou aeroplano é qualquer aeronave que necessita de asas fixas para se sustentar no ar.
	  Een aangedeven vliegmaschine die zwaarder is dan lucht met vaste vleugels die lift krijgt door het Bernoulli-effect en gebruikt wordt voor transport.

## +Wikipedia categories

**Categories:**   Aeronautics, Aircraft configurations,   Aeroplani,  Aeronáutica, Configuraciones de aeronaves, Inventos,  Luftfahrzeug,  Type d'aéronef,  Avions,  Samoloty,  Vliegtuig,  Uçaklar, Fransız icatları,  航空, 航空器, 航空器种类, 飛機, 大氣污染,  Máy bay, Hàng không, Khí cụ bay, Vận tải, Phương tiện giao thông, Profile khí động lực học, Lực nâng máy bay,  Самолети, Обществен транспорт,  الملاحه الجوية, تشكيلات الطائرات, طائرات, كلمات من وضع مجمع اللغة العربية,  اختراعات أمريكايي, هوايما, فناوری هوافضا, فناوری هوایی, هوايماي, תעופה מסוים,  הונורדי,  विमानन, वायु परिवहन के साधन, वायुयान, यान,  Літаки, Повітряні судна,  Fly, Kollektivtrafikk,  Lentokoneet, 1900-luvun keksinnöt,  Letadla,  Repülőgépek, Repüléstechnika,  항공기,  Aeronautika, Pesawat, Konfigurasi Pesawat,  Flyvemaskiner, Luftfart, Maskiner,  Авиони, Ваздухоплови,  Lietadlá,  Orlaiviai,  Zrakoplovi,  Vliegtuie,  Zrakoplovi, Letala,  Technologie, Aeronautikë, Komunikacion,  Lennundus, Lennuliiklus,  Aerárthach,  Awyrennau,  Flugvélar,  Chomboanga,  Aeronautika,  Aviadiloj, Transportaj rimedoj, Aviado,  Aviación,  Aerinavigatio, Res volantes





## RDF-Lemon encoding of BabelNet

- RDF representation based on:
  - *lemon*, a Lexicon Model for Ontology
  - SKOS, Simple Knowledge Organization System
  - LexInfo Ontology 2.0
- 1,026,780,575 triples
- Interlinking: DBpedia, Wikipedia, *lemon* WordNet and *lemon* OmegaWiki (English), Wikipedia

# License issues and current solution

CC BY 3.0	Permette la creazione di opere derivate, con licenza differente da quella originale (purché compatibile), ma bisogna fare riferimento all'autore dei dati	<a href="http://creativecommons.org/licenses/by/3.0/">http://creativecommons.org/licenses/by/3.0/</a>
CC BY SA 3	Permette la creazione di opere derivate, ma con la stessa licenza dei dati utilizzati, consente scopi commerciali	<a href="http://creativecommons.org/licenses/by-sa/3.0/">http://creativecommons.org/licenses/by-sa/3.0/</a>
CC BY-NC-SA 3.0	<i>Niente scopi commerciali e stessa licenza per opere derivate e riferimento all'autore (direi di scartarla direttamente)</i>	<a href="http://creativecommons.org/licenses/by-nc-sa/3.0/">http://creativecommons.org/licenses/by-nc-sa/3.0</a>
	Nessuna restrizione sui dati	

Open Multilingual WordNet		
	als	GPL3, CC BY 3.0
	arb	CC BY-SA 3.0 <a href="http://creativecommons.org/licenses/by-sa/3.0/">http://creativecommons.org/licenses/by-sa/3.0/</a>
<b>Wikipedia</b>	cmn	WordNet
CC BY-SA 3.0	dan	WordNet
	eng	WordNet
	fas	No restriction
	fin	CC by 3.0 <a href="http://creativecommons.org/licenses/by/3.0/">http://creativecommons.org/licenses/by/3.0/</a>
<b>Wiktionary</b>	fra	CeCILL-C <a href="http://www.casta-net.jp/~kuribayashi/multi/wns/fra/LICENSE">http://www.casta-net.jp/~kuribayashi/multi/wns/fra/LICENSE</a>
CC BY-SA 3.0	heb	WordNet
	ita	CC by 3.0
	jpn	WordNet
	cat	CC BY 3.0
<b>WordNet</b>	eus	CC BY-NC-SA 3.0
WordNet licence	glg	CC BY 3.0
	spa	CC BY 3.0
	ind	MIT
	zsm	MIT
	nno	WordNet
	nob	WordNet
	pol	WordNet
	por	CC BY SA 3.0 <a href="http://creativecommons.org/licenses/by-sa/3.0/br/">http://creativecommons.org/licenses/by-sa/3.0/br/</a>
	tha	WordNet

## License issues and current solution

CC BY 3.0	Permette la creazione di opere derivate, con licenza differente da quella originale (purché compatibile), ma bisogna fare riferimento all'autore dei dati	<a href="http://creativecommons.org/licenses/by/3.0/">http://creativecommons.org/licenses/by/3.0/</a>
CC BY SA 3	Permette la creazione di opere derivate, ma con la stessa licenza dei dati utilizzati, consente scopi commerciali	<a href="http://creativecommons.org/licenses/by-sa/3.0/">http://creativecommons.org/licenses/by-sa/3.0/</a>
CC BY-NC-SA 3.0	<i>Niente scopi commerciali e stessa licenza per opere derivate e riferimento all'autore (darei di scartarla direttamente)</i>	<a href="http://creativecommons.org/licenses/by-nc-sa/3.0/">http://creativecommons.org/licenses/by-nc-sa/3.0</a>
	Nessuna restrizione sui dati	

- Create an onion-like distribution with:
  1. BabelNet core, including unrestricted data and the alignments between resources
  2. CC-BY 3.0 data
  3. CC-BY-SA 3.0 data
  4. CC-BY-SA-NC 3.0 data (only if BN is used for non commercial purposes)



# Babelfy: Bringing together Entity Linking and Word Sense Disambiguation (TACL, 2014)

- «Thomas and Mario are strikers playing in Munich»

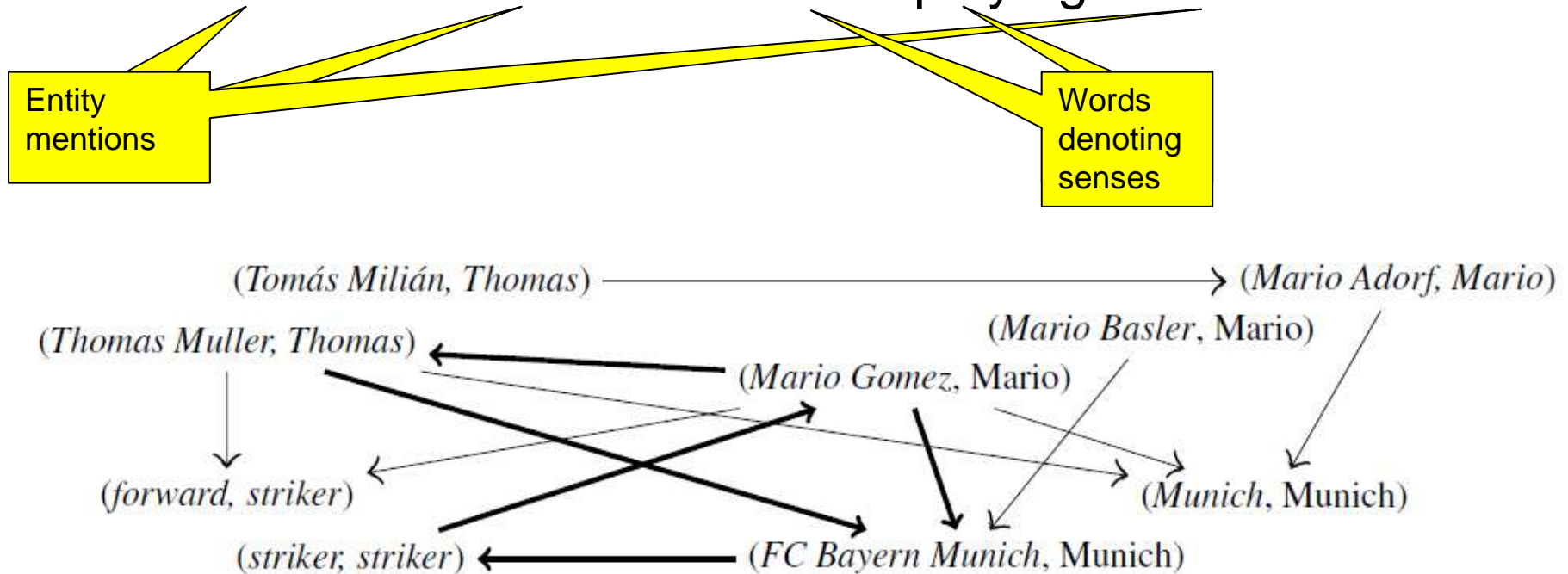


Figure 1: An excerpt of the semantic interpretation graph automatically built for the sentence *Thomas and Mario are strikers playing in Munich* (the edges connecting the correct meanings are in bold).



SAPIENZA  
UNIVERSITÀ DI ROMA

**Roberto Navigli**

Linguistic Computing Laboratory

<http://lcl.uniroma1.it>

