



Análisis Textual de Argumentos en Escritos Académicos (Creación de corpus)

Proyecto Doctoral

18 de Enero del 2017



National Institute of Astrophysics, Optics and Electronics
Department of Computational Sciences
Laboratory of Language Technologies



Contenido

1. Introducción
2. Objetivo y modelo
3. Construcción del corpus
4. Sistema de anotación BRAT
5. Conclusión

Introducción I

- La escritura de textos de investigación implica presentar **las ideas de estudiantes** en documentos como:
 - Tesis
 - Proyectos de final de curso
 - Reportes de Investigación
- Una buena argumentación es necesaria para **transmitir y sustentar** claramente dichas ideas.
- Existe una **deficiencia** en la habilidad de argumentar entre algunos estudiantes (Bañalez et al, 2015).

Introducción II

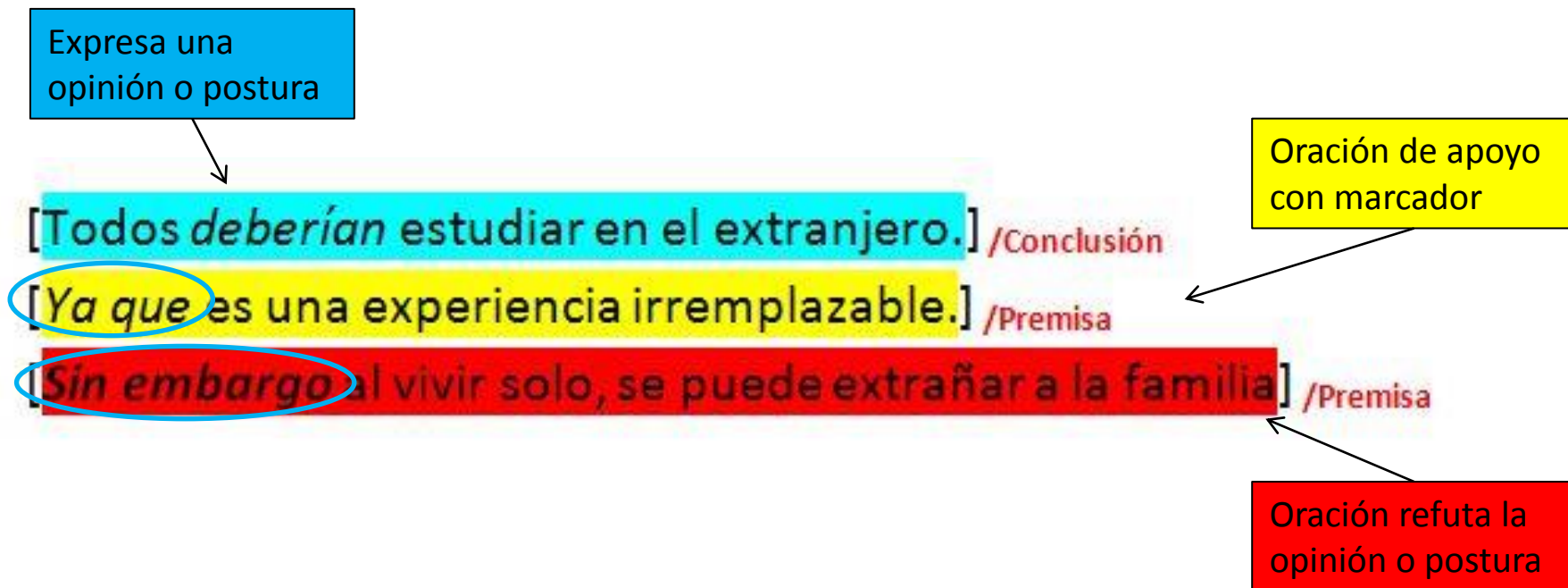
- Argumentar significa **fundamentar con evidencias** lo que se pretende probar (Argudin, 2005).
- Existen diversas razones para analizar la argumentación
 - Extracción de evidencias (Mochales & Moens, 2011).
 - Identificación de posturas (a favor o en contra) (Cabrio & Villata, 2012).
 - Evaluación de ensayos (Persing & Ng, 2015).

Argumentación I

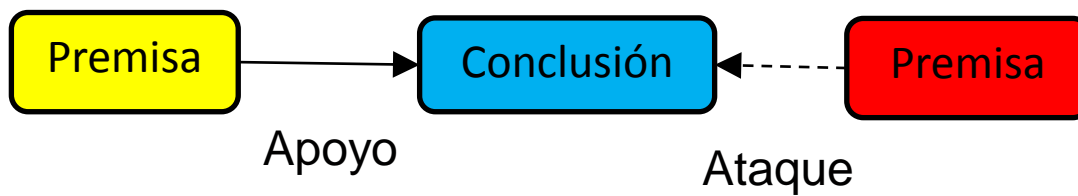
- Un argumento puede definirse como una serie de aseveraciones (oraciones, proposiciones) que individualmente o en su conjunto apoyan a otra afirmación (Capaldi, 2000).
 - La aseveración que resulta **apoyada** se denomina **conclusión** (postura, opinión).
 - La aseveración que **apoya o ataca** se denomina **premisa** (evidencia) (Freeman, 2011).

Argumentación II

Ejemplo



Estructura:



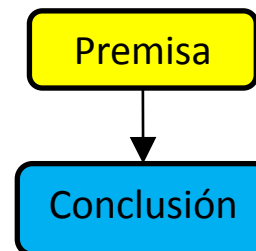
Ejemplo de escrito académico

Sección Justificación

Tarea 1 Identificar componentes:

[En la actualidad existen un gran número de **empresas de servicios**, públicas y privadas, que proveen insumos y/o **servicios** a otras **empresas** que exportan sus productos a **mercados** donde la **calidad** es fundamental,]/Premisa [por ello éstas últimas están exigiendo a las **empresas de servicios** la implantación de un Sistema Administrativo de **Calidad** bajo la normatividad de ISO-9000, y su certificación por un organismo externo ya sea nacional o internacional, dependiendo del **mercado** o giro de la **empresa**, que sea reconocido por la Organización Internacional para la Estandarización, quien es la encargada de emitir y revisar las normas ISO-9000.]/Conclusión

Tarea 2 Determinar relaciones entre componentes:



Tarea 3 Evaluar el argumento:

Nivel de argumentación: media (2)

Ejemplo de anotación

Sección Justificación

1 Componentes:

[La autenticidad en una WSN debe ser una propiedad inherente de la propia red,]/C1 [ya que como se ha establecido en la sección anterior, un ataque originado a raíz de la falta de ésta puede llegar a comprometer el funcionamiento de la red completa.]/P2 [Sin embargo, a diferencia de otros entornos, las WSN tienen como un marco de operación muy reducido en recursos,]/P3 [por lo que la selección de los mecanismos de autenticidad a utilizar debe ser acorde a dichas limitantes.]/P4

2 Relaciones:

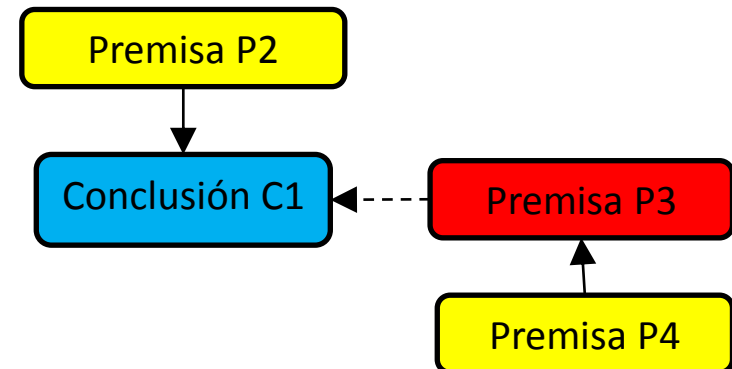
[Relaciones de apoyo (Quién apoya a quién)]

P2, C1

P4, P3

[Relaciones de ataque (Quién ataca a quién)]

P3, C1



3 Evaluar:

<nivel_argumentacion= 0 (no_tiene), 1 (débil), [2] (media), 3 (fuerte)>

<tipo_argumento= Autoridad, Ejemplos, [Causal], Analogía, [Comparación]>

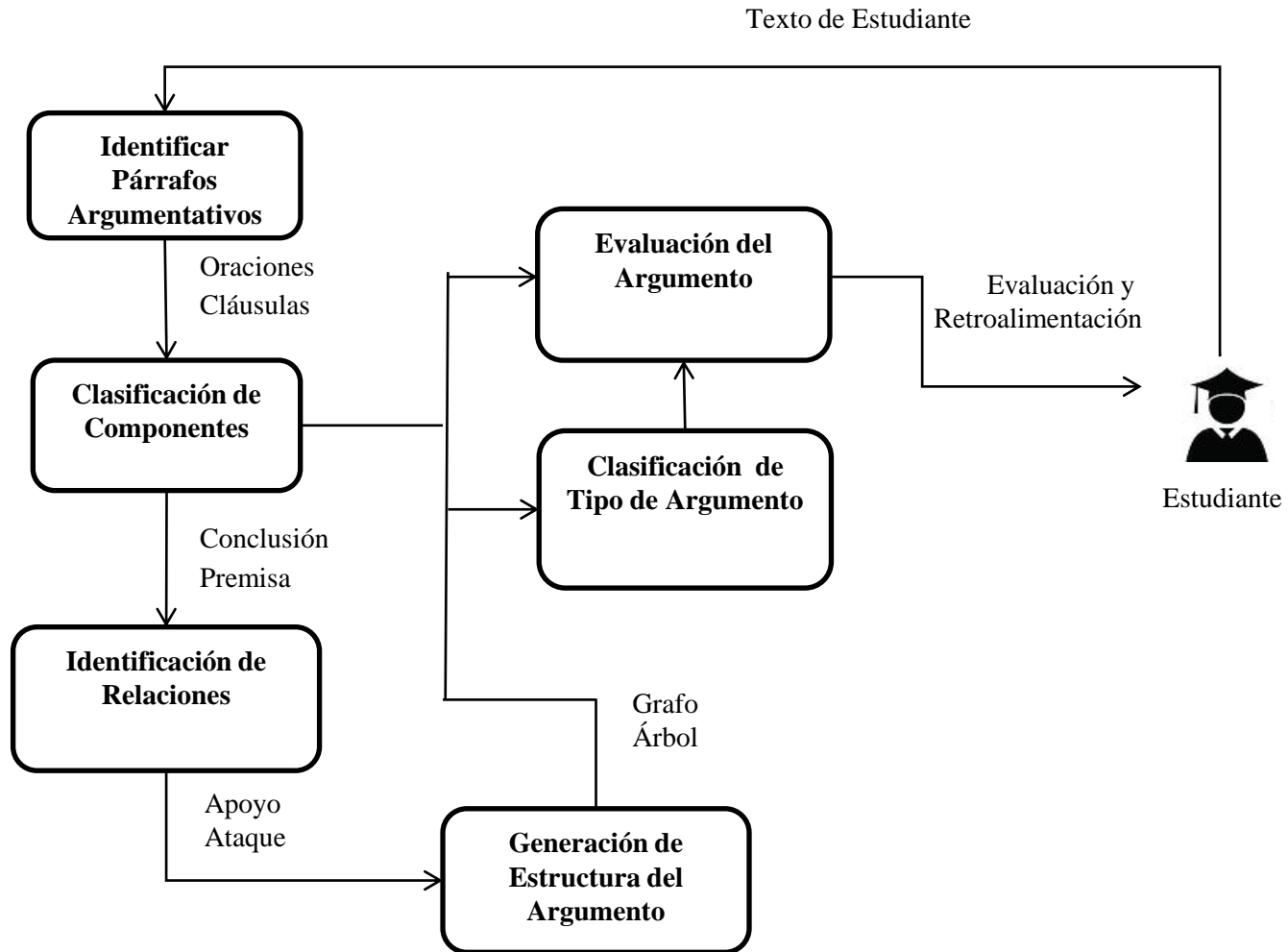
Contenido

1. Introducción
2. **Objetivo y modelo**
3. Construcción del corpus
4. Sistema de anotación BRAT
5. Conclusión

Objetivo general

- Desarrollar **métodos para extraer y evaluar** los argumentos en escritos académicos de investigación
 - aplicando técnicas de procesamiento del lenguaje natural para generar representaciones que permitan **modelar su información lingüística argumentativa**,
 - alcanzando niveles de acuerdo aceptables comparados con los **revisores humanos** (anotadores)
 - con la finalidad de **apoyar al alumno** en la escritura de su texto así como de asistir al profesor en su proceso de revisión.

Modelo



Contenido

1. Introducción
2. Objetivo y modelo
3. Construcción del corpus
4. Sistema de anotación BRAT
5. Conclusión

Identificación y selección de conjuntos de datos con escritos académicos de investigación

- **Colección Coltypi** (González-López & López-López, 2015)
 - 468 tesis y propuestas de investigación en español del área de computación de nivel:
 - pregrado (TSU y Licenciatura)
 - posgrado (Maestría y Doctorado)

Grado	Documentos
Doctorado	59
Maestría	181
Licenciatura	150
TSU	78
Total	468



Avance corpus I

- Se tienen anotadas 300 secciones por dos expertos.
- Mas de la mitad de los párrafos tienen argumentos.

	Párrafos con argumento	Párrafos sin argumento
Planteamiento del Problema	164	93
Justificación	151	81
Conclusión	250	98
Total	565 (68%)	272 (32%)

Avance corpus II

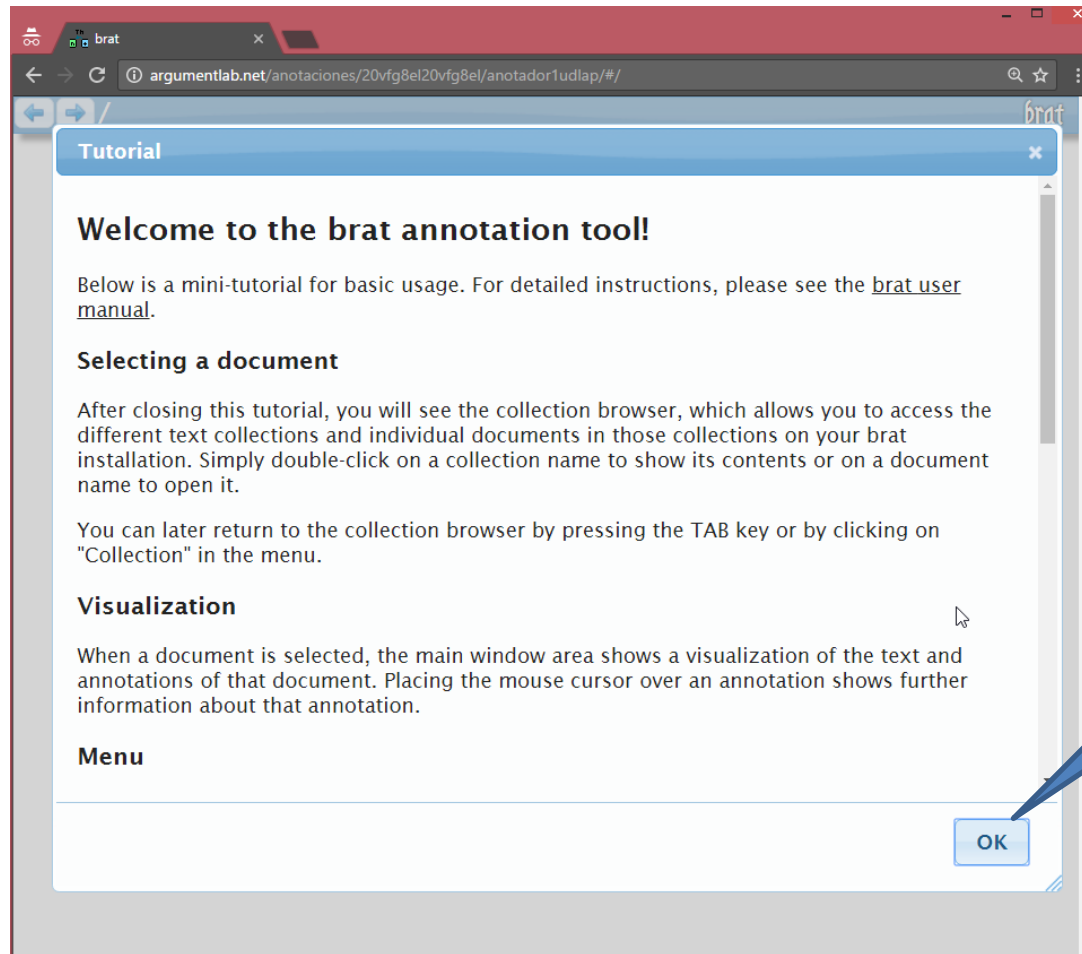
- Componentes anotados en las 300 secciones.
- Se observa una cantidad de premisas de casi el doble de conclusiones.

	Conclusiones	Premisa	Ninguna
Planteamiento del Problema	144	338	200
Justificación	135	252	141
Conclusión	243	423	268
Total	522 (24.4%)	1013 (47.2%)	609 (28.4%)

Contenido

1. Introducción
2. Objetivo y modelo
3. Construcción del corpus
4. Sistema de anotación BRAT
5. Conclusión

Sistema de anotación BRAT bienvenida

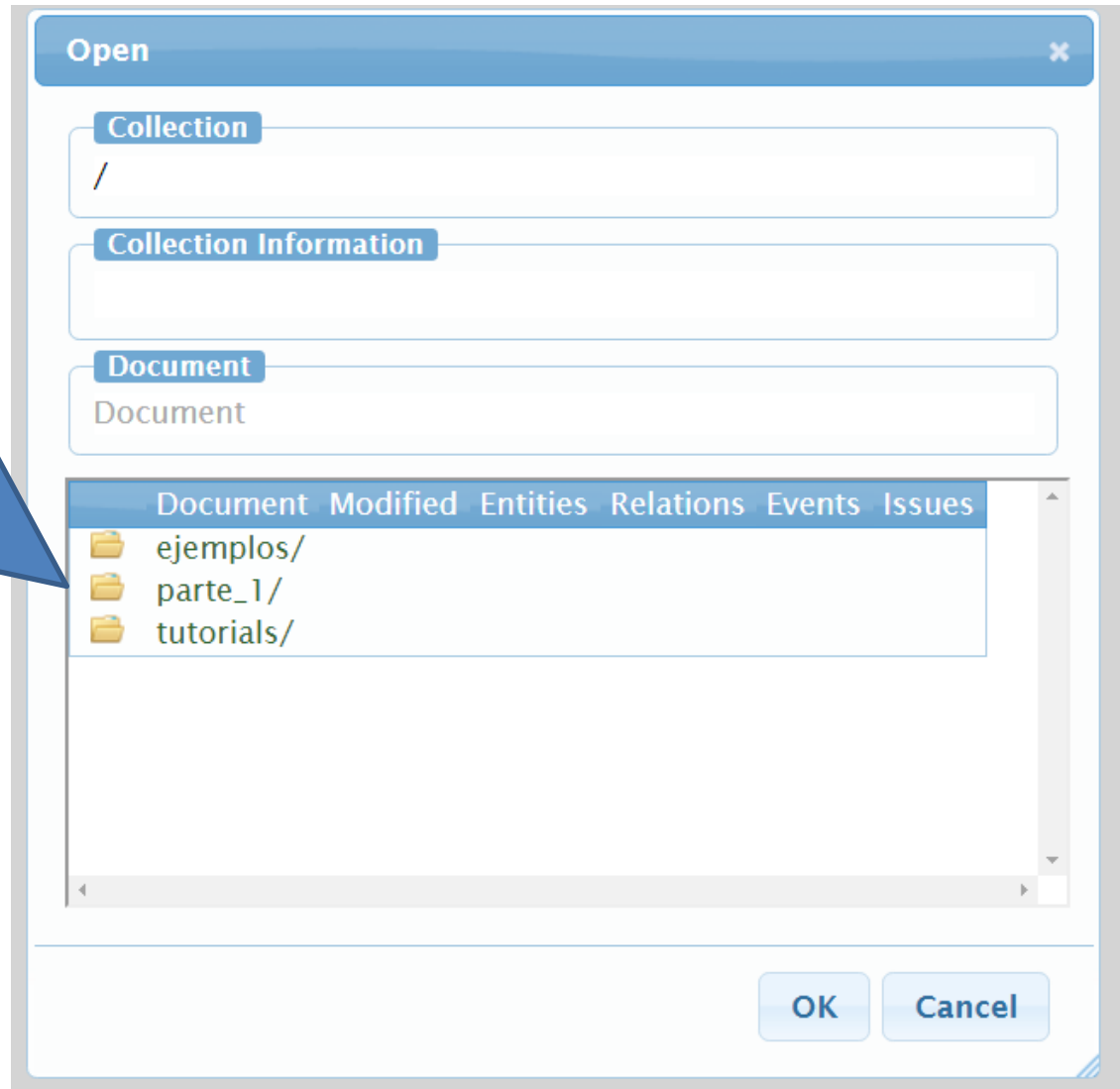


BRAT directorios

Ejemplos: se presentan tesis anotadas.

Tutorials : es la guía en ingles para usar BRAT

parte_#: Estos directorios se irán agregando gradualmente y contienen las tesis que serán anotadas.



BRAT archivo anotar

Login para hacer anotaciones

/ejemplos/ejemplo001

brat

Tesis de nivel: Maestria

Título: METODOLOGIA PARA LA IMPLANTACION DE UN SISTEMA ADMINISTRATIVO DE CALIDAD EN EMPRESAS DE SERVICIOS, BAJO LA NORMATIVIDAD DE ISO-9000

Objetivo: Proporcionar a los directivos de las empresas de servicios en una forma sencilla y práctica, la estructura de las normas ISO-9000, la esencia de los veinte requerimientos de la norma ISO-9001, así como sugerencias de cómo interpretarla y cómo afrontarla de manera efectiva.

@-----@

@ANOTAR_SECCION_Justificacion@

En la actualidad existen un gran número de empresas de servicios, públicas y privadas, que proveen insumos y/o servicios a otras empresas que exportan sus productos a mercados donde la competencia es fuerte. **Premisa** → **apoya** →
por ello éstas últimas están exigiendo a las empresas de servicios la implantación de un Sistema Administrativo de Calidad bajo la normatividad de ISO-9000, y su certificación por un organismo acreditado. **apoya** →

Nivel [Medio]

Tipo [Causal]

@CalificaPorNivelParrafoAnteriorJustificacion@ @CalificaPorTipoParrafoAnteriorJustificacion@

Por las empresas de servicios portuarios y servicios turísticos del estado de Colima, que cada día tiene un desarrollo tangible y un crecimiento sostenido, y que la expectativa a mediano y largo plazo es de un crecimiento sostenido. **Premisa** → **apoya** →
me siento motivado y quiero ayudar a contribuir con el presente trabajo a proporcionarles una metodología con fundamentos sólidos y comprobados, que les ayude y oriente objetivamente a implementar un Sistema Administrativo de Calidad. **Conclusion**

Nivel [Medio]

Tipo [Causal]

@CalificaPorNivelParrafoAnteriorJustificacion@ @CalificaPorTipoParrafoAnteriorJustificacion@

Antes de anotar

- Antes que nada, es necesario leer los todos los elementos a anotar para comprender el tema central del trabajo.
 - Leer el **título del trabajo y el objetivo** para ubicar el tema de estudio.
 - Después de haber leído en su totalidad los elementos a anotar, **releer un párrafo a la vez** y determinar si es o no argumentativo.
 - Es decir, si se identifica una conclusión y por la menos una premisa.
- NOTA: Para decidir si estamos ante un argumento o no, apelaremos al sentido común y a un sencillo análisis del texto sobre el que hayamos de decidir, centrándonos en los siguientes aspectos:
 - ¿El texto tiene una conclusión? Si es así, ¿cuál es?
 - ¿El texto ofrece pruebas que apoyen la conclusión?
 - ¿El texto presume que hay una relación inferencial entre premisas y conclusiones? (marcadores argumentativos)
- Revisar el material :
 - http://argumentlab.net/anotaciones/capaldi_como_id_argumentos.pdf
 - <http://proyectoafri.es/ffia1/cd23logicasimplex/aprenlogicamadrid/logica/actividades/1ref1.html>
 - <http://www.proyectoafri.es/ffia1/cd23logicasimplex/aprenlogicamadrid/logica/01concbasicos/121argum.html>

Identificar las conclusiones y premisas en los párrafos

- Procurar **anotar oraciones o clausulas** para cada componente argumentativo, evitar seleccionar varias oraciones en un solo componente argumentativo.
- Algunos marcadores discursivos para identificar las premisas son:
 - **ya que, puesto que, puesto, pues, como, en tanto que, dado que, viendo que, debido a, de acuerdo con, por cuanto, viendo que, a causa de, porque, se sigue de, como muestra, como es indicado por, la razón es que, se puede inferir de, se puede derivar de, se puede deducir de, en vista de que, debido a (que)...**
- Y para las conclusiones:
 - **por lo tanto, por ende, de ahí que, en consecuencia, por consiguiente, se desprende de, como resultado, llegamos a la conclusión, por (lo) tanto, lo cual apunta a la conclusión de que, lo cual muestra que, así (que), lo cual nos permite inferir que, correspondientemente, lo cual implica que, lo cual prueba que, lo cual significa que, se sigue que, por estas razones, por esta razón, podemos inferir que, concluyo que, consecuentemente...**
- También existen palabras para ubicar la relación de ataque (contraargumento) como son las siguientes:
 - **sin embargo, a pesar de, a menos (que), de todos modos, de cualquier modo, en cambio, en todo caso, no obstante, pero, por el contrario, mas, ahora bien, a pesar de todo, después de todo, antes bien, no parece, sino que, si bien...**

No es un argumento

- Lista de posibles alternativas cuando no encontramos en una porción de discurso premisas, conclusión o relación inferencial lógica entre ambas.
- **Advertencias:**
 - No se proporcionan razones (no hay premisas). P
 - redomina la función apelativa (llamar la atención) y conativa (conseguir algo).
- **Enunciación de una creencia u opinión**
 - No se proporciona un fundamento sólido, real para tal creencia u opinión.
 - Aunque puede que exista la pretensión de que se reconozca tal creencia u opinión como verdadera, no hay un desarrollo sistemático de premisas-inferencia-conclusión en apoyo de lo enunciado.
- **Proposiciones vagamente relacionadas**
 - Las proposiciones no están conectadas por relación inferencial alguna.
- **Informes**
 - Son simples enumeraciones de hechos, del tipo que aparecen en las noticias de los periódicos.
 - No hay intención de probar nada, simplemente, se proporciona información sobre los hechos.
- **Ilustración**
 - Simplemente se ofrecen ejemplos de algo.
- **Explicaciones**
 - Consiste en una aclaración de por qué algo es el caso.
 - Una explicación a veces es difícil de distinguir de un argumento porque también involucra razones (similares a las premisas).
 - Pero, a diferencia de los argumentos, donde la conclusión es "nueva" información, en una explicación el enunciado que es explicado (el *explanandum*, la parte que parece la conclusión) es normalmente un hecho comúnmente aceptado.

Realizar anotaciones en BRAT I

- Realizar login
- Entrar a la carpeta de parte_#
- Anotar premisa y conclusiones
 - **Seleccionar texto** de premisas y conclusiones.
 - Aparece un recuadro
 - **Seleccionar la opción** correspondiente
 - Pulsar OK

New Annotation

Text

En la actualidad existen un gran número de empresas de servicios, públicas y privadas, que proveen insumos y/o servicios a otras empresas que exportan sus productos a mercados donde la calidad es fundamental,

Search

Google, Wikipedia

Entity type

Premisa
 Conclusion
 Nivel
 Tipo

Entity attributes

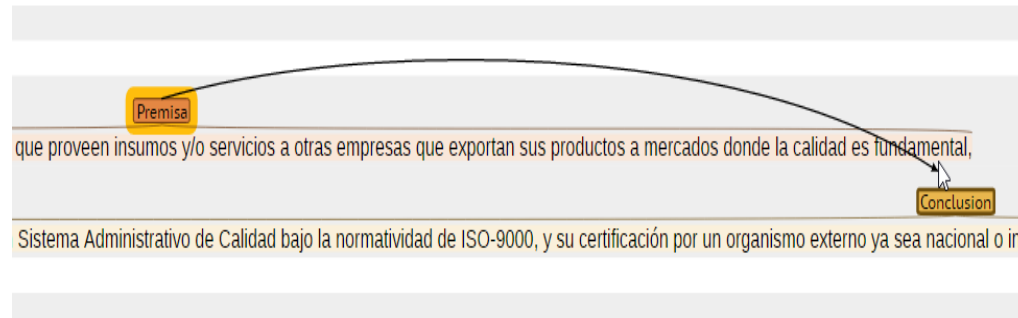
Nivel: ? Tipo: ?

Notes

OK Cancel

Realizar anotaciones en BRAT II

- Anotar relaciones
 - Seleccionar el recuadro del componente y dejarlo caer en el componente a enlazar.
 - Aparece un recuadro
 - Seleccionar la opción correspondiente (ataca o apoya)
 - Pulsar OK



The screenshot shows the "New Annotation" dialog box in BRAT II. The dialog has a title bar "New Annotation" with a close button. It contains the following fields:

- From:** Premisa ("En la actualidad existen un gran número de empresas de servicios, públicas y privadas, que proveen insumos y/o servicios a otras empresas que exportan sus productos a mercados donde la calidad es fundamental,")
- To:** Conclusion ("por ello éstas últimas están exigiendo a las empresas de servicios la implantación de un Sistema Administrativo de Calidad bajo la normatividad de ISO-9000, y su certificación por un organismo externo ya sea nacional o internacional, dependiendo del mercado o giro de la empresa y, que sea reconocido por la Organización Internacional para la Estandarización, quien es la encargada de emitir y revisar las normas ISO-9000.")
- Type:** Radio buttons for "apoya" (selected) and "ataca".
- Notes:** An empty text area.

At the bottom right of the dialog are "OK" and "Cancel" buttons.

Realizar anotaciones en BRAT III

- Indicar el nivel y el tipo de argumento observado
 - Pulsar doble clic en @CalificaPorNivelParrafoAnteriorPlanteamiento@ o @CalificaPorTipoParrafoAnteriorPlanteamiento@
 - Seleccionar la opción de Nivel o Tipo
 - Selecciona en la parte de atributos el nivel o tipo
 - Pulsar OK

12 | @CalificaPorNivelParrafoAnteriorJustificacion@

Edit Annotation

Text
CalificaPorNivelParrafoAnteriorJustificacion [Link](#)

Search
Google, Wikipedia

Entity type

- Premisa
- Conclusion
- Nivel
- Tipo

Entity attributes

- Nivel: Medio
- Nivel: ?
- Nivel: NoTiene
- Nivel: Medio
- Nivel: Fuerte
- Nivel: Debil

Add Frag. **Delete** **Move** **OK**

Cancel

Niveles de argumentos

Puntaje	Escala	Descripción	Ejemplos	
[0]	No tiene	No hay argumento ya que no se proveen justificaciones.	Estoy de acuerdo con él.	No es argumento
[1]	débil	No es un argumento completamente bueno (Se tiene una conclusión sin premisa)	[El problema de establecimiento de identidad tiene solución]/C1. [Los productos orgánicos deben ser marcados como buenos productos]/C1	No es argumento
[2]	media	Argumento con un razonamiento	En términos generales, [el problema de establecimiento de identidad tiene solución]/C1 [mediante un proceso de autenticación entre pares.]/P2 Relación de apoyo P2,C1	Argumento completo
[3]	fuerte	Argumentos con dos o más razonamientos (apoyo o ataque)	Podría comentar que [no estoy de acuerdo con la generación de alimentos transgénicos]/C1 [no han demostrado combatir la hambruna]/C2 [ya que este problema tiene un origen económico y político que viven los países pobres en cuanto a la desigualdad de repartición de recursos y no por problemas de agricultura.]/P3 Relación de apoyo C2,C1 P3,C2	Argumento completo

Selecciona Tipo de argumento

Edit Annotation [x]

Text
CalificaPorTipoParrafoAnteriorJustificacion [Link](#)

Search
[Google](#), [Wikipedia](#)

Entity type

- Premisa
- Conclusion
- Nivel
- Tipo

Entity attributes

- Tipo: Causal
- Tipo: ?
- Tipo: Ejemplos
- Tipo: Analogia
- Tipo: Ninguno
- Tipo: Autoridad
- Tipo: Comparacion
- Tipo: Causal

Cancel Move OK

Tipos de argumentos

- **Argumento por autoridad.** Se apoya en otras referencias —personas, organizaciones u obras de referencia más documentadas— para que nos expliquen gran parte de lo que necesitamos saber sobre el mundo (Weston 1994). Este tipo de argumentos son muy comunes en el ámbito académico, ya que generalmente el autor se apoya en las publicaciones realizadas por investigadores del área para sustentar sus aseveraciones.
- **Argumento de ejemplos.** Ofrece uno o más ejemplos específicos en apoyo de una generalización. Un ejemplo simple puede ser usado como ilustración, pero uno solo no ofrece suficiente apoyo. Se necesita más de un ejemplo para otorgar mayor fuerza al argumento.
- **Argumento por analogía.** Establece el punto de vista comparando conceptos o situaciones diferentes. Se trata de poder inferir conclusiones de una situación que no se conoce mediante la comparación con una situación que resulta familiar. De modo que lo que es aplicable para uno, es aplicable para otro.
- **Argumento por causalidad.** También llamado instrumental. Se establece una conexión causal entre dos hechos. Los argumentos son un medio para alcanzar la conclusión o constituyen una causa que la provoca.
- **Argumentos por comparación:** las comparaciones confrontan varios objetos para evaluarlos uno con relación a otro. Las comparaciones pueden efectuarse por oposición (lo pesado y lo ligero), por ordenación (lo que es más pesado que) y por ordenación cuantitativa (en este caso la pesada por medio de unidades de peso).

Revisar el material para tipos de argumentos en:

http://argumentlab.net/anotaciones/weston_tipos_argumentos.pdf

Contenido

1. Introducción
2. Objetivo y modelo
3. Construcción del corpus
4. Sistema de anotación BRAT
5. **Conclusión**

Conclusiones

- Existe la necesidad de apoyar a estudiantes y profesores en la evaluación de la argumentación.
- Las técnicas computacionales pueden ayudar en el **análisis estructural de los argumentos** para mejorar la evaluación de los escritos académicos.

Referencias I

- Argundi, Yolanda. (2005). Aprender a pensar escribiendo bien: desarrollo de habilidades para escribir. México, Trillas.
- Bañales Faz, G., Vega López, N. A., Araujo Alvinada, N., Valladares, A., & Rodríguez Zamarripa, B. (2015). La Enseñanza de la Argumentación Escrita en la Universidad. *Revista Mexicana de Investigación Educativa*, 20(66).
- Cabrio, E., & Villata, S. (2012). Combining textual entailment and argumentation theory for supporting online debates interactions. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2* (pp. 208-212). Association for Computational Linguistics.
- Capaldi, N. (2000). *Cómo ganar una discusión*. Gedisa.
- Carstens, L., & Toni, F. (2015). Towards relation based argumentation mining. In *Proceedings of the 2nd Workshop on Argumentation Mining*, (pp. 29-34). Association for Computational Linguistics, Denver, CO, USA.
- Florou, E., S. Konstantopoulos, A. Koukourikos, & P. Karampiperis. (2013). Argument extraction for supporting public policy formulation. In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 49-54).
- Freeman, J. B. (2011). *Argument Structure: Representation and Theory* (Vol. 18). Springer Science & Business Media.
- González-López, s., & López-López A. (2015). Colección de Tesis y Propuesta de Investigación en TICs: un recurso para su análisis y estudio. XIII Congreso Nacional de Investigación Educativa (pp. 15).
- López Ferrero, C. (2003). La argumentación en los géneros académicos. In *Actas del Congreso Internacional "La Argumentación"* (pp. 1121-1129). Buenos Aires: Universidad de Buenos Aires.
- Lawrence, J., & Reed, C. (2014). Combining Argument Mining Techniques. In *Proceedings of the 2nd Workshop on Argumentation Mining*, (pp. 127-136). Association for Computational Linguistics, Denver, CO, USA.
- Mochales, R., & Moens, M. F. (2011). Argumentation mining. *Artificial Intelligence and Law*, 19(1), 1-22.

Referencias II

- Moens, M. F., Boiy, E., Palau, R. M., & Reed, C. (2007). Automatic detection of arguments in legal texts. In Proceedings of the 11th international conference on Artificial intelligence and law (pp. 225-230). ACM.
- Nguyen, H. V., & Litman, D. J. (2015). Extracting argument and domain words for identifying argument components in texts. In Proceedings of the 2nd Workshop on Argumentation Mining (pp. 22-28). Association for Computational Linguistics, Denver, CO, USA.
- Ong, N., Litman, D., & Brusilovsky, A. (2014). Ontology-based argument mining and automatic essay scoring. In Proceedings of the First Workshop on Argumentation Mining, (pp. 24–28). Association for Computational Linguistics, USA.
- Persing, I., Davis, A., & Ng, V. (2010). Modeling organization in student essays. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing (pp. 229-239). Association for Computational Linguistics.
- Persing, I., & Ng, V. (2015). Modeling argument strength in student essays. In Proceedings of the 7th International Joint Conference on Natural Language Processing (pp. 543–552). Association for Computational Linguistics, Beijing, China.
- Persing, I., & Ng, V. (2016). End-to-End Argumentation Mining in Student Essays. In Proceedings of NAACL-HLT (pp. 1384-1394).
- Reed, C., Mochales Palau, R., Rowe, G., & Moens, M. F. (2008). Language resources for studying argument. In Proceedings of the 6th conference on language resources and evaluation-LREC 2008 (pp. 91-100).
- Stab, C., & Gurevych, I. (2014). Identifying Argumentative Discourse Structures in Persuasive Essays. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), (pp. 46–56).
- Wachsmuth, H., Kiesel, J., & Stein, B. 2015. Sentiment Flow—A General Model of Web Review Argumentation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (pp. 601-611).
- Weston, Anthony. (2005). Las claves de la argumentación. Ed. Ariel, Barcelona, España.

Gracias por su atención

Dudas