UNIVERSIDADE DE SANTIAGO DE COMPOSTELA
Departamento de Electrónica e Computación

Tese de doutoramento

APPROXIMATE SYLLOGISTIC REASONING: A CONTRIBUTION TO INFERENCE PATTERNS AND USE CASES

Presentada por:
Martín Pereira Fariña

Dirigida por:
Alberto J. Bugarín Diz
Alejandro Sobrino Cerdeiriña

Santiago de Compostela, Novembro 2013

Alberto J. Bugarín Diz, Catedrático de Universidade da Área de Ciencia da Computación e Intelixencia Artificial da Universidade de Santiago de Compostela

Alejandro Sobrino Cerdeiriña, Profesor Titular da Área de Lóxica e Filosofía da Ciencia da Universidade de Santiago de Compostela

FAN CONSTAR:

Que a memoria titulada **APPROXIMATE SYLLOGISTIC REASONING: A CONTRIBUTION TO INFERENCE PATTERNS AND USE CASES** ven de ser realizada por **D. Martín Pereira Fariña** baixo a nosa dirección no Departamento de Electrónica e Computación da Universidade de Santiago de Compostela, e constitúe a Tese que presenta para optar ao grao de Doutor.

Santiago de Compostela, Novembro 2013

Alberto J. Bugarín Diz
Codirector da tese

Alejandro Sobrino Cerdeiriña
Codirector da tese

Martín Pereira Fariña
Autor da tese

Á miña familia e a Noa,
polo seu constante apoio

¿Cantas veces lle teño dito que se eliminamos o imposible, o que queda, por improbable que pareza, ten que ser a verdade?

Sherlock Holmes, *O Signo dos Catro*

Agradecementos

Unha vez rematado este longo e duro traballo pero á súa vez estimulante e enriquecedor que é facer unha tese de doutoramento, é de xustiza dedicar as primeiras liñas da mesma a expresar o meu máis sincero agradecemento a todas aquelas persoas e institucións que fixeron posible o seu desenvolvemento.

Quero encomezar polos codirectores da tese, os profesores Alberto José Bugarín Diz e Alejandro Sobrino Cerdeiriña, quen seguiron este traballo dende o primeiro ata o último día; sen a súa confianza e axuda esta investigación non tería sido posible. Tamén quero agradecer especialmente a axuda dos doutores Félix Díaz Hermida e Juan Carlos Vidal Aguiar, pois sen a súa colaboración este volume tampouco tería visto a luz. Por último, non podo deixar de mostrar o meu agradecemento a Alejandro Ramos e a Pedro Montoto, pois colaboramos cóbado con cóbado no desenvolvemento da libraría software.

Tamén quero expresar o meu agradecemento ao Departamento de Electrónica e Computación, principalmente ao Grupo de Sistemas Intelixentes, ao de Lóxica e Filosofía Moral e ao Centro Singular de Investigación en Tecnoloxías da Información (CITIUS), pois foi no seu seo e baixo os seus teitos nos que desenvolvín a maior parte do meu traballo. Tampouco podo esquecerme do European Centre for Soft Computing nin da xente que alí traballa, pois durante o outono do 2011 pasei entre eles unha tremendamente agradable estadía durante a cal puiden manter longas conversas co doutor Gracián Triviño Barros e das que xurdiron frutíferas ideas; nin da Universidade Técnica de Viena, pois durante o outono do 2012 acolléronme moi agarimosamente para unha nova estadía durante a cal se me brindou a oportunidade de coñecer o traballo do Prof. Dipl.-Ing. Dr. techn. Christian Georg Fermüller e do seu doutorando Chris Roschger.

Tampouco podo esquecerme dos vellos e novos amigos labrados durante estes anos, pois eles fixeron moito máis levadía esta tarefa. E, por suposto, agradecer o apoio da miña familia e de Noa, que dende sempre me apoiaron para que acadara os meus soños e metas con plena confianza no meu criterio.

Para rematar, non podo deixar de mostrar o meu agradecemento ao apoio económico recibido durante estes anos. En primeiro lugar, destacar o soporte ofrecido pola Xunta de Galicia a través das súas axudas para a “Realización de estudos de terceiro ciclo nas universidades do sistema universitario de Galicia” e polo Ministerio de Educación, Cultura y Deporte a través das bolsas do seu programa “Formación del Profesorado Universitario (FPU)”. Ao Ministerio de Economía y Competitividad, polo soporte recibido a través do proxecto “Descubrimiento automático de reglas borrosas no convencionales: Bases de Conocimiento Temporal Borroso y TSK de Estructura Variable” (TIN2008-00040) e ao mesmo ministerio xunto cos Fondos Europeos para o Desenvolvemento Rexional (ERDF/FEDER) tamén polo proxecto “Descripción Lingüística de Fenómenos Complejos: Cuantificadores Borrosos Generalizados en Proposiciones Temporales” (TIN2011-29827-C02-02); de novo aos fondos ERDF/FEDER, neste caso en colaboración coa Xunta de Galicia, polos proxectos CN2012/151 e CN2011/058; ao Grupo de Sistemas Intelixentes e ao CITIUS polo seu apoio económico á hora de realizar as estadías e asistir a congresos.

Santiago de Compostela, Novembro 2013

Contents

Resumen de la tesis

La investigación que presentamos en esta memoria de tesis doctoral se encuadra en el marco de la Computación con Palabras y del Razonamiento Aproximado. En ella abordamos el tratamiento de un tipo de inferencia deductiva, el silogismo, cuyo patrón de razonamiento se basa en el encadenamiento de términos mediante el uso de oraciones cuantificadas de la forma *Q S son P*, donde *Q* designa un cuantificador binario, *S* el término sujeto y *P* el término predicado.

En su aproximación clásica y nítida, desarrollada por Aristóteles [7], se contemplan argumentos compuestos por tres términos, dos premisas y una conclusión. De los tres términos, dos son los llamados *extremos*, que aparecen en la conclusión, y el otro es el *término medio*, que sirve de conexión o encadenamiento entre ellos y únicamente aparece en las premisas. Los cuantificadores considerados son los cuatro clásicos: "todos", "ningún", "alguno", "alguno... no" (organizados en el Cuadrado Lógico de la Oposición); por ello todo silogismo se reduce a una combinación de las cuatro oraciones generadas por estos cuantificadores y las distintas posiciones que puede ocupar el término medio en las premisas (las cuatro Figuras aristotélicas).

De la silogística aristotélica cabe destacar que la forma lógica de las proposiciones es su forma gramatical, por lo que la construcción de un silogismo resulta intuitiva y sencilla para cualquier usuario. Sin embargo, dado que se limita a sólo cuatro tipos de oraciones y argumentos de dos premisas, su capacidad expresiva también resulta muy restringida. Además, tampoco se consideran cuantificadores imprecisos o vagos como "pocos", "muchos", "casi todos",... lo que limita todavía más su expresividad. Por otra parte, el número de términos aceptados en los argumentos (tres) también es insuficiente ya que, aunque los humanos no elaboramos argumentos muy largos, si pueden aparecer fácilmente algunos más.

La hipótesis de trabajo de nuestra investigación ha sido que los modelos existentes de razonamiento silogístico sólo consideran una pequeña parte de los silogismos que los humanos podemos manejar, tanto desde el punto de vista de los cuantificadores como de los patrones de inferencia. Como explicaremos, estos siguen siendo bastante limitados y, por esta razón, hemos establecido como principal propósito de nuestra investigación el desarrollo de un nuevo modelo de silogismo que contribuya a mejorar la capacidad de la Computación con Palabras para tratar inferencias usando lenguaje natural en este contexto. Hemos definido siete objetivos (detallados en el prefacio de esta memoria de tesis), que en este resumen agrupamos en tres: i) analizar los modelos silogísticos propuestos e identificar los puntos débiles a mejorar así como los diversos aspectos a desarrollar; ii) definir un nuevo modelo para la representación de las oraciones cuantificadas y un nuevo patrón de inferencia que no sólo incluya cuantificadores vagos sino también imprecisión en el encadenamiento; iii) analizar casos de uso del silogismo no previamente considerados.

Análisis de diversos modelos de silogismo

En la bibliografía podemos encontrar diversas propuestas que tratan de superar las mencionadas limitaciones. Analizamos aquellas que se centran en introducir los cuantificadores vagos como “muchos”, “pocos”,... en los mismos (o prácticamente) patrones de inferencia. En función de cómo estos modelos interpretan las proposiciones, los clasificamos en dos categorías: a) interpretación basada en conjuntos; b) interpretación condicional.

La interpretación basada en conjuntos asume que los términos de las proposiciones denotan conjuntos de elementos y los cuantificadores una relación de cardinalidad entre ellos. Por ejemplo, “todos los hombres son mortales” significa que el conjunto *hombres* está incluido totalmente en el conjunto de *mortales*. Describimos cinco modelos diferentes: la *Intermediate Syllogistics*, la *Exception Syllogistics*, la *Interval Syllogistics*, la *Fuzzy Syllogistics* y la *Generalized Intermediate Syllogistics*.

La *Intermediate Syllogistics* [63] define un nuevo Cuadrado Lógico de la Oposición en el que se introducen nuevos cuantificadores, los llamados intermedios (por ejemplo, “muchos”,...), que se sitúan en el espacio entre el universal y el existencial. La *Exception Syllogistics* [39] destaca porque introduce los cuantificadores no proporcionales, en particular los llamados de excepción (por ejemplo, “todos menos tres”,...), y una reescritura de los proporcionales en estos términos. Este modelo también admite la conclusión múltiple, lo que le da un cariz heurístico a este tipo de inferencias (hay varias conclusiones válidas posibles). Sin embargo, ambos modelos se limitan a las figuras aristotélicas y a usar definiciones nítidas para los cuantificadores.

Los modelos *Interval Syllogistics* [18], *Fuzzy Syllogistics* [93] y *Generalized Intermediate Syllogistics* [37] pertenecen a la lógica borrosa. Todos comparten como objetivo el modelado de los cuantificadores imprecisos mediante conjuntos borrosos aunque con diferentes aproximaciones. La *Interval Syllogistics* modela los cuantificadores mediante sub-intervalos del intervalo $[0,1]$ e interpreta el proceso de razonamiento como un proceso de optimización, donde las premisas del silogismo constituyen un conjunto de restricciones y la conclusión el máximo intervalo compatible con ellas. La *Fuzzy Syllogistics* identifica los cuantificadores con números borrosos (proporcionales cuando los cuantificadores son proporcionales y absolutos cuando son absolutos) obtenidos de aplicar la medida de cardinalidad borrosa $\Sigma - count$ a los términos e interpreta el proceso de razonamiento aplicando el Principio de Extensión de Cuantificadores y la aritmética borrosa. Por último, en la *Generalized Intermediate Syllogistics* se propone una generalización de la *Intermediate Syllogistics* en el marco de una teoría de tipos borrosos pero usando definiciones borrosas para los cuantificadores.

La interpretación condicional, por su parte, asume que una oración cuantificada está expresando, en su estructura profunda, una oracional condicional donde el cuantificador denota el grado de fortaleza de la conexión entre el antecedente y el consecuente. Por ejemplo, "todos los hombres son mortales" significa que para cualquier elemento x del universo, "si x tiene la propiedad de ser hombre, entonces x es mortal con un grado de fortaleza 1". Los cuantificadores borrosos, en este caso, denotan un vínculo más débil.

Nosotros analizamos tres modelos: *Probabilistic Syllogistics* [43], *Support Logic Syllogistics* [79] y *Qualified Syllogistics* [71]. La *Probabilistic Syllogistics*, desarrollada en el ámbito de la psicología del razonamiento, asume que los cuantificadores expresan realmente probabilidad; por ejemplo, "todos" significa "100 %" o "casi todos" es "más del 90 %". Desde el punto de vista de los patrones de inferencia, aunque se limita a analizar las cuatro Figuras aristotélicas, asimila el razonamiento a un proceso heurístico aunque a veces nos conduzca a conclusiones erróneas. La *Support Logic Syllogistics* sustituye la probabilidad por la teoría de la evidencia de Dempster-Shafer; así pues, un cuantificador viene a denotar el grado de confianza del usuario en la relación condicional. Sobre los patrones de inferencia, considera sólo algunos de la *Fuzzy Syllogistics* pero obteniendo resultados diferentes. Por último, la *Qualified Syllogistics* establece que los cuantificadores pueden denotar cantidad, frecuencia o probabilidad puesto que existe cierta equivalencia entre estos conceptos; por ejemplo, "casi todos" $\equiv$ "casi siempre" $\equiv$ "casi seguro". Respecto de los patrones de inferencia, simplemente asume el *Modus Ponens* sin ningún tipo de referencia al patrón de silogismo típico.

A partir del análisis de estos modelos hemos establecido los puntos a desarrollar en nuestro modelo de silogismo. Varios de los modelos analizados tratan los cuantificadores borrosos mediantes definiciones nítidas; aquellos que efectivamente usan definiciones borrosas no pueden tratar correctamente los modos clásicos. Además, casi todos los modelos consideraron únicamente cuantificadores proporcionales, descartando otros definidos por la Teoría de los Cuantificadores Generalizados (TGQ) como los absolutos, comparativos,... que también son útiles para hacer inferencias. Por otra parte, los patrones de inferencia propuestos se limitaron prácticamente (excepto en los modelos borrosos) a las cuatro figuras aristotélicas. Por último, los patrones heurísticos, que permiten más inferencias que las deducciones, o bien sólo consideraban los patrones aristotélicos (como la *Probabilistic Syllogistics*) o no lo manejaban correctamente (como la *Interval Syllogistics*).

Desarrollo de nuestro nuevo modelo de silogismo

A partir de las conclusiones anteriores definimos tres puntos a satisfacer por nuestro nuevo modelo de silogismo [56]: i) incluir todos los tipos de cuantificadores binarios definidos por la TGQ; ii) desarrollar un patrón de inferencia que permita más inferencias que las deductivas; iii) modelar argumentos con cualquier número de términos y premisas sin necesidad de seguir un patrón previamente predefinido.

Adoptamos la interpretación basada en conjuntos, pues es compatible con la TGQ, nos permite el tratamiento de cuantificadores no proporcionales y puede ser gráficamente representada de forma intuitiva y simple mediante diagramas de Venn. Respecto del patrón de inferencia, tomamos como punto de partida la *Interval Syllogistics* [18] pues, a pesar de no considerar el caso clásico, si es compatible con él, considera cuantificadores vagos y precisos y nos permite adoptar una perspectiva heurística al interpretar el razonamiento como un proceso de optimización matemática.

Nuestro primer paso en el desarrollo del modelo ha sido la definición de las oraciones cuantificadas según la TGQ. Así pues, cualquier oración cuantificada binaria se expresa mediante una ecuación entre los términos de la proposición y la cardinalidad denotada por el cuantificador. Por lo tanto, para cualquier tipo de cuantificador binario de la TGQ obtenemos una expresión que se refiere unívocamente a una región del universo del discurso. Por ejemplo, en el caso de los cuantificadores precisos, “quince estudiantes son altos” se define como $|estudiantes| \cap |altos| = 15$, donde $|estudiantes|$ denota la cardinalidad del conjunto estudiantes y $|altos|$ la cardinalidad del de personas altas; en los cuantificadores borrosos, “pocos estudiantes son altos” se define como $\frac{|estudiantes| \cap |altos|}{|estudiantes|} = pocos$.

El siguiente paso fue el patrón de inferencia. Siguiendo la *Interval Syllogistics*, las premisas del argumento se interpretan como un conjunto de restricciones y la conclusión como el valor que tiene que ser máximamente compatible con esas restricciones. Así pues, transformamos el problema de razonamiento en un problema de optimización, dividiéndolo en tres etapas: i) división del universo en conjuntos disjuntos; ii) definición de las proposiciones como sistemas de inecuaciones; iii) aplicación del método de optimización necesario en cada caso para resolver el problema del razonamiento.

En el paso i), apoyándonos en los diagramas de Venn, identificamos todos los conjuntos disjuntos de los términos del silogismo. Cada premisa del argumento designa a un subconjunto o a una operación booleana entre ellos y el cuantificador la cardinalidad correspondiente. Por lo tanto, cada una de las premisas genera un sistema de inecuaciones.

En el paso ii) se organizan las inecuaciones obtenidas en el paso i). Si los cuantificadores son precisos, la inecuación se transforma en una igualdad dado que el cuantificador denota un valor preciso; si son borrosos o imprecisos, cada una de las inecuaciones denotará los límites superiores e inferiores de los cuantificadores, definidos como intervalos o conjuntos borrosos (por ejemplo, aproximación mediante núcleo-soporte). Esto nos permitió identificar unívocamente regiones del universo del discurso del silogismo definidos por cada una de las premisas.

En el paso iii) se selecciona el método de optimización adecuado para cada caso. En esta investigación usamos dos: Simplex, para los casos en los que el cuantificador de la conclusión es absoluto y programación fraccional para los casos en los que el cuantificador de la conclusión es proporcional. Para la correcta aplicación de dichos métodos es necesario añadir tres restricciones adicionales: i) que no hay ningún conjunto con un número negativo de elementos; ii) en las premisas con forma fraccional, no hay ningún denominador igual a 0; iii) la suma de las cardinalidades de todos los subconjuntos es igual a la cardinalidad del referencial.

Evaluamos este modelo en el tratamiento del silogismo aristotélico y de los otros modelos silogísticos analizados, obteniendo unos resultados consistentes. Además, también obtuvimos otras mejoras significativas, como el uso de combinaciones booleanas de términos, el tratamiento de diferentes tipos de cuantificadores en un mismo argumento y el modelado de argumentos con cualquier número premisas o términos. Por otra parte, al adoptar una aproximación heurística, abrimos la posibilidad de usar el silogismo en contextos no previamente considerados. También cabe destacar que este modelo ha sido parcialmente implementado en la librería software SEREA (Syllogistic Epistemic REAsoner) con la que se realizaron la mayor parte de los ejemplos.

La segunda parte del objetivo ii) fue el tratamiento de los términos vagos en el encadenamiento. Esto es especialmente significativo, puesto que es el nexo entre los términos extremos. Aquellos silogismos con un encadenamiento no exacto son considerados falaces (*falacia de los cuatro términos*) por la lógica clásica, aunque sean aceptables desde el punto de vista del razonamiento ordinario. Este tipo de casos se da, por ejemplo, cuando el término medio está desempeñado por dos palabras de significado similar.

En la lógica borrosa existen diferentes aproximaciones al tratamiento de la similitud entre conceptos (como la inclusión borrosa o el solapamiento borroso), pero nosotros la abordamos usando la sinonimia, que es la relación de similitud lingüística por excelencia. Para el desarrollo de nuestra propuesta [55] asumimos tres principios: i) la sinonimia es una relación aproximada que refleja un parecido de familia entre términos; ii) dos términos son sinónimos si aparecen como tales en un diccionario de sinónimos (*thesaurus*; como WordNet en nuestro caso); iii) utilizamos diversas medidas de distancia entre conjuntos para medir o establecer el grado de sinonimia entre dos palabras.

Cada palabra incluida en WordNet está descrita por una serie de glosas (breves descripciones conceptuales de cada una de las acepciones del término), y cada una de ellas tiene asociado un conjunto de sinónimos (*synset*). Si dos términos son sinónimos es plausible pensar que sus synsets compartirán un elevado número de sinónimos, por lo que asumimos esta relación como principio. Para obtener un valor cuantitativo de esta medida, aplicamos alguna de las métricas de distancia (coeficiente de Jaccard, de Solape,...) sobre los synsets de los términos en comparación: si el resultado es 1, los términos son idénticos; si es 0 los términos no se parecen en nada.

Sin embargo, las distintas medidas generan resultados significativamente diferentes, por lo que surge la cuestión de cuál usar en cada caso. Para abordar este problema, proponemos usar las glosas. Si los términos en comparación comparten muchas de ellas, deben utilizarse métricas que ofrezcan unos valores más elevados, como el coeficiente de Solape; si comparten pocas, se usarán las de resultado más conservador, como el coeficiente de Jaccard.

Respecto del patrón de inferencia, apuntamos que el encadenamiento aproximado tiene que verse reflejado en la fiabilidad del argumento. En este caso proponemos cualificar la conclusión del argumento mediante un grado de confianza bajo el siguiente principio: cuanto mayor sea el grado de sinonimia entre los términos, mayor es el grado de confianza en la conclusión. Así pues, la conclusión de un silogismo con un grado de similitud alto en el encadenamiento aproximado será más fiable que uno con uno bajo. El desarrollo de este modelo también es una contribución efectiva a la mejora de la expresividad del silogismo, pues nos permite tratar un tipo de inferencia aproximada usando un concepto típicamente lingüístico como la sinonimia.

Nuevos casos de uso del silogismo

Como indicamos, nuestro modelo silogístico tiene una mayor capacidad expresiva, que nos permite considerar usos del silogismo no contemplados hasta ahora. En particular, abordamos el caso del razonamiento probabilístico, dado que es un tipo de formalismo ampliamente estudiado y utilizado en la Inteligencia Artificial y dada la equivalencia entre las oraciones cuantificadas proporcionales y las probabilidades condicionales. Analizamos tres modelos distintos: el Teorema de Bayes Generalizado, las redes bayesianas y el razonamiento probabilístico en el ámbito judicial.

El Teorema de Bayes Generalizado se caracteriza por usar exclusivamente probabilidades condicionadas. Para razonar con él, se necesita de una combinación entre el Patrón I de la *Interval Syllogistics* y el propio teorema y aplicar una serie de pasos iterativos para obtener los resultados óptimos. Usando nuestro modelo de silogismo mostramos como construir un argumento silogístico en el que se incluyen todas las probabilidades condicionadas y así obtener en un único paso la mayoría de los resultados óptimos.

Las redes bayesianas son ampliamente utilizadas en el modelado de situaciones con incertidumbre. Destacan por su gran versatilidad y su capacidad para ejecutar diferentes tipos de inferencias. A partir de un ejemplo de un diagnóstico médico, se ilustra como elaborar un silogismo equivalente y obtener los mismos resultados en la ejecución de las inferencias básicas de las redes bayesianas (predictiva simple, predictiva con evidencia, diagnóstico e intercausal). Un silogismo, basado en oraciones cuantificadas del lenguaje natural y una estructura de premisas-conclusión, resulta más sencillo para usuarios no familiarizados, dado que las redes bayesianas son un formalismo que usualmente requiere la intervención de expertos. Por otra parte, el silogismo también nos permite incluir vaguedad o imprecisión con sencillez, simplemente substituyendo los cuantificadores precisos por cuantificadores vagos.

Por último, analizamos el uso del razonamiento probabilístico en la argumentación jurídica. En ciertas ocasiones, esta no se aplica correctamente y aparece la llamada *falacia del fiscal*, que consiste en un uso erróneo del razonamiento probabilístico con el objetivo de probar la culpabilidad de un acusado. Usando nuestro modelo, todas las probabilidades pueden ser expresadas en un único argumento mediante oraciones cuantificadas y, dado que el cuantificador de la conclusión busca el intervalo de máxima compatibilidad, si el conjunto de premisas no está adecuadamente definido se obtendrá una indeterminación. Así pues, la expresión de este tipo de argumentos usando nuestro modelo facilitaría la identificación de argumentos mal fundamentados.

Para concluir, consideramos que hemos desarrollado satisfactoriamente todos los aspectos relevantes de nuestra hipótesis de trabajo. El modelo desarrollado presenta una capacidad expresiva significativa, que permite el tratamiento de argumentaciones complejas que incluyen tanto cuantificadores borrosos como un encadenamiento aproximado. Esto, conjuntamente con la perspectiva heurística adoptada, permite abrir nuevas posibilidades de uso para el silogismo, como la expresión de las redes bayesianas usando oraciones cuantificadas.

Como trabajo futuro apuntamos la incorporación de otros cuantificadores de la Teoría de los Cuantificadores Generalizados, como los ternarios o cuaternarios; el tratamiento de redes bayesianas más complejas (más allá de los casos básicos); explorar otros conceptos para evaluar la similitud entre términos, como la analogía o la antonimia y el modelado del contexto. Por último, también proponemos avanzar en el desarrollo de SEREA, desde su versión actual (librería, que también debe ser mejorada), a una herramienta software gráfica más completa y análoga a las existentes para el modelado de las redes bayesianas.

Preface

Reasoning takes part in any human activity and, by this omnipresence, its study has been a constant in human thinking and in our understanding about the world. We reason using inferences, which are expressed in natural language using arguments, and they have the following standard structure: i) *premises*, which are a set of statements; ii) *conclusion*, a statement inferred from these premises. However, preserving this structure, the validity of the inferences can vary significantly depending on the context or on its aim. For instance, let us consider the case of B. Russell's *Inductivist Turkey* [70]:

> The turkey found that, on his first morning at the turkey farm, he was fed at 9 a.m. Being a good inductivist turkey he did not jump to conclusions. He waited until he collected a large number of observations that he was fed at 9 a.m. and made these observations under a wide range of circumstances, on Wednesdays, on Thursdays, on cold days, on warm days. Each day he added another observation statement to his list. Finally he was satisfied that he had collected a number of observation statements to inductively infer that "I am always fed at 9 a.m.". However on the morning of Christmas eve he was not fed but instead had his throat cut.

Inductive arguments allow us to make a prediction in the future according to the information collected in the past about the same event. In this case, the premises are all evidence that the turkey gathered during the days in the farm and the (wrong) conclusion was "I am always fed at 9 a.m.". Let us now consider the following example of inference [46]:

> All the beans from this bag are white. These beans are white. Therefore, these beans are from this bag.

In this case, the premises are "all the beans from this bag are white" and "these beans are white" and the conclusion "these beans are from this bag". This is not an inductive argument since there is not a collection of evidence that allows us to infer something about the future but the conclusion is rather a hypothesis which explains the premises. This is the so-called abductive reasoning, and, as we can see, with the same structure of premises and conclusion it generates a different type of argument.

The type of reasoning *par excellence* is perhaps deductive reasoning. It is characterized by being sure, definitive and free of controversy [65], which gives us a solid base for supporting our knowledge. For this reason, it is the dominant one in the scientific field, particularly in probative fields such as mathematics.

The first systematic study of deductive reasoning in the history of logic was developed by Aristotle in Ancient Greece [4]. He focused on the so-called *syllogism*. This is a deductive argument of a very specific sort: it only deals with statements having the form *All S are P* or *Some S are P* (and their respective the negated versions), where *S* stands for the subject and *P* stands for the predicate and with arguments comprising three terms or concepts, two premises and one conclusion. It was the hegemonic logical system from Ancient Greece until the nineteenth century, when mathematical logic appears [29]. At this moment, Aristotelian Syllogistics was completed rejected. In [70, p. 202], B. Russell says:

> I conclude that the Aristotelian doctrines with which we have been concerned in this chapter are wholly false, with the exception of the formal theory of the syllogism, which is unimportant. Any person in the present day who wishes to learn logic will be wasting his time if he reads Aristotle or any of his disciples. None the less, Aristotle's logical writings show great ability, and would have been useful to mankind if they had appeared at a time when intellectual originality was still active. Unfortunately, they appeared at the very end of the creative period of Greek thought, and therefore came to be accepted as authoritative. By the time that logical originality revived, a reign of two thousand years had made Aristotle very difficult to dethrone. Throughout modern times, practically every advance in science, in logic, or in philosophy has had to be made in the teeth of the opposition from Aristotle's disciples.

Nevertheless, in the second half of the twentieth century new perspectives about Aristotelian Syllogistics appear [33, 13] and it was partially recovered. It is relevant to note that all syllogistic statements are quantified statements and, quantification is a typical phenomenon of natural language. In [86], D. Westersthål shows the compatibility between Aristotelian Syllogistics and the Theory of Generalized Quantifiers (TGQ), the most complete and up-to-date theory of quantification in linguistics, opening thus the possibility of interpreting syllogism as a form of deductive reasoning about quantities.

Nonetheless, quantification not only denotes precise quantities but also vague ones. For instance, terms such as “most”, “many”, “few”, etc. are quantifiers that are very common in natural language and refer to a more or less indeterminate quantity, but are usually sufficiently informative to the individuals, even for making inferences. However, as stated above, Aristotelian Syllogistics only dealt with two types of propositions (which comprise the universal quantifier “all” and the existential one “some” with their corresponding oppositions); therefore, a relevant field for the development and expansion of syllogistics beyond Aristotle’s scope with the combination of deductive reasoning involving vague quantities appears.

There are several different proposals that expand syllogistic reasoning following the Aristotelian concept of quantified statement; i.e., a quantifier denotes a quantity relationship between two sets. Some of these proposals are based on an extension or generalization of the Logic Square of Opposition (LSO), the Aristotelian classification of syllogistic quantified statements, such as Intermediate Syllogistics [63] of Generalized Intermediate Syllogistics [37, 36]; others are based on a reinterpretation of quantifiers in terms of precise quantities, such as the Exception Syllogistics [38, 39]; others on the use of intervals to represent vague quantifiers, such us the model developed by D. Dubois et al. [17, 18, 16]; and, another on the interpretation of vague quantifiers through fuzzy sets such us Zadeh’s proposal [92].

However, during the twentieth century alternative interpretations for the quantifiers also appear, where a quantifier can also denote probabilities, beliefs or even modal references. The most used one is the interpretation of quantified statements in terms of conditional probabilities, as it opens up the possibility of a combination between syllogism and Bayesian reasoning [3, 16], a type of probabilistic reasoning that has been deeply studied. In this case, it is assumed that any proportional quantifier, such as “all”, “most”, “few”, etc. denotes probability. Thus, for instance, “all mammals are animals” is equivalent to saying “if something is a mammal then it is an animal with 100% probability”, in the case of vague quantifiers, such as “most mammals live on land” is equivalent, for instance, to saying “if something is a mammal then it lives on land with more than 80% probability” [43].

Quantifiers can also be interpreted as beliefs. In [79], M. Spies proposes to use Dempster-Shafer’s Theory, where the belief is represented through an interval. Thus, a quantifier denotes the belief of a person in the relationship between the terms involved in the statement. For instance, "all mammals are animals" means that "I believe that if something is a mammal then it is an animal with $[1,1]$ degree", in the case of vague quantifiers, such as "most mammals live on land" means "I believe that if something is a mammal then it is an animal with $[0.8,0.9]$ degree". Starting from this concept, M. Spies [79] built a syllogistic framework that give us a new perspective about this kind of reasoning.

Other approach, elaborated by D. Schwartz [71], proposes interpreting quantifiers in terms of modality. For instance, "most mammals live on land" means, for instance, "usually, mammals live on land", where the modifier "usually" assumes the meaning of the quantifier "most". However, the treatment of inference patterns, as we shall explain, does not follow the standard characteristics of syllogistic reasoning but is based on *Modus Ponens*. Despite this, it constitutes an alternative approach to syllogistic reasoning to be considered.

All the previous approaches, irrespective of how they model the meaning of a quantified statement, only deal with a very limited number of inference patterns. For instance, Intermediate Syllogistics deals with more than one hundred inference patterns, but all of them are grouped into the four Aristotelian figures. Fuzzy proposals and those with an alternative interpretation of quantified statements also manage a reduced number of patterns, although some of them are different from Aristotle's ones. Thus, although the capability for managing vague statements of natural language was notably increased, they were only focused on a particular type of quantifiers (mainly on proportional ones), disregarding the other ones considered by the TGQ, and on a very reduced number of patterns of inference with no more than three or four premises, thus significantly limiting the capability of the syllogism to manage moderately complex problems.

On the other hand, vagueness can also appear in the terms of the syllogism, not only in quantifiers. In this case, the syllogism's probative character is replaced by other approximated or probable one since the links between vague terms are based on similarity rather than equality. Thus, we here find another way of extending syllogistic reasoning. We think in addition that, since we shall approach syllogism from the perspective of natural language, the concept of synonymy is the linguistic relationship that best represents semantic similarity between two terms. There are currently tools available on line for obtaining synonym relationships between terms, such us WordNet [1] or Galnet [26], which offer us the environment for proposing a way of establishing the synonymy between two terms using measures of similarity.

Thus, taking into account all the ideas explained above, we propose in this research extending syllogistic model of reasoning dealing with vagueness into two ways: managing vague quantifiers and vague terms. In the first case, we use the framework of the TGQ to model the meaning of the quantifiers statements; Venn diagrams to build the knowledge representation and algebraic methods to build the reasoning process. In the second case, we use the concept of synonymy, taken from linguistics and the philosophy of language, to measure the similarity between two terms; distance measures to calculate the similarity between two concepts and fuzzy inferences rules to develop the reasoning procedure associated to this kind of arguments. As we can see, this research constitute a multidisciplinary work where several fields such us logic, philosophy of language, linguistics, pragmatics and computer science are involved to develop a model of reasoning for dealing with arguments expressed in natural language that comprise fuzzy quantified statements.

Hypothesis

The current models for syllogistic reasoning only deal with a small part of syllogisms that human beings can manage and exhibit certain limitations, in the quantifiers, terms and inferences schemas that they manage. This makes them hardly applicable to realistic scenarios related to everyday reasoning. The existing proposals only consider arguments of a very specific type many of them differing greatly from the use that people do in their daily life. The development of a new model for syllogistic reasoning compatible with the TGQ and dealing with vagueness would contribute to improve the capability of Computing with Words to manage natural language arguments.

Objectives

Thus, the main aim of this research is **to define a model for dealing with syllogistic arguments that is able to manage the quantifiers defined by TGQ and approximate terms to make inferences expressed in natural language**. The achievement of this aim is supported by the following objectives:

1. **Studying the existing proposals about syllogistic reasoning to identify how quantified statements are interpreted**. The analysis of syllogistic reasoning was framed in the discipline of logic from the very beginning and the analysis of quantified statements

has changed gradually over time. The objective of this step is to analyse some of these proposals and, after an in-depth analysis, to identify which of them are the most adequate to our aim. We also adopt Aristotelian Syllogistics as the referential framework, since it is the classical system.

2. **Defining a model for the representation of quantified statements from the point of view of reasoning and the TGQ**. We assume the Sapir-Whorf hypothesis [28] and, hence, depending on the structure assigned to quantified statements, we fix a particular conceptualization of the world. Thus, in this step, we define the knowledge representation associated to quantified statements which also determines the general form of our inference procedure.

3. **Elaborating a procedure to infer a valid conclusion that manages different quantifiers simultaneously**. Attending to the two previous objectives, we define the steps that constitute the reasoning procedure and how it operates to infer a valid conclusion from a set of premises.

4. **Studying different alternatives to measure the similarity between two terms in approximate syllogisms.** The validity of syllogisms with an approximate chaining directly depends on how similar the involved terms are. Hence, in this stage we address different proposals that analyse this concept.

5. **Proposing a mechanism to execute inferences with approximate chaining in the middle term**. Incorporating vagueness or approximation in middle terms entails a reconfiguration of syllogistic inference as it focuses principally on the deduction of the quantifier. We provide an approach for measuring the similarity between two concepts and how its results are interpreted in terms of validity or reliability of the syllogism instead of in terms of quantifiers, as in the previous problem.

6. **Analysing the compatibility between some types of Bayesian reasoning and syllogism supported in the equivalence between quantified statements and conditional probabilities**. It is accepted that quantified statements can be interpreted in terms of conditional probabilities. Thus, if we extend this equivalence to reasoning, certain types of Bayesian reasoning can also be modelled using syllogisms. We shall explore this possibility and the limitations that it can have and we shall illustrate how our proposal

works with some examples taken from real scenarios; particularly, medical diagnosis and legal argumentation.

7. **Implementing our proposal of syllogistic system in a software library for testing it and making it available for any possible user**. This library allows us to check the behaviour of our model and facilitates its use by any user.

Structure of the thesis

To adequately develop the formulation of our problem and achieve the explained objectives, this thesis is organized as follows:

- **Chapter 1**: We introduce syllogistic reasoning as a model of deductive reasoning based on the use of quantified statements. First of all, we explore the different interpretations about them described in the literature as quantified statements are the propositions managed in this type of reasoning. They are classified into two groups: i) those where the terms designate sets and the quantifier denotes a quantity link between them; and ii) those where the terms constitute the antecedent and consequent of a conditional proposition and the quantifier denotes the strength of the link between them. In the first case, we describe the main characteristics of Aristotelian Syllogistics as the first syllogistic system and assume it as the set of inferences that any further proposal must satisfy. Thus, the collection of syllogistic frameworks described in this chapter are complemented with a critical analysis to check their compatibility with the Aristotelian system. In the second case, we describe three possible models for the conditional interpretation; i.e., as probabilities, as beliefs or as modality. We conclude the chapter with an analysis about which is the best model for satisfying our objective.

- **Chapter 2**: We present the definition of quantifiers according to TGQ and develop the first part of the main objective of this research; i.e., a proposal of syllogistic reasoning that can manage the different quantifiers proposed by TGQ simultaneously. We include some illustrative examples to show how it works and conclude by summarizing its main characteristics, and its strengths and weaknesses.

- **Chapter 3**: The question of approximate middle terms is proposed, including also a brief description of how this problem is addressed in the literature. The first part is a brief description of different environments and strategies for measuring the similarity between two terms. We choose synonymy as the most interesting relationship since it is characteristic of natural language. We adopt a definition based on the use of a *thesaurus*, a dictionary of synonyms, and different mathematical measures of similarity. Over these ideas, we propose an inference pattern, compatible with the one devoted to quantifiers, which allows us to obtain qualified syllogisms according to their reliability or confidence. We also include some illustrative examples and conclude by summing up our main contributions.

- **Chapter 4**: We study the compatibility between our model of syllogism with different types of probabilistic reasoning. We analyse the so-called Generalized Bayes' Theorem, Bayesian networks probabilistic reasoning in legal argumentation. We focus principally on the basic Bayesian inference patterns and how they can be modelled using syllogisms, providing some guidelines. We illustrate the behaviour of our method with an example taken from medicine and other from law. The objective is to show the possibilities of syllogism as an alternative tool, easier and clearer for non-specialized users, to model some probabilistic problems that involve Bayesian reasoning.

- **Conclusions**: This is devoted to explaining the conclusions reached during this research. We summarize the main contributions made during this work, emphasizing its strengths and weaknesses, and also the most relevant open questions that we consider will constitute our future work.

Publications

By way of conclusion, we would like to emphasize that the main results achieved from this research have been described in several papers:

- The analysis about the behaviour of fuzzy approaches involving fuzzy quantified statements to deal with Aristotelian Syllogistics, the classical system of syllogism, was published in *Fuzzy Sets and Systems* [51] concluding that any of them could manage it adequately. A number of initial results were presented at at *IEEE World Congress on Computational Intelligence* (WCCI2010) [50]. Both papers are our starting point to explore the possibility of other models of syllogistic reasoning that involves both systems.

- Our proposal of syllogistic reasoning was explained in detail in *Fuzzy Sets and Systems* [51] and constitutes main heart of the Chapter 2. A number of preliminaries results have been previously presented at the *XVI Congreso Español sobre Tecnologías y Lógica Fuzzy* (ESTYLF 2012) [57].

– The first approach to a syllogism involving approximate middle terms was presented at the *Congreso Español de Lógica, Metodología y Filosofía de la Ciencia en España* [54] and subsequently extended in the paper presented at the *2013 IEEE International Conference on Fuzzy Systems* [55]. Both papers constitute the basis of the proposal developed in chapter 3. A new paper comparing and analysing different procedures for measuring synonymy was submitted to the *XVII Congreso Español sobre Tecnologías y Lógica Fuzzy* (ESTYLF 2014) [53].

– We propose the field of argumentation theory as an adequate environment for the use of syllogistic arguments as it deals with any human situation that involves the intentional acts of persuading or dissuading somebody. The first contact this field was through pragmatics, the field that deals with *utterances*, to analyse the most superficial layer of arguments. This analysis allowed us to collaborate with the research unit of *Cognitive Computing: Computing with Perceptions* of the *European Centre for Soft Computing* writing two papers, [52] and [19].

– In chapter 4, a brief section about the relevance of implicit premises in arguments is included. The links between the concept of intentionality and the use of arguments in an approximated framework were initially considered in [48]. An extended version of this paper was published in *Mathware & Soft Computing* journal [49].

– The first steps of this research also led us to make a number of reflections on the impact and the role of vagueness in fuzzy logic from the point of view of philosophy. They were presented in the *2009 IFSA World Congress - 2009 EUSFLAT Conference* (IFSA-EUSFLAT 2009) [49] and contributed to define the perspective adopted in this research.

Chapter 1

State of the Art in Syllogistic Reasoning

Human beings, in our daily lives, have to deal with quantitative aspects of reality; that is, we need to refer to the number or the amount of elements endowed with a particular feature or which fulfil a property. Natural language, a characteristic capability of *zoon politikon* [5][1], offers us a very useful tool for expressing quantities: linguistic quantifiers. The most common are, perhaps, terms such as "no", "many", "few", etc. that denote proportional values; others like "thirty", "ten", "five", etc. for referring to absolute quantities; "all but five" or "all but three" are useful for designating exceptions, focusing on the number of elements that *do not fulfil* the property; "double" or "triple" are relationships of multiplication, etc. The TGQ [8] identifies all these different types of quantifiers while classical logic only consider two: the universal (usually expressed as "all") and the existential (usually expressed as "some") ones.

Linguistic quantifiers can be used directly in simple statements (i.e., "How old are you? I am *twenty eight* years old"); however, they are mostly used inside in propositions that involve a relationship between two terms or sets: the so-called *binary quantified statements*[2]. For instance, in order to tell somebody about how many students are attending a conference, one has to refer to a quantity, and, hence, use a quantified statement such as "twenty-five students attended the conference" or "almost all students attended the conference". Depending on the

[1] Although the usual translation of this expression is "political animal", the characterization of the human being as a "linguistic animal" is understood by it.

[2] The research presented in this thesis is mainly focused on binary quantifier statements and, for brevity, they will be referred simply as quantified statements.

available information, the quantities can be precisely known (as in the first statement) and referred to using precise quantifiers such as “one hundred”, “twenty five”, “50%”, “no”, etc.; or they can be vague or imprecisely known (as in the second statement), using, in this case, fuzzy quantifiers such as “few”, “many”, “most”, etc.

Regarding the pragmatic dimension of quantified statements, it is relevant to note their expressive wealth for transmitting connotations and a palette of shades that are presenting in any event. For instance, considering again the previous example, instead of using the absolute quantifier “twenty-five” and knowing how many students are in the classroom (i.e.; twenty-eight), one could use “all but three students attended the conference”, thus emphasizing, like this, that the conference caught the attention of almost all students. However, if somebody wants to emphasize that three students did not attend the conference, he could say “some students did not attend the conference” pointing out implicitly, in this way, that the conference may not have been interesting enough for the students.

On the other hand, quantified statements can also be used for making inferences where all the involved statements are quantified; i.e., to deduce a conclusion (a quantified statement) from a set of premises (also quantified statements). The standard form of the quantified statements for this purpose is *Q S are P*, where *Q* stands for the quantifier (either precise or vague), *S* stands for the subject of the statement, *P* stands for its predicate and *are* stands for the corresponding form of the copulative verb *to be*. For instance, the statements “most students attended the conference” or “twenty-five students attended the conference” follow this structure. Let us now consider a very basic example of inference involving quantified statements; i.e., if we say “almost all students attended the conference” we can infer that “very few students did not attend the conference” since the quantifier “almost all” denotes a high quantity, while “very few” denotes the complementary quantity. The most common form of syllogism is composed by two premises and a conclusion and the engine of the inference is the chaining; i.e., it is possible to conclude that two terms (*extreme terms*) are linked because there is a third term that links them: the so-called *middle term*. The usual name for this kind of inference is *syllogism* and the logical system generated by its study is known as *syllogistics*.

Thus, the aim of this chapter is to introduce the topic of syllogistic reasoning and review the state of the art to provide a highly detailed characterization of the different models proposed, both in philosophical and fuzzy logic. We also assume Aristotelian Syllogistics as the referential framework that any further extension of syllogism must include as a particular case. We classify the syllogistic systems analysed into two groups according to their interpretations of quantified statements: i) set-based interpretation (section 1.1); and ii) conditional interpretation (section 1.2). Finally, section 1.3 presents an analysis comparing both interpretations.

1.1 Set-Based Interpretation

Set-Based interpretation assumes that the terms that constitute a quantified statement in its standard form (Q S are P) are sets. A set, adopting an intuitive definition [12], refers to a collection of objects that share a particular characteristic or property. Let us consider the statement “all natural numbers are real numbers”; the term *natural numbers* denotes the set composed by natural numbers; the term *real numbers* denotes the set of real numbers and *all* how the elements of both sets are linked. Hence, any quantified statement denotes a relationship of quantity between sets, the terms that comprise it being *term-sets*. Thus, the subject of the statement is called *subject-term* and its predicate, the *predicate-term*. Furthermore, they constitute precise assertions although they can involve vague quantities or a vague reference to a quantity. Thus, the general rule for interpreting the schema *Q S are P* is that *there is a quantity-relation between S and P denoted by Q*.

This section describes different syllogistic systems that can be found in the literature and it is organized as follows: section 1.1.1 describes Aristotelian Syllogistics, the first syllogistic system in the history of logic; section 1.1.2 analyses the so-called Intermediate Syllogistics, the first extension of Aristotelian Syllogistics; section 1.1.3 characterizes the named Exception Syllogistics, which is based on the exclusive use of exception quantifiers; section 1.1.4 describes a fuzzy syllogistics based on the use of intervals; section 1.1.5 details the main characteristics of another fuzzy syllogistics based on fuzzy numbers and fuzzy arithmetic; and, finally, section 1.1.6 presents a generalization of Intermediate Syllogistics in fuzzy logic.

1.1.1 Aristotelian Syllogistics

Aristotelian Syllogistics [4], developed in the fifth century B.C., is the first approach to syllogistic reasoning and also the first logical system in history. Aristotle defined syllogism as a deductive argument “in which certain things having been supposed, something different from those supposed results of necessity because of their being so”[7, I.2, 24b18-20]. The *things supposed* are the premises of the argument and *results of necessity* are the conclusion. It is worth noting that syllogism has a probative and deductive character because talking about "results of necessity", where the premises are true, prevents the conclusion from being false.

The systematic study of syllogism was developed by Aristotle in the books that comprise the *Organon* [6]. For the analysis of syllogism, the most relevant books are *De Interpretatione*, where Aristotle's theory of categorical statement is described, and *Prior Analytics*, which includes Aristotelian theory of deductive inference.

This section is organized as follows: section 1.1.1.1 addresses the theory of Aristotelian categorical statements; section 1.1.1.2, describes how the classical version of the Logic Square of Opposition is generated; section 1.1.1.3, describes the Aristotelian reasoning patterns and, finally, section 1.1.1.4 concludes with a summary analysing the strengths and weaknesses of this proposal.

1.1.1.1 Theory of Aristotelian Categorical Statements

The theory of categorical statement describes the so-called *categorical statements*, the typical propositions of syllogisms. These are defined as "assertions about classes[3] that affirm or deny that a class is included in other total or partially" [12, p. 168]. As we can see, the concept of *set* is the key point of this framework.

It is worth noting that Aristotle assumes that the logical structure and the syntactical structure of a statement are the same and, hence, the logical terms (subject-term, predicate-term and copulative verb) coincide with the syntactical categories (subject, attribute and copulative verb, respectively). At the beginning of the twentieth century, the new logic developed by G. Frege and B. Russell dismissed this idea. For them, the logic structure of propositions does not coincide with their syntactical structure but corresponds with the mathematical structure of function-argument [29]. It is also relevant to say that every categorical statement is quantified; i.e., it always involves some kind of quantity. Those statements that do not involve quantifiers but singular terms, such as proper names (i.e.; "Arthur Conan Doyle is a writer"), are not categorical statements and, thus, are not valid for making categorical syllogisms.

The properties of categorical statements are as follows [6]:

1. Quantity: It is *universal* if the statement refers to each of the elements that belong to the subject-term and it is *particular* if the statement only refers to a portion of them.

[3] For the context of this research, *class* and *set* are equivalent terms.

Universal quantity is expressed by the quantifiers “all” and “no” and particular quantity by the quantifiers “some” and “some. . . not”.

2. Quality: It is *affirmative* if the assertion affirms a universal or a particular quantity and it is *negative* if it denies the quantity.

3. Distribution: It is a property of the statements respect to the terms that it involves. A statement distributes a term “if it refers to each member of the class designated by that term” [12, p. 168]. As will be shown, this is a key concept for the development of the theory of syllogism.

The combination of these properties generates the four standard categorical statements [6]:

– **All S are P:** This is a *universal affirmative* and means that every member of the subject-term is also a member of the predicate-term; i.e., “all natural numbers are real numbers”.

– **No S is P:** This is a *universal negative* and means that the subject-term and the predicate-term do not share any element; i.e., “no stone is a vegetable”.

– **Some S are P:** This is a *particular affirmative* and means that certain members of the subject-term are also members of the predicate-term; i.e., “some animals are pets”.

– **Some S are not P:** This is a *particular negative* and means that certain members of the subject-term are not members of the predicate-term; i.e., “some animals are not mammals”.

Table 1.1 summarizes the four Aristotelian categorical statements. The third column from the left shows the capital letters commonly used for referring to them. This nomenclature comes from Scholastic philosophers in the Middle Ages and it has a mnemonic purpose, since the letters **A** and **I** stand for affirmative statements (universal and particular ones, respectively) and **E** and **O** for negative ones (universal and particular ones, respectively). **A** and **I** come from the Latin word, “**A**ff**I**rmo”, that means “I affirm” and **E** and **O** come from, “n**E**g**O**”; that means, “I deny”.

The property of distribution [12, 61] arises in each categorical statement from the properties of quantity and quality according to the following description (see Table 1.2):

– **All S are B**: *S* is distributed and *P* is not distributed; i.e., in “all human beings are mortal”, *human beings* is distributed because we know that each of the elements in the

Type of proposition	Form	Capital letter
Universal affirmative	All *S* are *P*	A
Universal negative	No *S* are *P*	E
Particular affirmative	Some *S* are *P*	I
Particular negative	Some *S* are not *P*	O

Table 1.1: Four standard categorical statements.

Statement	Subject-term	Predicate-term
A	Distributed	Undistributed
E	Distributed	Distributed
I	Undistributed	Undistributed
O	Undistributed	Distributed

Table 1.2: Distribution in categorical statements.

set *human beings* belongs to the set of *mortal beings*; however, we do not have the same information regarding the elements of the predicate-term.

– **No S is B**: *S* and *P* are distributed; i.e., "no stone is a vegetable" classifies every element of both sets because no element in the set *stones* belongs to the set *vegetables* and vice versa.

– **Some S is B**: *S* and *P* are not distributed; i.e., "some animals are pets", neither the elements of *animals* nor the elements of *pets* are classified.

– **Some S is not B**: *S* is not distributed and *P* is distributed; i.e., "some animals are not mammals" classifies every element of the set *mammals* as all of them belong to the set *animals* but in this term-set there are unclassified elements.

1.1.1.2 Logic Square of Opposition

Aristotle's Syllogistics is a binary logic, and therefore, only two truth-values are managed: truth and falseness. So, each one of the four categorical statements can be true or false, which generates a determined set of dependences among them. These are the so called *relationships of opposition* [12, p. 214-216]:

– **Contradictory (A-O, E-I):** Two propositions are contradictory if one is the denial or negation of the other; whenever one is true, the other is false and vice versa. For

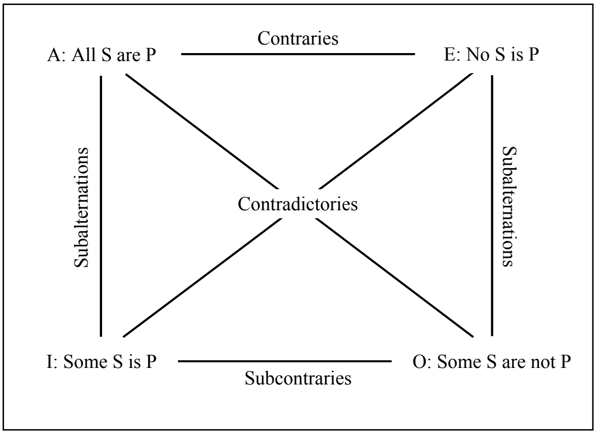

Figure 1.1: Classical LSO.

instance, if “all mammals are animals” is true, then “some mammals are not animals” is false.

– **Contrariety (A-E):** Two propositions are contrary if they cannot both be simultaneously true; i.e., the truth of either of them entails the falseness of the other; however, both can be false. For instance, “all animals are mammals” and “no animal is a mammal” are simultaneously false.

– **Subcontrariety (I-O):** Two propositions are subcontrary if they cannot both be false, although they may both be true. For instance, “some animals are mammals” and “some animals are not mammals”.

– **Subalternation (A-I, E-O):** Two propositions with the same subject and predicate-terms and quality (affirmative or negative); the truthfulness of the the universal one -named *superaltern*- entails the truthfulness of the particular one -named *subaltern*. For instance, “no cat is a stone” entails “some cat is not a stone”.

In the Middle Ages, scholastic philosophers proposed a graphical representation of categorical statements according to these properties. The result, called the *Logic Square of Opposition* (LSO), is shown in Figure 1.1.

The LSO allows us to generate some simple inferences known as *immediate inferences* as they are constituted by a single premise and two terms. They do not belong to the general

If	**A**	is true, then	**E** and **O** are false, while **I** is true.
If	**E**	is true, then	**A** and **I** are false, while **O** is true.
If	**I**	is true, then	**A** and **O** are undetermined, while **E** is false
If	**O**	is true, then	**A** is false, while **E** and **I** are undetermined.
If	**A**	is false, then	**O** is true, while **E** and **I** are undetermined.
If	**E**	is false, then	**I** is true, while **A** and **O** are undetermined.
If	**I**	is false, then	**A** is false, while **E** and **O** are true.
If	**O**	is false, then	**A** and **I** are true, while **E** is false.

Table 1.3: Basic immediate inferences.

Convertend	**Converse**
All *S* is *P*	Some *P* is *S* (by limitation)
No *S* is *P*	No *P* is *S*
Some *S* is *P*	Some *P* is *S*
Some *S* is not *P*	Conversion invalid

Table 1.4: Conversion inferences.

theory of syllogism but they play a fundamental role in its adequate development since they are used by Aristotle for proving the validity of his inference schemas. Two types of immediate inferences can be distinguished [12]: i) basic ones (see Table 1.3) and ii) derived ones, which involve more than one property of the LSO. The latter are:

– **Conversion (E,I,A):** This allows us to obtain equivalent statements by interchanging their terms. **E** and **I** are *simply convertible*; i.e., subject-term and predicate-term can be interchanged with no restriction. For instance, “no vegetable is a stone” is equivalent to “no stone is a vegetable”, as this quantity relationship is based on the intersection operation, which is commutative. **A** statements are not simply convertible as the inclusion operation is not commutative. However, in this case we can say that **A** statements are convertible with restrictions because it is true that if “all animals are mammals” is true, then “some animals are mammals” is also true; i.e., the conversion of the subaltern. Finally, the conversion of **O** statements is not possible as the statement and its converted are completely independent. Table 1.4 summarizes this inference.

– **Obversion:** This allows us to obtain statements that are logically equivalent through changing only the quality of the statement and substituting the predicate-term by its complementary. Thus, for instance, the obverted form “no vegetable is a stone” is

Obverted	**Obverse**
All *S* is *P*	No *S* is *non* − *P* (by limitation)
No *S* is *P*	All *S* is *non* − *P*
Some *S* is *P*	Some *S* is not *non* − *P*
Some *S* is not *P*	Some *S* is *non* − *P*

Table 1.5: Obversion inferences.

Premiss	**Contrapositive**
A: All *S* is *P*	**A**: All *non* − *P* is *non* − *S*
E: No *S* is *P*	**O**: Some *non* − *P* is not *non* − *S*(by limitation)
I: Some *S* is *P*	Contraposition is invalid
O: Some *S* is not *P*	**O**: Some *non* − *P* is not *non* − *S*

Table 1.6: Contraposition inferences.

"all vegetables are no-stones". As we can see, the quality of the statement changes (it is affirmative instead of negative) and the predicate term (*stone*) is substituted by its complementary (*no-stone*). Table 1.5 summarizes this inference.

– **Contraposition:** This allows us to obtain statements that are logically equivalent by using the complementary terms of the subject and the predicate-terms interchanging its positions. Thus, for instance, the contrapositive of "all mammals are animals" is "all no-animals are no-vegetables". Table 1.6 summarizes this inference.

This version of the LSO shows the so-called problem of *existential import* [58]. A statement has existential content if it affirms that the sets denoted by each one of the involved terms have at least one member; in other words, all the designated sets are not empty [12]. **I** and **O** have existential import because their assertion necessarily entails the existence of elements in the involved sets. For instance, the truthfulness of "some animals are mammals" entails that there is, at least, one element in the set *animals* that is also a member of the set *mammals*. Let us consider now the case of **A** and **E** statements. For instance, "all mammals are animals" is true because we know that every member of set *mammals* belongs to the set *animals*; now, if we say "all Martian mammals are animals" is this true or false? From a strict logical point of view, it cannot be false because every member of the subject-term *Martian mammals* ($\emptyset$) belongs to the predicate-term *animals* ($\emptyset$ is subset of every set); hence, it is true. Something similar happens with **E** statements. Thus, if we accept this interpretation for the **A** and **O** statements (not involving the existential import), the subalternation property is not valid given the existential meaning of **I** and **O**. For this reason, the most accepted interpretation within the Aristotelian LSO is that **A** and **E** statements have existential import.

The debate on existential import has been raging in Philosophy since Aristotle's day and was very important during the Middle Ages [29]. Nevertheless, the most accepted interpretation of this question today is that the existential import is not always assumed in ordinary language and, sometimes, we reason without assuming it [29, 58].

An alternative version of the LSO, according to this last idea, is proposed by S. Peters and D. Westersthål [58]. Figure 1.2 shows the so-called Modern LSO. The quantifier "some... not" is substituted by "not all" and the relationships expressed in the LSO are modified because the concept of 'opposition' is substituted by the concept 'negation' and, hence, the concept of truth ceases to be the core of the LSO. Negation can adopt three forms:

Outer negation This is the equivalent of the contradictory property in the Aristotelian LSO. The outer negation of *Q S are P* is *not (Q S are P)*; for instance, the outer negation of "all mammals are animals" is "not all mammals are animals". In the Modern LSO, the outer negation is in the diagonals.

Inner negation This is the equivalent of the contrariety and subcontrariety properties of the classical LSO. Hence, the inner negation of *Q S are P* is *Q S are not P*; for instance, the inner negation of "some animals are mammals" is "some animals are not mammals". In the Modern LSO, the inner negation is in the horizontal line.

Dual This is the combination of the outer and inner negation and it is the equivalent to the obversion property. Hence, dual of *Q S are P* is *not (Q S are not P)*; for instance, the dual of "no vegetable is a stone" is "not all vegetables are stones". It is worth noting that a sentence and its dual are logically equivalent; i.e., if one of them is true or false the other is also true or false.

1.1.1.3 Theory of Aristotelian Syllogism

As we have already stated, the basic inference schema of Aristotelian syllogistics is the *syllogism*. It is defined as a piece of deductive reasoning in which two general terms are linked by a middle term [7, I, 23 (40b16)] and requires three terms, two premises and a conclusion. This

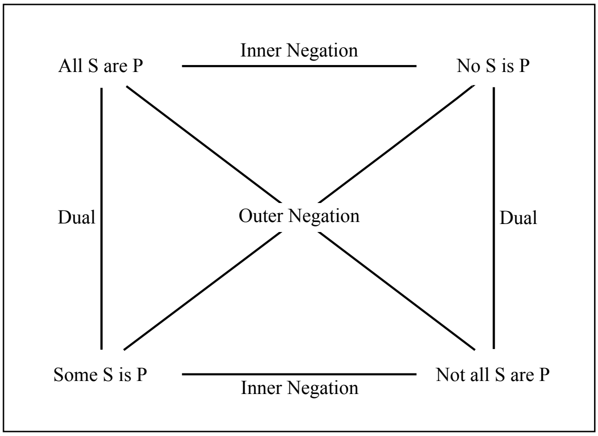

Figure 1.2: Modern LSO.

(a) Aristotelian syllogistic inference.	(b) Inference schema.	
All human beings are mortal	**MP**	Q_1 *DT* are *MT*
All Greeks are human beings	**NP**	Q_2 *NT* are *DT*
All Greeks are mortal	**C**	Q_3 *NT* are *MT*

Table 1.7: Example of Aristotelian syllogism.

defines the syllogism as a form of mediate inference, in opposition to immediate inferences explained above, which only included two terms and a single premise.

The term shared by the premises is the so-called *Middle Term* (DT) and the other two are the so-called *Extremes*, where the subject-term of the conclusion is the *Minor Term* (NT) and the predicate-term is the *Major Term* (MT). The most famous example of Aristotelian syllogism is shown in Table 1.7(a). This example follows the inference schema shown in Table 1.7(b), where Q_1, Q_2 and Q_3 are crisp quantifiers ("all","some", "no" or "some. . . not"). The *Major Premise* (MP) is the premise in which the *MT* appears and the *Minor Premise* (NP) in which the *NT* appears.

As we have previously mentioned, the DT only appears in the premises although it can be the subject-term or the predicate-term. For instance, in the schema of Table 1.7(b) it is the subject-term in MP and the predicate-term in NP. The different positions of DT in the premises generates the so-called four Aristotelian *Figures* (see Table 1.8).

Figure I	**Figure II**	**Figure III**	**Figure IV**
Q_1 DT are MT	Q_1 MT are DT	Q_1 DT are MT	Q_1 MT are DT
Q_2 NT are DT	Q_2 NT are DT	Q_2 DT are NT	Q_2 DT are NT
Q_3 NT are MT	Q_3 NT are MT	Q_3 NT are MT	Q_3 NT are MT

Table 1.8: Figures of Aristotelian Syllogistics.

Figure I	**Figure II**	**Figure III**	**Figure IV**
AAA (Barbara)	EAE (Cesare)	AAI (Darapti)	AAI (Bramantip)
AIE (Celarent)	AEE (Camestres)	IAI (Disamis)	AEE (Camenes)
AII (Darii)	EIO (Festino)	AII (Datisi)	IAI (Dimaris)
EIO (Ferio)	AOO (Baroco)	EAO (Felapton)	EAO (Fesapo)
AAI (Barbari)	EAO (Cesare)	OAO (Bocardo)	EIO (Fresison)
EAO (Celaront)	AEO (Camestrop)	EIO (Ferison)	AEO (Camenos)

Table 1.9: Moods of Aristotelian Syllogistics.

(MP)	No vegetable is a stone
(NP)	Some vegetables are green
(C)	Some stones are not green

Table 1.10: *Ferison* syllogism.

Since there are four Figures and four possibilities for each one of the three premises (the four categorical statements of the LSO), there are $4 \times 4^3 = 256$ possible combinations of inference patterns. Nevertheless, only 24 are valid: the so-called *Moods* (see Table 1.9). The *Moods* are labelled with three letters, the first capital letter refers to the MP, the second one to the NP and the last one to the C. As an example, syllogism of the Table 1.7(a) is of the type *AAA (Barbara)* and it belongs to Figure I. Another example is *Ferison (EIO)* Mood (see Table 1.10), which belongs to Figure III (the DT is the subject-term in both premises) and its MP is an **A** statement and the NP and C are **I** statements.

Aristotle handles different ways for testing the validity of a syllogism: proofs, counterexamples and rules.

Regarding the proofs, Aristotle distinguishes between *perfect proofs* and *imperfect proofs*. Perfect proofs, or perfect syllogisms, are in deductions that "need no external term in order to show the necessary result" [7, 24b23-24]; imperfect proofs, or imperfect syllogisms, are in deductions that "need one or several in addition that are necessary because of the terms

(a) Invalid syllogism.	(b) Counterexample.
All roses are flowers	All human beings are mortal
Some flowers are red	Some mortal beings are dogs
Some roses are red	Some human beings are dogs

Table 1.11: Aristotelian counterexample.

supposed but were not assumed through premises" [7, 24b24-25]. In [33], J. Łukasiewicz, adopting a contemporary point of view, considers that those syllogisms that only need perfect proofs are the primitive rules or axioms of the system, while the others are derived from them.

Regarding the use of counterexamples, it is a well known procedure in first-order logic. The procedure is simple: following the same inference schema for the syllogism which is under evaluation, trying to build an example with true premises and a false conclusion. Table 1.11 shows an example.

The last of the Aristotelian ways of determining the validity of a syllogism is to check whether it violates any of the rules of the syllogism. If all of them are satisfied, the syllogism is valid. The syllogistic rules are the following [7]:

- **Distribution**

  R1 The middle term is distributed as least once.

  R2 No term is distributed in the conclusion unless it is distributed in one premise.

- **Quality**

  R3 At least one premise is affirmative.

  R4 The conclusion is negative if and only if one of the premises is negative.

- **Quantity**

  R5 At least one premise is universal.

  R6 If either premise is particular, the conclusion is particular.

Apart of the Aristotelian ways of proving the validity of syllogisms, in the nineteenth century J. Venn [85] developed a graphical procedure, known as Venn Diagrams, for representing categorical statements and categorical syllogisms. It was subsequently improved by

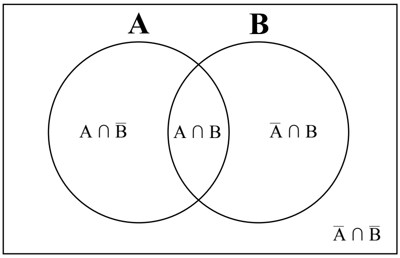

Figure 1.3: Primary diagram of the term-sets A and B.

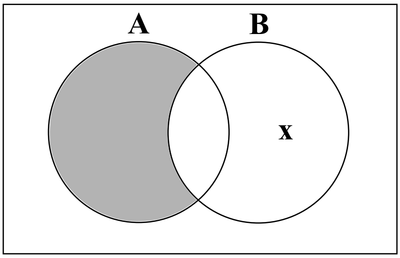

Figure 1.4: Venn-Peirce diagram for "All A are B and Some B are not A".

C. S. Peirce [47]. Its main concept is the *primary diagram*, which represents all the possible set-theoretic relationships between the sets involved in a categorical statement identifying the four possible quantities referred to by the quantifiers; i.e., $A \cap B$, $\bar{A} \cap \bar{B}$, $\bar{A} \cap B$ and $A \cap \bar{B}$ (see Figure 1.3). Only two cardinalities can be represented: empty sets, denoted by shadow areas, and sets with at least one element, denoted by 'x'. Figure 1.4 shows the representation for "all A are B and Some B are not A".

Adding an additional set intersected with the standard two, categorical syllogisms also can be represented. For instance, Figure 1.5 shows the *Ferio (EIO)* Mood of Figure I (see 1.5(a)) and its corresponding representation using Venn-Peirce diagrams (see 1.5(b)).

1.1.1.4 Summary

Aristotle focused on a type of deductive reasoning, syllogism, which deals with a very specific class of propositions, the so-called categorical statements, which are quantified. He distinguishes between four types according to the properties of quantity (universal or particular), quality (affirmative or negative) and distribution (distributed and not distributed). These

(a) *Ferio* Syllogism.

(MP)	No DT is MT
(NP)	Some NT is DT
(C)	Some NT is not MT

(b) Diagram.

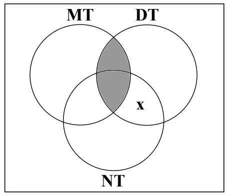

Figure 1.5: *Ferio* syllogism and its representation using Venn-Peirce diagrams.

properties generate a set of dependences among them using the concept of opposition that constitute the LSO, one of the principal contributions of Aristotelian Syllogistics to the study of natural language and quantification, and the basis for his theory of deductive arguments: syllogisms.

Syllogisms are arguments which comprise three terms, two premises and a conclusion. Their validity is supported by the chaining of the extreme terms, which appear in the conclusion, through a middle term, that only appears in the premises. Attending to these constraints, twenty-four valid inferences schemas are obtained: the so-called *Moods*.

Aristotle only dealt with a very restricted set of quantifiers and a very specific kind of arguments. Hence, the capability of this approach for managing many of the arguments expressed in natural language is very limited as many of them involve other quantifiers and vagueness in its formulation. Nevertheless, Aristotelian studies establish the basis for the study of reasoning involving quantities, preserving its natural language form, and we shall adopt it as the referential framework for subsequently analysing syllogistic systems.

1.1.2 Intermediate Syllogistics

During the Middle Ages, one of the main topics of Scholastic Philosophy was Aristotelian Syllogistics, making a number of relevant contributions, such as the LSO. However, it is in the twentieth century when the first proposals for increasing the quantities dealt with in a syllogism appear: P. L. Peterson [60, 59, 61, 62, 63] and B. Thompson [81, 82] introduce the concept of *Intermediate Quantifier*, which extends syllogistic inferences beyond the four Aristotelian crisp quantifiers. In the following sections, we shall describe this concept, how it is built and the valid syllogisms that can be generated using them.

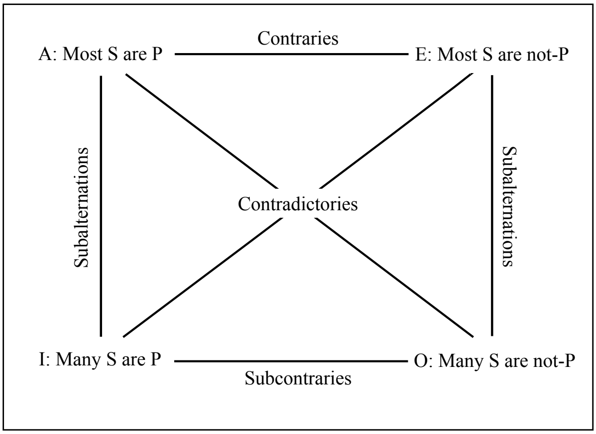

Figure 1.6: LSO with “most” and “many”.

This section is organized as follows: section 1.1.2.1 addresses the theory of intermediate categorical statements and the new LSO generated; section 1.1.2.2 describes the reasoning patterns and, finally, section 1.1.2.3 concludes with a summary analysing the strengths and drawbacks of this proposal.

1.1.2.1 Theory of Intermediate Categorical Statements

The notion of *intermediate categorical statement* is an extension of the Aristotelian categorical statement incorporating the concept *intermediate quantifier*. It is defined as any quantity/quantifier between the universal (“all”, “no”) and the particular (“some”, “some... not”) quantifiers and consistent with the LSO [63]. This means that all the relationships defined in the Aristotelian LSO (contrary, subcontrary, contradictory and subalternation; see 1.1.1.2) are preserved. However, in natural language there a lot of quantifiers that can appear in the space between the universal and particular quantities and, hence, there are multiple candidates to be incorporated. In [60], three possibilities for the combination of the quantifiers “few”, “many” and “most” are considered : i) “few” and “many”; ii) “few” and “most”; iii) “most” and “many”. We shall focus only on option iii) because, as will be explained, it shows the best behaviour.

Figure 1.6 shows the LSO built with “most” and “many” (Most-Many LSO), where all the typical properties of a LSO can be observed. If, furthermore, the relationship shown in

Categorical statements	
Few S are P	Most S are not-P
Few S are not-P	Most S are P

Table 1.12: Equivalence between "few" and "most".

Table 1.12 is also assumed, where the equivalence between categorical statements involving "few" and "most" is established, option ii) (a Most-Many LSO) and option i) (a Few-Most LSO) can be generated; therefore, the Most-Many LSO integrates the other two. This is a heuristic conclusion that was achieved after comparing the three generated LSO. In addition, the are other two reasons that make the Most-Many LSO preferable [60]:

- It emulates the Aristotelian LSO exactly: the affirmative statements are on the left-hand side (A and I) and the negative ones are on the right (E and O).
- Most-Many LSO and Aristotelian LSO can be integrated into a single LSO, preserving all the properties and attending to the following order [60, p. 168]:

$$\text{All S are P} \rightarrow \text{Most S are P} \rightarrow \text{Many S are P} \rightarrow \text{Some S are P} \tag{1.1}$$

$$\text{No S is P} \rightarrow \text{Most S are not-P} \rightarrow \text{Many S are not-P} \rightarrow \text{Some S are not-P} \tag{1.2}$$

A LSO built with intermediate quantifiers, in addition to the existential import (see section 1.1.1.2; it is solved in the same way; i.e., subjects-terms cannot be empty sets), exhibits other problematic questions: i) multiple quantities for the same quantifier [60, 81]; and ii) a constant reference class [60].

Question i) arises owing to the vague and context-dependent character of intermediate quantifiers. Let us consider the following statement, "most students are dedicated"; a possible definition for "most" is that the number of students that are dedicated is greater than the number of students that are not dedicated. This definition opens a wide range of possibilities in the effective interpretation of "most" or any other intermediate quantifier. Thus, in [60, 81], three possible senses for delimiting this meaning are proposed:

- Minimal sense: This means *at least* or *no less than* the quantity named.
- Maximal sense: This means *only* or *no more than*.
- Exact sense: This is the combination of the minimal and the maximal senses; i.e., *no more nor less than*.

The preservation of the subalternation property requires a particular combination of these senses. Thus, "some", "many" and "most" are interpreted in the minimal sense and "few" in the maximal one, given that their minimal sense is too close to the usual interpretation for "some" [60, 81].

Question ii), the so-called constant reference to a class, is also characteristic of an LSO with intermediate quantifiers. It supports the two aforementioned assumptions and refers to the possible readings of a quantifier; i.e., *absolute* or *proportional*. In the absolute one, the quantifier refers to the cardinality of the relationships involved in a proposition saying that it is *at least n*, where *n* is a quantity value. In the proportional reading, the quantifier refers to a relative proportion of the sets involved in the corresponding proposition saying *at least k*, where *k* is a fraction between 0 and 1 or a percentage [44]. Let us consider the following example "many students are athletes" in the context of the University of Santiago de Compostela (USC) and in the academic year 2011/2012: $28,438$. We can say that the proposition is true using the absolute reading if $4,000$ or $5,000$ students play sports regularly, as $4,000$ or $5,000$ students constitute a lot of people; however, if we adopt the proportional reading, this quantity is around 15% and it is more difficult to accept that "many students are athletes" is true, given that the usual *k* value for "many" is over 50% [76]. A deep analysis of this question about the quantifier "many" can be found in [44, 21].

Coming back to the LSO, maintaining the same reference is fundamental for preserving the properties as, otherwise, they are not fulfilled. For instance, attending to Most-Many LSO (see Figure 1.6 and Table 1.12), from *most S are P* we can infer *few S are not-P* whenever we maintain the proportional reading in both statements. If we use the proportional reading in "most" and the absolute one in "few", then this inference is not valid. For instance, considering again the students of the USC, if "most students are athletes" means "around 70% of students are playing sports regularly", there are around $8,000$ of them that do not play sports regurlaly, and, hence, we cannot say that $8,000$ students are "few students".

To increase the expressive capacity of the LSO, Peterson merges the Aristotelian LSO with the Most-Many LSO and adds two additional quantifiers: "almost all" and "few". This generates a 5-quantity square [59, 63] as involves the five quantifiers "all", "almost-all", "most", "many" and "some" preserving all the standard relationships (see Figure 1.7). The new in-

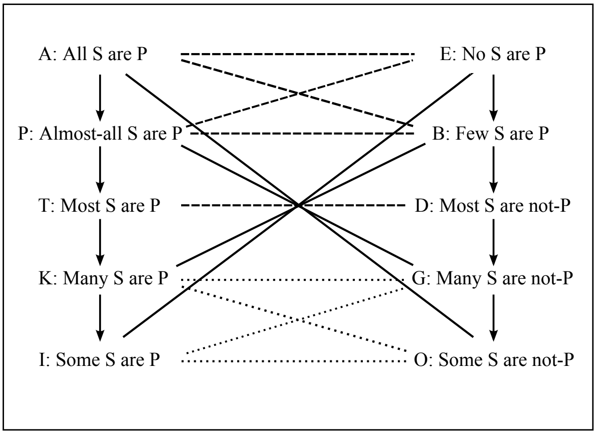

Figure 1.7: 5-Quantity LSO.

termediate categorical statements are named according to the following labels, following the mnemonic medieval spirit[4]:

- Predominant: *almost-all S are P* (P) and and *few S are P* (B)
- Majority: *most S are P* (T) and *most S are not P* (D)
- Common: *many S are P* (K) and *many S are not P* (G)

It is clear that the aforementioned quantifiers do not exhaust the space between the universal and particular ones. Thus, given that there is a procedure for generating LSOs, Peterson [63, p. 233] proposes to use fractions to generalize the concept of intermediate quantifier. In this way, any intermediate quantifier can be expressed by a fraction m/n, being $n \geq m$, and the symbols $>, \geq, <, \leq$ for selecting the corresponding sense of the quantifier. For instance, $> 2/3$ *the S are P* denotes the intermediate quantifier "more than two thirds of S are P". The general schema for any fractional quantifiers LSO is shown in Figure 1.8. It is worth noting that only proportional readings of quantifiers can be used, and, therefore, any consideration of absolute ones lies outside this framework.

[4]P, T and K are consonant letters contrasting with the vowels of Aristotelian LSO. Furthermore, they are widespread unvoiced consonantal stops which point of articulation moving from front to back. The letters for the negative statements, B, D and G are their voiced counterparts.

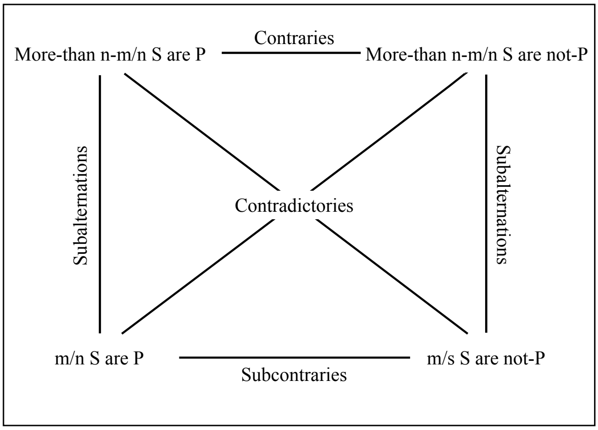

Figure 1.8: General schema for generating the LSO for fractional quantifiers.

1.1.2.2 Theory of Intermediate Syllogisms

P. Peterson and R. Carnes build their doctrine of intermediate syllogism over the 5-quantity LSO (see Figure 1.7) [63, chap. 2]. Thus, any intermediate syllogism can be obtained by "taking any traditional syllogism (valid or not) and replace one or more of its categorical propositions with some intermediate categorical propositions" [63].

The syllogistic system involving the 5-quantity LSO can generate around $4,000$ different inference schemas but only 105 are valid, these being the 24 the Aristotelian ones and 81 new ones. Table 1.13 shows the list of the new valid syllogisms, adhering to the capital letter notation explained above.

The validity of these new syllogisms is proved through an extension of Venn Diagrams [63] inspired by P. Geach's work [22], which consists of an algebraic method. The first step is the division of Venn Diagram of a syllogism into sub-classes, as shown in Figure 1.9. Each lower case letter $(a, b, \ldots, h)$ denotes a subsection of the Venn Diagram and the capital letters $(NT, DT$ and $MT)$ denotes the terms that configure a syllogism.

The heart of this method is to interpret each intermediate categorical statement as making a statement about the relative size of two sub-classes. Table 1.14 shows all the algebraic interpretations of the intermediate categorical statements for a 5-quantity LSO, where S stands for NT and P for MT [59] and $>$ is read as *exceeds* and $>>$ as *greatly exceeds*. We shall

Figure I	**Figure II**	**Figure III**	**Figure IV**
syllogisms containing majorities			
AAT	AED	ATI	AED
ATT	ADD	ETO	ETO
ATI	ADO	TAI	TAI
EAD	EAD	DAO	
ETD	ETD	TTI	
ETO	ETO	DTO	
syllogisms containing majorities and commons			
AAK	AEG	AKI	AEG
ATK	ADG	EKO	EKO
AKI	AGO	KAI	KAI
AKK	AGG	GAO	
EAG	EAG		
ETG	ETG		
EKO	EKO		
EKG	EKG		
syllogisms containing majorities, commons and predominants			
AAP	AEB	PAI	AEB
APP	ABB	EPO	PAI
APT	ABD	BAO	EPO
APK	ABG	GAO	
API	ABO	API	
EAB	EAB	PPI	
EPB	EPB	TPI	
EPD	EPD	KPI	
EPG	EPG	PTI	
EPO	EPO	PKI	
		BPO	
		DPO	
		GPO	
		BTO	
		BKO	

Table 1.13: Moods of Intermediate Syllogistics.

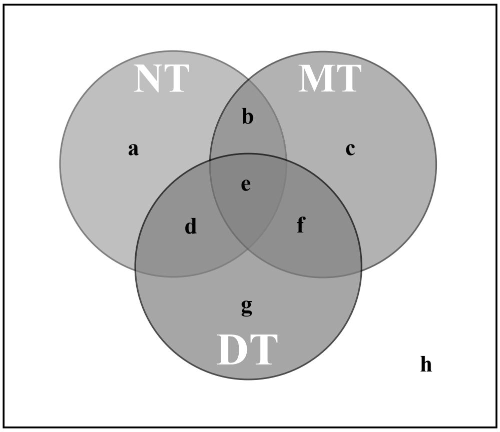

Figure 1.9: Venn Diagram subsections.

Intermediate Proposition	Algebraic Interpretation
all S are P	$a = 0, d = 0$ with $b \neq 0$ or $e \neq 0$
almost all S are P	$b + e >> a + d$ with $b \neq 0$ or $e \neq 0$
most S are P	$b + e > a + d$ with $b \neq 0$ or $e \neq 0$
many S are P	$\neg(a + d >> b + e)$ with $b \neq 0$ or e $\neq 0$
some S are P	$b \neq 0$ or $e \neq 0$
no S are P	$b = 0$ and $e = 0$
few S are P	$a + d >> b + e$ with $b \neq 0$ or $e \neq 0$
many S are not-P	$(a + d >> b + e)$ with $b \neq 0$ or $e \neq 0$
some S are not-P	$a \neq 0$ or $d \neq 0$

Table 1.14: Algebraic interpretation of intermediate categorical statements.

use this same version of Venn diagrams to develop our proposal of syllogistic system (see Chapter 2).

To demonstrate the validity of a syllogistic argument two alternative paths can be taken: i) deduce the algebraic representation of the conclusion from the premises; ii) show that the premises and the denial of the conclusion are inconsistent. For instance, Table 1.15 shows an example of intermediate syllogism proved by path ii).

The Aristotelian meta-rules (see section 1.1.1.3) are still valid for intermediate syllogisms, but they must be extended in order to involve the new schemas. The new rules are mainly focused on the notion the of distribution as, while in Aristotelian syllogistics there are only two possible choices (distributed and undistributed), new possibilities appear in intermediate

TTI Syllogism	
Most DT are MT	$e+f>d+g$
Most DT are NT	$d+e>g+f$
Some NT are MT	$b\neq 0$ or $e\neq 0$
Proof	
$(1) e+f>d+g$	premise
$(2) d+e>g+f$	premise
$(3) b=0$ and $e=0$	denial of conclusion
$(4) f>d+g$	from (1) and (3)
$(5) f>d$	from (4)
$(6) d>f$	from (5) and (3)
(5) and (6) are contradictory	

Table 1.15: TTI syllogism and proof.

syllogisms; in other words, the crisp Aristotelian concept is converted into a graduated one. Peterson proposes the concept of '*more nearly distributed*' expressed as *Distribution Index* (DI) with the following scale:

$DI=5$ Subjects of universals and predicates of negatives.

$DI=4$ Subjects of P and B.

$DI=3$ Subjects of T and D.

$DI=2$ Subjects of K and G.

$DI=1$ Subject of particulars and predicates of affirmatives.

Following the numerical indices, the order from the distributed terms ($DI=5$) to the undistributed terms ($DI=1$) is as follows:

$$DI=5>DI=4>DI=3>DI=2>DI=1$$

Thus, the new rules for intermediate syllogisms are the following ones:

– **Distribution**

R1 In a valid syllogism, the sum of DIs for the middle term must exceed 5.

R2 No term may be more closely distributed in the conclusion than it is in the premises (i.e., no term may bear a higher DI in the conclusion than it bears in the premises).

- **Quality**

  R3 At least one premise must be affirmative.

  R4 The conclusion is negative if and only if one of the premises is negative.

- **Quantity**

  R5 At least one premise must have a quantity of majority (T or D) or higher.

  R6 If any premise is non-universal, then the conclusion must have a quantity that is less than or equal to the premise.

1.1.2.3 Summary

Intermediate Syllogistics incorporates new quantifiers between the universal and particular ones in syllogistic reasoning, many of them being vague. This entails a relevant innovation with respect to the Aristotelian one; which only managed four crisp quantifiers. It also shows that these new quantifiers also can be ordered into an LSO, although some additional constraints must be taken into account, such as, for instance, identifying the adequate sense in the use of the quantifiers. The authors make their proposal of syllogistic system using an LSO with 5 quantifiers.

The incorporated innovations enable Peterson and Thompson to generate a vast number of new syllogisms that contributes to expand the expressive power of the Aristotelian Syllogistics. However, it still limited to proportional quantifiers (others, such as exception or absolute quantifiers, are not considered) and arguments constituted by three terms, two premises and a conclusion. Many arguments involving quantities that human beings use in their daily lives are far beyond the scope of this framework.

It is worth noting that the algebraic method based on Venn Diagrams will be used as the base for our proposal of syllogistic system compatible with TGQ.

1.1.3 Exception Syllogistics

As we have already stated, in natural language there are other kinds of quantifiers different to proportional ones, such as exception ("all but eight", etc.) or comparative ("double", "three more", etc.), which human beings also use in their inferences, even though they are not considered in syllogistics. We have shown that Intermediate Syllogistics significantly increases

Traditional formulation	Exception formulation	Symbolization
All S are P	All but 0 S are P	0SAP
No S are P	At most 0 S are P	0SEP
Some S are P	At least 1 S is P	1SIP
Some S are not P	At least 1 S is not P	1SOP

Table 1.16: Recasting A, E, I and O as exception statements.

the number of quantifiers of Aristotelian Syllogsitics, but considering exclusively proportional ones, such as “most”, “many”, “few”, etc.

In [39], W. Murphree proposes a new approach to syllogistic reasoning based exclusively on exception quantifiers (i.e.; “all but three”, “all but five”, etc.) and on categorical statements with the basic form *all but x S are P*, where x stands for a $\mathbb{N}$. In this way, the proportional reading of the quantifiers is substituted by an absolute one, as it involves a direct reference to an absolute value. Consequently, new schemas of syllogism emerge, one of the principle novelties of which is the consideration of several possible conclusions instead of one single one.

This section is organized as follows: section 1.1.3.1 addresses the theory of exception categorical statements, describing how quantified statements are interpreted in terms of exception quantifiers; section 1.1.3.2 describes the characteristics of the LSO; section 1.1.3.3, describes the reasoning patterns and the procedure for inferring the conclusion and, finally, section 1.1.3.4 concludes with a summary analysing the strengths and weaknesses of this framework.

1.1.3.1 Theory of Exception Categorical Statements

W. Murphree assumes that a categorical statement can be expressed using exception quantifiers, whose basic form is *all but x*, with usually $x \in \mathbb{N}$. For instance, “all human beings are mortal” is equivalent to “all but 0 human beings are mortal”, where the numerical value 0 indicates the “exception” of the quantifier. However, depending on the quantifier under consideration and its associated presuppositions, the basic form of exception quantifiers can be modified. Table 1.16 shows the categorical exception statements equivalent to the four traditional categorical statements (i.e.; A, E, I and O). For instance, “some” is interpreted as “at least one” as it entails the existential import.

Exception formulation	**Symbolization**
At least all but *x* S are P	*x*SAP
At most *x* S are P	*x*SEP
At least *x* S is P	*x*SIP
At leas *x* S is not P	*x*SOP

Table 1.17: General form of exception statements.

Intermediate quantifiers also can be generated using this form, by modifying the value of x, which goes from 0 to $|S|$. Table 1.17 shows the general form exception statements for intermediate quantifiers depending on the universal and particular ones. It is worth noting that this formulation can require the assignation of a cardinality to S as, otherwise, the transformation cannot be properly executed. For instance, let us consider "few students are tall". If we assume $|S| = 10$ and that "few" means "no more than three", we can use $xSEP$ symbolization and say "at most three students are tall".

On the other hand, the introduction of the numerical value x allows Murphree to talk about statements in terms of strength and weakness. Thus, statements with "the same quantity and quality are stronger or weaker than another according to their exception values" [39, p.51]. For universal statements ($xSAP$ and $xSEP$), if the exception value increases, then the claim is weaker; in the case of particular ones ($sSIP$ and $xSOP$), if the exception value increases, then the claim is stronger. For instance, "at most ten students are tall" is stronger than "at most five students are tall" or "at least five dogs are barking" is weaker than "at least ten dogs are barking".

Regarding the distribution property, this does not have modifications with respect to Intermediate Syllogistics (see Table 1.2).

1.1.3.2 Logic Square of Opposition

The introduction of new numerical values in the quantifiers entails a new way of building the LSO. All its relationships are preserved, although Murphree focuses on the so-called equivalent relationships: conversion, obversion and contraposition. Table 1.18 shows the relationships among these properties. Subalternation is explained in terms of strength and weakness; the strongest statement entails the weakest one.

Nevertheless, the problem of existential import is not avoided with categorical statements of this type. Although Murphree [39, p.52] explicitly considers this question, he rejects it

Original	**Conversion**	**Obversion**	**Contraposition**
xSAP	xPAS	xSE$\overline{P}$	$x\overline{P}$A$\overline{S}$
xSEP	xPES	xSA$\overline{P}$	$x\overline{P}$E$\overline{S}$
xSIP	xPIS	xSO$\overline{P}$	$x\overline{P}$I$\overline{S}$
xSOP	xPOS	xSI$\overline{P}$	$x\overline{P}$O$\overline{S}$

Table 1.18: Equivalent relationships between exception statements.

because, in general, the subject-term of universal statements is not necessarily not empty. The generalization of this idea is defined by two observations:

1. A statement is inherently existential to the magnitude of x if and only if the existence of at least x S is a precondition for the truth of the statements.

2. The existence of at least x S is a precondition for the truth of xSIP and xSOP, but xSAP and xSEP are compatible with the existence of any number of S whatsoever.

For instance, a statement with the form *at least x S are P* entails that $|S| > 0$ to be true; while, *at most x S are P* only entails $|S| < x$ which is compatible with $|S| = 0$. Nevertheless, the meaning of exception value x demands the existence of a number of elements in S and, hence, a number of presuppositions appear. Murprhee distinguishes between *minimum presuppositions* and *maximum presuppositions*.

Minimum presuppositions belong to particular statements and are expressed as *at least x*. For instance, if we consider $|S| \leq 10$, "all but 0 S is P" (0SAP) entails "at least 10 S are P" (10SIP); and in addition;

1SAP entails 9SIP by presupposition,
2SAP entails 8SIP by presupposition,
...
9SAP entails 1SIP by presupposition,

On the other hand, 9SIP also implies 8SIP, 7SIP, etc. and, therefore, they are also entailed by 1SAP. Thus, assuming that $S \geq y$, xSAP implies $y - xSIP$ where y and x are the value of exception in particular and universal statement, respectively. In this way, minimum presuppositions give existential import to universal statements.

Maximum presuppositions are expressed as *at most x*. Thus, let us now consider $|S| \geq 10$, "at least 1 S is P" (1SIP) entails "all but 9 S are P" (10SAP); and in addition;

1SIP entails 9SAP by presupposition,
2SIP entails 8SAP by presupposition,
...
10SIP entails 0SAP by presupposition,

As in the minimum presuppositions, 10SAP also implies the weaker claims 9SAP, 8SAP, etc. and, therefore, they are also entailed by 10SIP. To avoid any possible incongruousness, the values of the exception in particular statements (denoted by y) must be complementary to the value of exception in universal statements (denoted by x) regarding S; i.e., $S \leq y + x$.

Four different considerations can be made regarding minimum and maximum presuppositions [39, p. 55-56]:

1. Universal statements entail particular statements by virtue of minimum presuppositions and particular ones entail universal ones by maximum presuppositions.
2. Minimum presuppositions concern the distributed terms in universal statements and maximum presuppositions concern the undistributed ones in particular propositions.
3. Minimum presuppositions are not a truth criterion either for particular or for universal statements, while maximum presuppositions can render universal ones true and particular ones false.
4. The cardinality of S is a relevant datum. Assuming some undetermined minimum cardinality of S is insufficient for making inferences; however, every universal proposition follows from some maximum cardinality of S; for instance, from *at most x S*, we can deduce xSEP and xSAP for any P.

The combination of minimum and maximum presuppositions is a necessary condition to build the LSO based on exception statements [39, p. 56-57]. Thus, from xSAP, xSEP, xSIP, xSOP (see Table 1.17), minimum presuppositions allow us to infer, for instance, $xSIP$ and $x\overline{S}I\overline{P}$ from $xSAP$; maximum presuppositions supports the inference of $x\overline{S}A\overline{P}$ from secondary implications of the particular ones. In addition to presuppositions, the derivations also require the properties of obversion (Ob), conversion (Cv) and contraposition (Cp) (see Table 1.18). Table 1.19 shows the derivation of $aSIP$.

	$1.aSIP$	
Presuppositions	$2.S \leq a+x$	
	$3.\overline{S} \geq b+y$	
	$4.P \leq a+y$	
	$5.\overline{P} \geq c+x$	
Derivations	$6.xSAP$	Ps 1,2
	$7.aPIS$	Cv 1
	$8.yPAS$	Ps 4,7
	$9.y\overline{S}A\overline{P}$	Cp 8
	$10.b\overline{S}I\overline{P}$	Ps 3,9
	$11.x\overline{P}A\overline{S}$	Cp 6
	$12.c\overline{P}I\overline{S}$	Ps 5,11
	$12.c\overline{S}I\overline{S}$	Cv 12

Table 1.19: Equivalent relationships between exception quantifiers.

1.1.3.3 Theory of Exception Syllogism

Arguments built with exception statements have the same form as the previous frameworks; i.e., syllogisms constituted by three terms, two premises and a conclusion. However, this system admits multiple alternative valid conclusions. It is an immediate consequence of interpreting quantifiers using absolute numerical values in combination with the presuppositions explained above.

All Aristotelian moods have been checked in this framework and they are valid, attending to the equivalence established between traditional categorical statements and exception ones (see Table 1.16).

Regarding the procedures for checking the validity of the syllogisms, only syllogistic rules [39] and the algebraic method based on the use of Venn diagrams [38] are considered. Hence, the new rules to be added depend on two possible situations: i) inferences without using presuppositions and ii) inferences based on presuppositions.

In case i), two new rules appear [39, p.58]:

- **R1** The exception for a universal conclusion must not be less than the sum of the exceptions in the premises.

- **R2** The exception for a particular conclusion must not be greater than the exception of the particular permise minus the exception of the universal premise.

MP	$xMTANT$
NP	$aNTIMT$
C	$a-xNTIMT$

Table 1.20: *Datisi* syllogism with exception propositions.

R1 generates three possible scenarios:

– Invalid argument: the value of the exception in the conclusion is less than the addition of the value of the exceptions in the premises.

– Valid argument: the exception of the conclusion is greater than the exception of the premises.

– The strongest valid argument: the value of the exception in the conclusion equals the exception of the premises.

R2 generates these same three possible scenarios but for arguments with a particular conclusion. For instance, let us consider AII Mood from Figure I (see Table 1.20). The quantifier of the conclusion (C) is $a-x$, where x denotes the value of the universal premise and a stands for the value of the particular one. The problem of the existential import is avoided substituting the particular premise ($xNTIDT$) by $a+xNTIDT$ and the conclusion by $aNTIMT$.

Exception Syllogistics define 64 syllogistic schemas, 40 of which are new (the remaining 24 are the Aristotelian ones). Table 1.21 shows all of them. The conclusion does not appear because there are multiple valid alternatives, which, as we have already stated, depend on the presuppositions assumed and the cardinality of the involved terms.

Furthermore of the division into Figures, these patterns are classified into four groups according to the following rules: i) at least one premise is universal; ii) the occurrences of DT in the premises have opposite distribution values. Syllogisms of Group 1 are the only ones fully valid as they satisfy both rules and four conclusions can be inferred: the Aristotelian one and its three corresponding equivalent statements. Group 2 comprises syllogisms that only satisfy condition i) and, hence, require an additional presupposition (minimum of maximum) for the *DT* to satisfy rule ii). Group 3 are syllogisms that only satisfy rule ii), the *DT* is properly distributed but the premises are not right; in this case, a maximum presupposition for the extreme terms (*NT* or *MT*) compatible with the distribution of *DT* must be assumed to satisfy rule i). Finally, Group 4 comprises syllogisms that do not satisfy any of the two rules;

Group	Fig. I	Fig. II	Fig. III	Fig. IV	Satisfy
1	AA	AE	AI	AA	i), ii)
	AI	AO	AO	AE	
	EA	EA	EI	EI	
	EI	EI	EO	EO	
	IE	IE	IA	IA	
	IE	IE	IA	IA	
	OE	OA	IE	IE	
			OA		
			OE		
2	AE	AA		AI	i)
	AO	AI		AO	
	EE	EE	AA	EA	
	EO	EO	AE	EE	
	IA	IA	EE	OA	
	OA	OE	IE	OE	
3	IO	IO		OI	ii)
	OO	OI		OO	
4			II		none
			IO		
	II	II	OI	II	
	OI	OO	OO	IO	
	16	16	16	16	

Table 1.21: Moods of Exception Syllogistics.

to satisfy them, it is necessary to assume a maximum presupposition for the DT, which must be an universal proposition where DT is the subject, being thus distributed.

Venn diagrams [64, 38] are also valid for representing arguments of Exception Syllogistics and it is not necessary to change the notation respect to Intermediate Syllogistics (see section 1.1.2).

In [64], N. Pfeifer proposes an initial attempt to combine exception and intermediate quantifiers into a single statement with the following form *intermediate statement and exception statement*. Statements of this type are also valid for building syllogisms. Table 1.22 shows an example of the intermediate syllogism TTI from Figure III [64, p.69][5].

[5] In this example, the verb *to be* is substituted by *to have* for simplicity, since “most birds have wings” is equivalent to “most birds are issuing of wings”.

MP	Most birds have wings and at least all but *x* birds have wings
NP	Most birds have feathers and at least all but *y* birds have feathers
C	More than *x* and more than *y* things have feathers and wings

Table 1.22: TTI syllogism combining intermediate and exception quantifiers.

The procedure for checking the validity of these syllogisms is the algebraic method in combination with Venn diagrams. The first step is the formalization of the premises according to Figure 1.9, obtaining:

$$\begin{aligned} MP &\equiv e+f > x \geq d+g \\ NP &\equiv d+e > y \geq f+g \end{aligned} \tag{1.3}$$

Applying the algebraic method explained in section 1.1.2, we obtain the indirect proof:

$$\begin{array}{lr}
1\ \neg(e > x) & \text{(indirect proof assumption)} \\
2\ \neg(e+f > x) & \text{(1)} \\
3\ e+f > x & \text{(MP)} \\
4\ e > x & \text{(1)} \\
5\ \neg(e > y) & \text{(indirect proof assumption)} \\
6\ \neg(d+e > d+y) & \text{(5)} \\
7\ \neg(d+e > y) & \text{(6)} \\
8\ d+e > y & \text{(NT)} \\
9\ e > y & \text{(indirect proof (5)-(8))} \\
10\ e > xande > y & \text{(4 and 9)} \\
11\ e > max(x,y) & \text{(10)} \\
12\ b+e > max(x,y) & \text{(conclusion)}
\end{array} \tag{1.4}$$

N. Pfeifer show us that the combination of different type of propositions in a single statement also can generate valid syllogisms, which, furthermore, are more informative than the typical ones.

1.1.3.4 Summary

Exception categorical statements introduce a new way of defining the space between the universal and the particular propositions. In this case, the quantifiers refer to a number of elements of the universe of reference that do not fulfil the property of the predicate-term (the so-called exception quantifiers) and they only have an absolute reading, an option not considered in the models previously described.

From the point of view of reasoning, this is the first framework that deals with absolute quantities in syllogisms. This opens a new conception regarding inference schemas, where the reasoning process is heuristic (i.e., selecting a conclusion among different valid conclusions), rather than probative (i.e., checking if a concrete conclusion follows from the premises). Nevertheless, on the other hand, additional assumptions not explicitly considered in the argument must be assumed for reasoning, such as the described maximum and minimum presuppositions or knowing the size of S.

In conclusion, Exception Syllogistics gives rise to relevant challenges in syllogistic reasoning, such as the use of non-proportional quantifiers or heuristic reasoning, although the problem of vagueness is sparingly analysed.

1.1.4 Interval Syllogistics

All the previous approaches extend Aristotelian Syllogistics introducing new quantifiers to make a subdivision of the space between the universal and the particular statements, either as proportional quantifiers (such as intermediate ones) or absolute ones (such as exception ones). Most of them are vague and, hence, have a context-dependent meaning. Nevertheless, both Intermediate Syllogistics and Exception Syllogistics consider only crisp definitions for quantifiers.

The main approaches for dealing with vague intermediate quantifiers have been proposed in the field of fuzzy logic. The concepts of fuzzy set and fuzzy number open up a new framework for managing syllogistic reasoning involving vague intermediate quantifiers using fuzzy definitions for them. In [17], D. Dubois and H. Prade introduce a distinction between imprecise quantifiers (such as "between 30% and 40%", etc.), which are defined as intervals, and fuzzy ones (such as "most", "few", etc.), which are defined using fuzzy numbers. Thus, preserving the typical form of categorical statements *Q S are P*, both type of vague quantifiers (which are proportional) are substituted by intervals expressed as percentages, or by fuzzy numbers defined through trapezoidal functions. Nonetheless, in subsequently papers [18, 3, 16], they only focus on syllogistic patterns involving interval quantifiers, not fuzzy ones, and, for that reason, we shall deal with them first.

This framework also incorporates three significant changes with respect to the previous approaches to syllogism: i) the concept of intermediate quantifier is substituted by a numerical percentage; ii) as a consequence, the eminent linguistic character of syllogism is substituted by a numerical one; and, iii) the reasoning process, based mostly on logical proofs, becomes a calculation procedure, endowing syllogism with a heuristic character. On the other hand, it is relevant to note that the classical models of syllogisms are not explicitly dealt with and, hence, their compatibility is not checked.

This section is organized as follows: section 1.1.4.1 addresses the theory of interval categorical statements, describing how quantified statements are interpreted in terms of intervals; section 1.1.4.2, describes the path for allocating an interval to a linguistic label and how they are structured; section 1.1.4.3, describes the reasoning patterns and the procedure for calculating the quantifier of the conclusion and, finally, section 1.1.4.4 concludes with a summary analysing the strengths and weaknesses of this proposal.

1.1.4.1 Theory of Interval Categorical Statements

We initially introduce the concept of *interval categorical statement* to refer to the statements managed by this approach. Typical examples are "between 30% and 50% students are blond" or "a half of the students are blond". As we can see, they preserve the typical form of categorical statements *Q S are P*, where *Q* denotes the interval quantifier "between 30% and 50%" or the linguistic label "a half of", *S* stands for the term-set *students* and *P* by the term-set *blond*. It is worth noting that only proportional quantifiers and affirmative terms are dealt with in this framework; the treatment of the negation of quantifiers or terms is not addressed.

Regarding the precision of the quantifiers, Dubois et al. distinguish three definitions [17]:

1. Imprecise: they denote proportional quantities that are not precisely known but have precise bounds. They are usually asserted by expressions taking the form *between* $\underline{q}$ *and* $\overline{q}$, *less than* $\overline{q}$, *more than* $\underline{q}$,... where $\underline{q},\overline{q} \in \mathbb{R}$. For instance, "between 30% and 50%" (see Figure 1.10) or "less than 70%". They are represented as $Q = \left[\underline{q},\overline{q}\right]$.

2. Precise: they denote proportional quantities that are precisely known and also have precise bounds. They are usually expressed by exact percentages such as "30%" (see Figure 1.11), "70%", etc. In this case, the lower and upper bounds of the interval coincide ($\underline{q} = \overline{q}$).

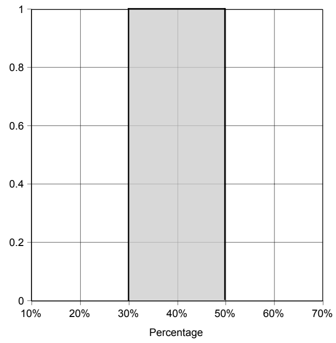

Figure 1.10: Graphic representation of “between 30% and 50%”.

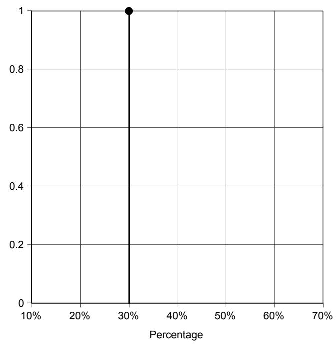

Figure 1.11: Graphic representation of “30%”.

3. Fuzzy: they denote proportional quantities that are vaguely known and whose bounds are also fuzzy. These are the properly called fuzzy quantifiers. They are usually expressed by percentages modified by expressions such as “around”, “more or less”, etc. or by linguistic terms such us “most”, “many”, “few”, etc. For instance, “around 40%” (see Figure 1.12), “aproximately 70%”, etc. They are represented by fuzzy numbers approximated through the usual four points trapezoidal representation $Q = \{q_*, \underline{q}, \overline{q}, q^*\}$, where $KER_Q = [\underline{q}, \overline{q}]$ corresponds with the kernel and $SUP_Q = [q_*, q^*]$ with the support. Figure 1.12 shows the graphical form for the quantifier “around 40%”=$\{0.25, 0.33, 0.47, 0.55\}$ with $SUP_{around40\%} = [q_* = 0.25, q^* = 0.55]$ and $KER_{around40\%} = [\underline{q} = 0.33, \overline{q} = 0.47]$.

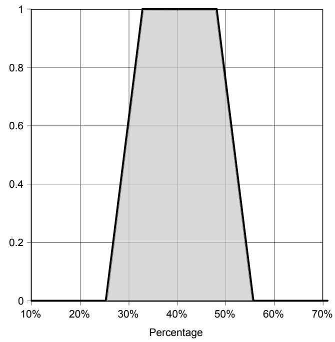

Figure 1.12: Graphic representation of “around 40%”.

1.1.4.2 Lattices of Linguistic Labels

As explained in the previous sections, the LSO is built over the relationships of opposition and it is fundamental to prove the validity of syllogisms. However, in Interval Syllogistics the quantifiers are numerically interpreted (in terms of intervals or in terms of fuzzy numbers) but without considering opposition relationships. Therefore, the development of an LSO here is not possible.

Nevertheless, in [16], authors propose a qualitative version of syllogism managing a set of elementary labels (there are seven: “none”, “almost none”, “few”, “about half”, “most”, “almost all”, “all”) of linguistic quantifiers where each label corresponds to an interval $[a,b] \subseteq [0,1]$. These labels are organized according to the concept of *order*, attending to two dimensions: specificity and certainty. This allows authors to build a biordered structure in tree form, which is reproduced in Figure 1.13. The allocation to each quantifier of its corresponding subinterval follows the path described below.

Let $\mathscr{P}$ be a partition of $[0,1]$ in adjacent subintervals and let $\mathscr{P}$ also be, by convention, the linguistic scale and the partition of the quantifiers (the seven ones mentioned above). It is also assumed that the linguistic scale is symmetric respect to 0.5 (which is reasonable) and the concept of *antonymy* is introduced. Although antonymy is an opposition relationship equivalent (in some sense) to the relationships of contrary and subcontrary in the classical LSO (see Figure 1.1) or to inner negation in the Modern LSO (see Figure 1.2), it is not introduced as a criterion for the initial selection of the quantifiers but as a condition for the selection of the numerical intervals, since it is defined as $ANT([\underline{q},\overline{q}]) = [1-\overline{q}, 1-\underline{q}]$. Thus, every quantifier has an antonym that also belongs to the scale and the partition; hence, it is

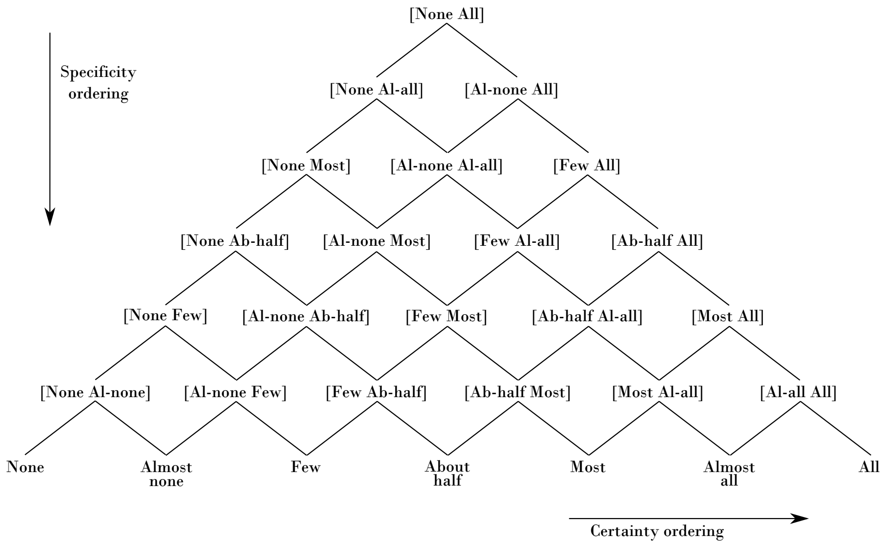

Figure 1.13: The structure ordered of linguistic quantifiers.

a fully symmetric relationship. For instance, $ANT(none) = all$ or $ANT(most) = few$ (see Figure 1.14).

To determine the corresponding subintervals to the pointed out seven linguistic quantifiers, we must take two parameters $a, b \in [0,1]$ where $a < 0.5, b < 0.5$ and $a < b$; thus, $a = \underline{q}$ and $b = \overline{q}$. The $[0,1]$ interval can be (non-strictly) symmetrically partitioned, as shown in 1.5, and each interval is assigned to a linguistic label. Table 1.23 shows this allocation and the corresponding result for $a = 0.2$ and $b = 0.4$.

$$\mathscr{P} = 0, (0,a], [a,b], [b,1-b], [1-b,1-a], [1-a,1), 1 \tag{1.5}$$

The interpretation of relationships between the linguistic quantifiers in terms of order instead of terms of opposition or negation is a new perspective in the analysis of categorical statements and, hence, also of syllogistic reasoning. However, it is also true that the key point of Interval Syllogistics is numerical quantifiers instead of linguistic ones, whose interpretation in terms of negation is closer to their use in natural language than those based on the order.

Another question that must be considered in the allocation of numerical values to linguistic terms is the so-called problem of *embarrassment of richness* in modelling vague natural language quantifiers [25]. The main idea is that there is a huge space of candidates for truth

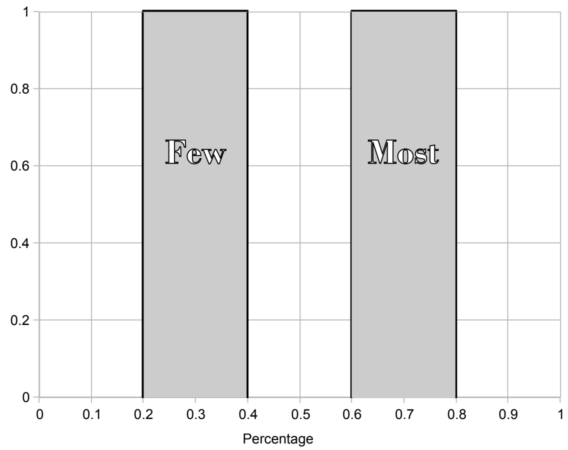

Figure 1.14: Antonymy relationship between “few” and “most” in a interval scale.

Linguistic quantifier	Interval	Numerical interval
None	0	0
Almost none	$(0,a]$	$(0,0.2]$
Few	$[a,b]$	$[0.2,0.4]$
About half	$[b,1-b]$	$[0.4,0.6]$
Most	$[1-b,1-a]$	$[0.6,0.8]$
Almost all	$[1-a,1)$	$[0.8,1)$
All	1	1

Table 1.23: Correspondence between the linguistic terms, their intervals in the scale and the values when $a = 0.2$ and $b = 0.4$.

functions that may be acceptable for defining a linguistic quantifier and it is excessively complex to manage. For instance, let us consider again the example “a half of the students are blond”. Which is the interval that best defines the quantifier? According to Table 1.23 it is *a half of* $= [0.4,0.6]$. However, if the elements in set of basic linguistic quantifiers are increased (for instance; from seven labels to nine) or the values of a and b changes or the intervals have a certain degree of overlap, then the interval allocated to “about half” also changes. Thus, the critical question is determining what the most accurate definition is or the location of the border over which the values are not acceptable. Although it is a significant problem, it is outside the framework of [16] and, for that reason, it is beyond the scope of our analysis.

1.1.4.3 Theory of Interval Syllogism

The previously analysed theories of inference focused mainly on identifying the valid inferences schemas through logical proofs, although Exception Syllogistics pointed out an heuristic aim for the arguments. However, the perspective of analysis changes in this framework. As shown in the previous section (1.1.4.2), the LSO is substituted by a tree structure based on the concept of order over a set of more or less arbitrarily chosen linguistic labels; and the logical proofs for a calculation procedure which aims to determine the numerical value of the quantifier of the conclusion in terms of the quantifiers of the premises. Hence, the premises constitute a set of restrictions over which the bounds of the quantifier of the conclusion must be calculated, maximizing the upper bound and minimizing the lower one; in other words, the reasoning process consist of obtaining the most favourable and most unfavourable proportions among the terms of the conclusion according to the proportions expressed in the premises.

Depending on the positions of the terms in the premises and the form of the conclusion, three reasoning patterns are defined [17, 18]: *Pattern I*, *Pattern II* and *Pattern III*. Although all of them involve three terms, as the classical models, the roles of MT, NT and DT are not considered. Only *Pattern I* preserves this structure, while *Pattern II* and *Pattern III* [18] have a different mechanism involving the logic operation of conjunction. In [79], both kinds of inference mechanisms are described:

1. Property inheritance (asymmetric syllogism): A term-set X and a term-set Z are linked via the concatenation of X with a term-set Y and Y with Z:

$$\begin{array}{rl} PR1 & Y \text{ in relation to } Z \\ PR2 & X \text{ in relation to } Y \\ \hline C & X \text{ in relation to } Z \end{array}$$

   This schema matches up with Figure I of Aristotelian Syllogistics. The other three figures change the position of Y in the premises.

2. Combination of evidence (symmetric syllogism): The ties between X and Z and between Y and Z are calculated separately and both are joined in the conclusion by a logic operator (conjunction or disjunction):

(a) Schema	(b) Example
Q_1 As are Bs;	$[0.85, 0.95]$ students are young
Q'_1 Bs are As	$[0.25, 0.35]$ young people are students
Q_2 Bs are Cs;	$[0.90, 1]$ young people are single
Q'_2Cs are Bs	$[0.60, 0.80]$ single people are young
Q As are Cs;	$[0.51, 1]$ students are single
Q' Cs are As	$[0.1, 0.36]$ single people are students

Table 1.24: Pattern I.

*PR*1	*Y* in relation to *Z*
*PR*2	*X* in relation to *Z*
C	*X*&*Y* in relation to *Z*

This schema is not considered in the classical approaches to syllogism as it is not based on the use of a middle term.

So, let A, B and C be the term-sets involved in a syllogism and Q_1, Q'_1, Q_2, Q'_2, Q and Q' the possible quantifiers (which can be precise, imprecise or fuzzy).

Pattern I is constituted by three pairs of premises and one pair of conclusions (see Table 1.24(a)) where Q_1 denotes the quantifier of the first premise and Q'_1 its converse, Q_2 denotes the quantifier of the second premise and Q'_2 its converse, and Q and Q' denote quantifiers for the conclusions. This is illustrated by the example in Table 1.24(b)[6].

Figure 1.15 shows a graphic representation of the Pattern I schema. The representation using Venn diagrams is substituted by a graph in which the circles denote the term-sets A, B and C; arcs with solid lines denote the quantifier Q_i for that premise; arrows with dotted lines denote the quantifiers for the conclusions; that is, Q and Q'.

Pattern I belongs to asymmetric syllogisms since B plays the role of the middle term; i.e., the link between the subject (A) and the predicate (C) of the conclusion is via concatenation of A with B and B with C. It is also worth noting that the links between all the three involved predicates have to be fully known.

[6]The data for this argument were taken from [16].

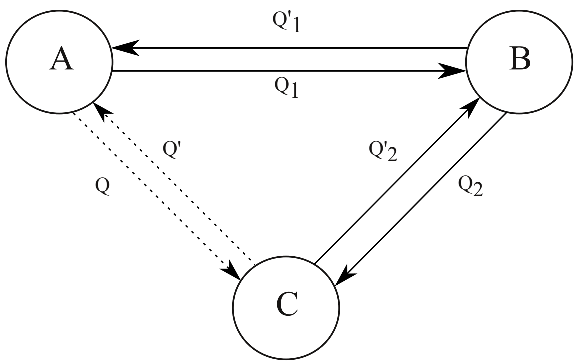

Figure 1.15: Graphic representation of Pattern I.

For precise quantifiers, the quantifier of the conclusion $Q := \left[\underline{q},\overline{q}\right]$ is calculated using the following expressions:[7]

$$\underline{q} = q_1 \cdot max\left(0, 1 - \frac{1-q_2}{q'_1}\right) \tag{1.6}$$

$$\overline{q} = min\left(1, 1-q_1 + \frac{q_1 \cdot q_2}{q'_1}, \frac{q_1 \cdot q_2}{q'_1 \cdot q'_2}, \frac{q_1 \cdot q_2}{q'_1 \cdot q'_2}[1-q'_2+q_1]\right) \tag{1.7}$$

For imprecise quantifiers, expression (1.6) should be minimized and (1.7) maximized, obtaining:

$$\underline{q} = q_1 \cdot max\left(0, 1 - \frac{1-\underline{q}_2}{\underline{q}'_1}\right) \tag{1.8}$$

$$\overline{q} = min\left(1, 1-\underline{q}_1 + \frac{\underline{q}_1 \cdot \overline{q}_2}{\underline{q}'_1}, \frac{\overline{q}_1 \cdot \overline{q}_2}{\underline{q}'_1 \cdot \underline{q}'_2}, \frac{\overline{q}_1 \cdot \overline{q}_2}{\underline{q}'_1 \cdot \underline{q}'_2}[1-\underline{q}'_2+\overline{q}_1]\right) \tag{1.9}$$

As has already been pointed out, the management of fuzzy quantifiers only appears in [17]. It consists of independently calculating the intervals corresponding to the support ($[q_*, q^*]$) and to the kernel ($[\underline{q},\overline{q}]$) and, subsequently, building the corresponding fuzzy number. Nevertheless, as pointed out in [56], this approach generates problems when non-normalized results are obtained as the truth function of the kernel is less than 1.

[7] Calculation of Q' is identical. For details of the proof, the reader is referred to [17, 18].

(a) Schema	(b) Example
Q_1 As are Bs	$[0.85, 0.95]$ students are young
Q'_1 Bs are As	$[0.25, 0.35]$ young people are students
Q_2 Bs are Cs	$[0.90, 1]$ young people are single
Q_3 As are Cs	$[0.2, 1]$ students are single
Q As and Bs are Cs	$[0.9, 1]$ young students are single

Table 1.25: Pattern II.

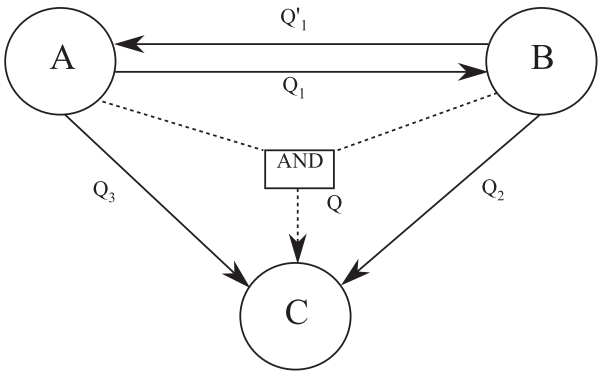

Figure 1.16: Graphic representation of Pattern II.

Pattern II is constituted by three terms and three premises (two of them being converse) and a single conclusion, which includes the three terms. Table 1.25(a) shows its schema for the quantifiers Q_1, Q'_1, Q_2 and Q_3 and Q is calculated in terms of them. Table 1.25(b) shows an illustrative example.

For precise quantifiers, the expressions that define the bounds of Q are the following ones:

$$\underline{q} = max\left(0, 1 - \frac{1-q_3}{q_1}, 1 - \frac{1-q_2}{q'_1}\right) \tag{1.10}$$

$$\overline{q} = min\left(1, \frac{q_3}{q_1}, \frac{q_2}{q'_1}\right) \tag{1.11}$$

Q_1 As are Bs	Q'_1 Bs are As
Q_2 Bs are Cs	Q'_2 Cs are Bs
Q_3 As are Cs	Q'_3 Cs are As
Q As and Bs are Cs	

Table 1.26: Pattern II'.

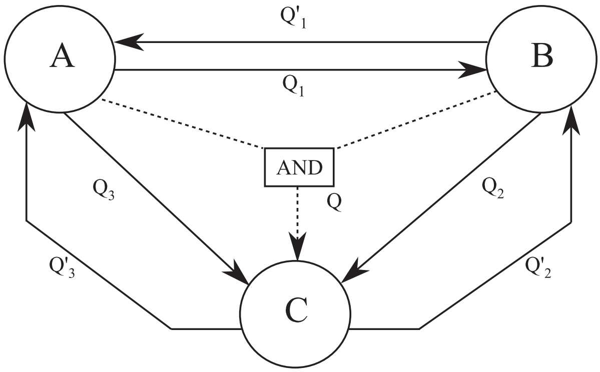

Figure 1.17: Graphic representation of Pattern II'.

For imprecise quantifiers, the previous expressions must be maximized and minimized respectively, obtaining:

$$\underline{q} = max\left(0, 1 - \frac{1-\underline{q}_3}{\underline{q}_1}, 1 - \frac{1-\underline{q}_2}{\underline{q'}_1}\right) \tag{1.12}$$

$$\overline{q} = min\left(1, \frac{\overline{q}_3}{\underline{q}_1}, \frac{\overline{q}_2}{\underline{q'}_1}\right) \tag{1.13}$$

If additional information is available, Pattern II has a general version called **Pattern II'**. It is obtained by adding the premises Q'_2 *Cs are Bs* and Q'_3 *Cs are As* to the schema of Table 1.25(a), obtaining the schema shown in Table 1.26 and the graph of Figure 1.17.

For precise quantifiers, the bounds for Q are:

$$\underline{q} = max\left(0, 1 - \frac{1-q_3}{q_1}, 1 - \frac{1-q_2}{q'_1}, \frac{q_3}{q_1} + \frac{q_2}{q'_1} \cdot 1 - \frac{1}{q'_2}\right) \tag{1.14}$$

$$\overline{q} = min\left(1, \frac{q_3}{q_1}, \frac{q_2}{q'_1}\right) \tag{1.15}$$

(a) Schema	(b) Example
Q'_2 Cs are Bs	$[0.3, 0.5]$ single people are young
Q'_3 Cs are As	$[0.7, 0.9]$ single people are students
Q Cs are As and Bs	$[0, 0.5]$ single people are young and students

Table 1.27: Pattern III.

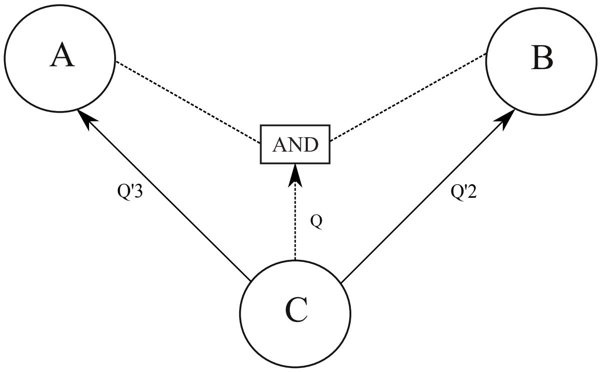

Figure 1.18: Graphic representation of Pattern III.

For imprecise quantifiers, the formula of the lower bound must be minimized and the formula of the upper bound must be maximized:

$$\underline{q} = max\left(0, 1 - \frac{1-\underline{q}_3}{\underline{q}_1}, 1 - \frac{1-\underline{q}_2}{\underline{q}'_1}, \frac{\underline{q}_3}{\overline{q}_1} + \frac{\underline{q}_2}{\overline{q}'_1} \cdot 1 - \frac{1}{\underline{q}'_2}\right) \tag{1.16}$$

$$\overline{q} = min\left(1, \frac{\overline{q}_3}{\underline{q}_1}, \frac{\overline{q}_2}{\underline{q}'_1}\right) \tag{1.17}$$

Pattern III is constituted by only two premises and a conclusion, which also includes the three terms. Table 1.27(a) shows its schema and Table 1.27(b) an example.

For precise quantifiers, the bounds for Q are:

$$\underline{q} = max\left(0, q'3 + q'2 - 1\right) \tag{1.18}$$

$$\overline{q} = min\left(q'3, q'2\right) \tag{1.19}$$

Q_1 As are Bs	Q'_1 Bs are As
Q_2 Bs are Cs	Q'_2 Cs are Bs
Q_3 As are Cs	Q'_3 Cs are As
Q Cs are As and Bs	

Table 1.28: Pattern III'.

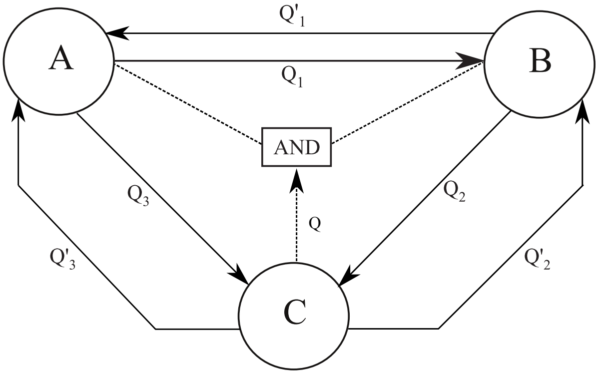

Figure 1.19: Graphic representation of Pattern III'.

In the case of imprecise quantifiers, the extension is obvious, as in the previous case:

$$\underline{q} = max\left(0, \underline{q}'_3 + \underline{q}'_2 - 1\right) \quad (1.20)$$

$$\overline{q} = min\left(\overline{q}'_3, \overline{q}'_2\right) \quad (1.21)$$

As in the case of Pattern II, there is a general version of this schema called **Pattern III'**. We can say that Pattern III' is complementary to Pattern II' as they share the same set of premises and the conclusion of the first one is the converse of the conclusion of the second one. The complete schema is shown in Table 1.28. Figure 1.19 shows the graphical representation thereof, where Q is calculated in terms of $Q_1, Q'_1, Q_2, Q'_2, Q_3$ and Q'_3.

For precise quantifiers, the lower and the upper bounds for Q are the following:

$$\underline{q} = max\left(0, q'_3 + q'_2 - 1, \frac{q'_3}{q_3} \cdot (q_1 - 1) + q'_3, \frac{q'_2}{q_2} \cdot (q'_1 - 1) + q'_2\right) \quad (1.22)$$

$$\overline{q} = min\left(q'_3, q'_2, q_1 \cdot \frac{q'_3}{q_3}\right) \quad (1.23)$$

For imprecise quantifiers, the extension of the lower bound is less simple than in the previous cases, for two reasons: a) the value for $\underline{q}_3 + \underline{q}_1 + 1$ can be positive or negative and according to it, we must use $\underline{q}'_3$ or $\overline{q}'_3$ to compute the lower bound and obtaining positive values; the same is true for q'_1, q_2 and q'_2; b) to calculate the lower bound in this pattern only five quantifiers are needed, not six; therefore, these six expressions must be used to tighten the bounds. Nevertheless, the extension for the upper bound is:

$$\overline{q} = min\left(\overline{q}'_3, \overline{q}'_2, \overline{q}_1 \cdot \frac{\overline{q}_3}{\underline{q}'_3}, \overline{q}'_1 \cdot \frac{\overline{q}_2}{\underline{q}'_2}\right) \tag{1.24}$$

1.1.4.4 Summary

One of the main contributions of this framework is to deal with the intrinsic vagueness of intermediate quantifiers directly instead of transforming them into precise definitions, as happens, for instance, in Exception Syllogistics (see section 1.1.3). Regarding the theory of inference, it is relevant to note that the reasoning process is a heuristic method for calculating the quantifier of the conclusions according to the restrictions of the premises rather than to the elaboration of a logical proof.

The authors also propose the use of subintervals of $[0,1]$ and the allocation thereof to a quantifier partition whose members are organized over the concepts of specificity and certainty, thus generating a biordered structure. The role assigned to this lattice is comparable with the role of the LSO in the other syllogistic proposals, insofar as between the number of linguistic quantifiers that can be managed is limited and they have some kind of links between them. Their differences reside in the type of relationships considered: the LSO is based on relationships of opposition, while the lattices of linguistic labels are based on relationships of specificity, certainty and antonymy.

Our contribution to this framework is to check its compatibility with Aristotelian Syllogistics [51]. Only Pattern I adheres to the structure of asymmetric syllogisms and only Figure I is compatible for the conclusion Q, as Table 1.29 shows, owing to the position of the middle term. Figures II, III and IV cannot be managed as all patterns have the middle term in the same position. Thus, if we analyse the six moods that constitute Figure I using consistent definitions for the four classical quantifiers ($A\,(all) := [1,1]$, $E\,(none) := [0,0]$, $I\,(some) := [\varepsilon,1]$, $O\,(not\ all) := [0,1-\varepsilon]$), with $\varepsilon \in (0,1]$, all of them exhibit correct behaviour (see Table 1.30).

A further weak point of Interval Syllogistics is that it only manages proportional quantifiers. Other quantifiers, such as exception or comparative ones, are avoided. On the other

Pattern	Major premise	Minor premise
Figure I	Subject	Predicate
Figure II	Predicate	Predicate
Figure III	Subject	Subject
Figure IV	Predicate	Subject
Pattern I	Subject	Predicate

Table 1.29: Position of the middle term in the Aristotelian figures and Pattern I.

	AAA	EAE	AII	EIO	AAI	EAO
Pattern I	Yes	Yes	Yes	Yes	Yes	Yes

Table 1.30: Behaviour of Pattern I with respect to the six moods in Figure I.

hand, the use of fuzzy numbers or fuzzy intervals to represent the meaning of the quantifiers has not been developed in the literature. Despite this, Interval Syllogistics constitutes a solid support for the development of framework for syllogistic reasoning compatible with the TGQ.

1.1.5 Fuzzy Syllogistics

L. Zadeh [90, 91, 93] and R. Yager [88] developed a genuine model of fuzzy syllogism, where quantifiers are always interpreted as fuzzy numbers (no intervals) and manipulated using fuzzy arithmetic, which represents a further step with respect to the previous approaches. They assume that the concept of quantifier is directly linked with the concept of cardinality and, more generally, with the concept of measuring the cardinality of a fuzzy set [93], expressed by means of $\Sigma - count$ cardinality [94]. Hence, the management of quantified statements involves the use of a fuzzy cardinality measure and fuzzy numbers.

The inference schemas preserve the basic structure of the Aristotelian ones; i.e, three terms, two premises and a conclusion. However, as in Interval Syllogistics (see section 1.1.4), there are models based on the chaining by a middle term (asymmetric syllogisms) and models based on the combination of evidence (symmetric syllogism)[8].

Given that fuzzy quantifiers must be manipulated through fuzzy arithmetic [93], the logical analysis of the arguments is substituted by an arithmetic one. This idea is in the same vein as the algebraic approaches to syllogistic reasoning, but using the fuzzy version thereof.

[8] They are explained in section 1.1.4.3

This implies a reasoning process, of a heuristic type, where the quantifier of the conclusion is calculated according to the fuzzy numbers of the premises.

On the other hand, it is worth noting that classical syllogistic proposals are not explicitly considered as a particular case, and the compatibility between them has not previously been checked. Our contribution to this framework is to analyse this question [50, 51].

This section is organized as follows: section 1.1.5.1 describes the form of quantified statements and how the linguistic quantifiers are interpreted as fuzzy numbers; section 1.1.5.2 shows the fuzzy syllogistic schemas and, finally, section 1.1.5.3, summarize the strengths and the weakness of this framework.

1.1.5.1 Theory of Fuzzy Categorical Statements

The standard form of fuzzy statements for syllogistic reasoning is the typical one; i.e, *Q S are P*, being *Q* a fuzzy quantifier, *S* the subject-term and *P* the predicate-term. For this reason, we call these statements *fuzzy categorical statements*.

Zadeh distinguishes two kinds of linguistic quantifiers: first type or absolute quantifiers, and second type or proportional quantifiers. First type refers to absolute quantities, those whose value is independent of the number of elements in the subject-term *S*; for instance, “around five”, “many”, “some”, etc. Second type refers to relative quantities, those whose value is relative to *S*; for instance, “many”, “more than a half”, “most”, etc. Thus, “*around ten* students are blond” is an absolute fuzzy categorical statement and “*most* of students are blond” is a proportional fuzzy categorical statement. It is worth noting that some linguistic quantifiers (e.g. “many”) can be absolute or proportional and, therefore, belong to one type or other according to the context [44].

The main point of Zadeh’s quantification framework is the identification between linguistic quantifiers and fuzzy numbers [90]. Hence, absolute quantifiers are interpreted as absolute fuzzy numbers and proportional quantifiers are interpreted as proportional fuzzy numbers. This is linked with the cardinality of the term-sets (fuzzy or not) over which it is quantified. In the case of fuzzy sets, Zadeh proposes a fuzzy cardinality measure, $\Sigma - Count$ [94][9], to establish that link.

[9] $\Sigma - Count$ generalizes the classical cardinality c of a subset $A \subset E$; that is $c(A) = \Sigma_{e \in E} I_A(e)$ where I_A is the membership of e to A and

$$I_A(e) = \begin{cases} 1 & \text{if } e \in A \\ 0 & \text{if } e \notin A \end{cases}$$

(P_1)	Most students are young
(P_2)	Most young students are single
(C)	$Most^2$ students are young and single

Table 1.31: An example of Zadeh's fuzzy syllogism.

$$\frac{\begin{array}{c} Q_1\ As \text{ are } Bs \\ Q_2\ Cs \text{ are } Ds \end{array}}{Q\ Es \text{ are } Fs}$$

Table 1.32: Zadeh's general fuzzy syllogistic scheme.

1.1.5.2 Theory of Fuzzy Syllogism

Zadeh defines fuzzy syllogisms as "an inference scheme in which the major premise, the minor premise and the conclusion are propositions containing fuzzy quantifiers" [93]. Thus, he preserves the same basic Aristotelian schema as Aristotle but expands the set of quantifiers that can be managed considering proportional fuzzy ones. A typical example of Zadeh's fuzzy syllogism is shown in Table 1.31. The quantifier of the conclusion, $Most^2 := Most \otimes Most$, is calculated from the quantifiers in the premises using the fuzzy arithmetic product.

Table 1.32 shows a general schema for fuzzy syllogism, where Q_1, Q_2 and Q are proportional fuzzy quantifiers ("few", "most", "many", etc.) and A, B, C, D, E and F are interrelated fuzzy properties or terms. All the valid fuzzy syllogistic inferences are built adding different constraints to this schema. For instance, the example of Table 1.31 assumes that $C = A \cap B$; $E = A$ and $F = B \cap D$.

R. Yager [88] extends Zadeh's framework incorporating immediate inferences; i.e., syllogisms with a single premise. Inferences of this kind are mainly based on the property of monotonicity. Table 1.33 summarizes them; the usual interpretation for $\geq Q_i$ being *at least* and for $\leq Q_i$ *at most*.

The reasoning process, as has already been said, is of a heuristic type for calculating a compatible conclusion rather than a logical test for checking the logic validity of the conclusion. This idea is developed through the *Quantifier Extension Principle* (QEP) [93]. Thus, let $P_1, \ldots, P_n$ be a collection of fuzzy statements which define an absolute or proportional char-

In fuzzy logic, the extension of I_A to $[0,1]$ is the membership function, μ_A. Therefore, the extension of the cardinality measure is

$$\sum -Count(A) = \sum_{e \in E} \mu_A(e)$$

Quantifier	Inference
Q is monotone non-decreasing	$if\ x > y \Rightarrow Q_x \geq Q_y$
Q is monotone non-increasing	$if\ x > y \Rightarrow Q_x \leq Q_y$

Table 1.33: Yager's immediate inferences.

(a) Schema	(b) Example
Q_1 As are Bs	*Most* students are blond
Q_2 As and Bs are Cs	*Most* blond students are tall
Q As are Bs and Cs	*Most*^{2} students are blond and tall

Table 1.34: Intersecion/Product syllogism.

acterization of a cardinality $P_i : C_i$ is Q_i; where C_i is the cardinality of a fuzzy set and Q_i is a quantifier. The QEP establishes that if $C = f(P_1;P_2;...;P_n)$ then $Q = \phi_f(Q_1;Q_2;...;Q_n)$, where C is the conclusion, $P_1;P_2;...;P_n$ are the premises, f is a function, Q is the quantifier of the conclusion, $Q_1;Q_2;...;Q_n$ are the quantifiers of the premises and ϕ_f is an extension of f by the extension principle. The main idea of the QEP is, simply, the application of the extension principle to f, obtaining a ϕ_f that can be applied to the fuzzy numbers defined by the quantifiers. Since the quantifiers are interpreted as fuzzy numbers, the operations between the quantifiers must be fuzzy arithmetic operations.

Now we analyse each of the syllogistic inference patterns proposed by Zadeh that emerge from the general pattern (Table 1.32).

Intersection/product is a symmetric syllogism and it is assumed $C = A \cap B, E = A, F = B \cap D$ [93]. Table 1.34(a) shows the schema, where Q_1 denotes the quantifier of the first premise, Q_2 the quantifier of the second one and Q the quantifier of the conclusion; in which $Q = Q_1 \otimes Q_2$, $\otimes$ being a fuzzy product. Its graphic representation is shown in Figure 1.20. Table 1.34(b) shows an example[10].

Multiplicative chaining (MC) is an asymmetric syllogism and it is assumed that $C = B, D = C, E = A, F = C$. Table 1.35(a) shows its linguistic expression [93], where Q_1 denotes the quantifier of the first premise, Q_2 the quantifier of the second premise and Q the quantifier of the conclusion; in which $Q \geq Q_1 \otimes Q_2$. Table 1.35(b) shows an example and Figure 1.21 its graphic representation.

[10]For details of the proof of all these schemas, reader is referred to [93].

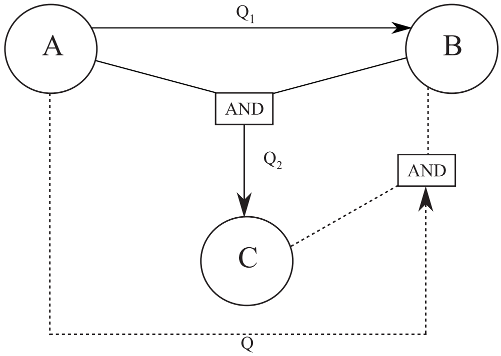

Figure 1.20: Graphic representation of Intersection/product Syllogism.

(a) Schema	(b) Example
Q_1 As are Bs (all Bs are As)	Most American cars are big
Q_2 Bs are Cs	Most big cars are expensive
Q As are Cs	*Most*^{2} American cars are expensive

Table 1.35: Multiplicative Chaining syllogism.

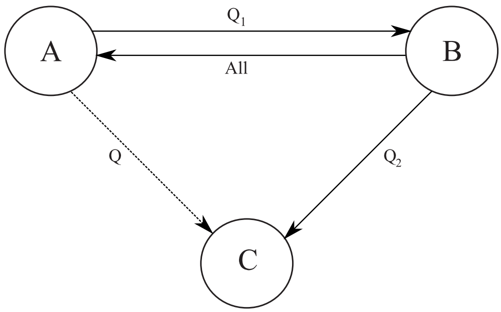

Figure 1.21: Graphic representation of Multiplicative Chaining syllogism.

The three terms involved have the following roles: A is the subject of the conclusion, C its predicate and B is the chaining between A and C. Since the terms of the conclusion are "chained" by a middle term, this pattern is called a chaining pattern.

It is worth noting that there is an additional constraint between A and B, namely is, $B \subseteq A$, i.e. $\mu_B(u_i) \leq \mu_A(u_i), u_i \in U, i = 1, \ldots$ It also can be expressed as the quantified statement *all*

(a) Schema	(b) Example
Q_1 Bs are As	Most big cars are American
Q_2 Bs are Cs	Most big cars are expensive
Q As are Cs	$\geq 0 \vee (2\ \text{most} \ominus 1)$ American cars are expensive

Table 1.36: Major Premise Reversibility Chaining syllogism.

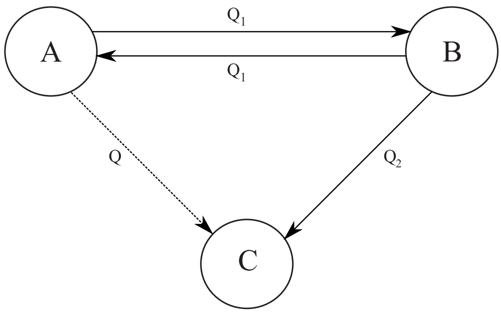

Figure 1.22: Graphic representation of MPR chaining syllogism.

B are A. This constraint is relevant as it allows us to ascertain the distribution of the elements in sets A and B. Without this information, the conclusion cannot be calculated.

Taking all the previous considerations into account, the procedure for calculating the conclusion is shown in Equation (1.25), where $\otimes$ denotes a fuzzy product and thus the chaining pattern is called multiplicative.

$$Q \geq (Q_1 \otimes Q_2). \tag{1.25}$$

Major premise reversibility chaining (MPR) is an asymmetric syllogism and it is assumed that $C = B, D = C, E = A, F = C$. The schema thereof is shown in Table 1.36(a), where Q_1 denotes the quantifier of the first premise, Q_2 the quantifier of the second premise and Q the quantifier of the conclusion. Table 1.36(b) shows an example and Figure 1.22 its graphic representation.

We can consider this model as a variant of the MC pattern, where the constraint $B \subseteq A$ is substituted by the reversibility of the first premise; i.e., Q_1 *A are B* $\leftrightarrow$ Q_1 *B are A*. Zadeh points out that this semantic equivalence is approximate rather than exact, but the calculation of how "approximate" it can be remains a non-trivial open question [93].

(a) A. Conjunction	(b) A. Disjunction
Q_1 As are Cs	Q_1 As are Cs
Q_2 Bs are Cs	Q_2 Bs are Cs
Q As and Bs are Cs	Q As or Bs are Cs

Table 1.37: Antecedent Conjunction/Disjunction schemas.

(a) A. Conjunction	(b) A. Disjunction
Many students are tall	*Many* students are tall
Most teachers are tall	*Most* teachers are tall
$[None, all]$ teachers and students are tall	*Many* students or teachers are tall

Table 1.38: Antecedent Conjunction/Disjunction examples.

Nevertheless, there is a notable obstacle for this constraint. For proportional quantifiers, it does not hold that a quantified sentence (e.g., "most American cars are big") and another sentence with the arguments interchanged ("most big cars are American") are semantically equivalent. However, it is true when we are talking about absolute quantifiers; e.g., "around a hundred thousand American cars are big" is semantically equivalent to "around a hundred thousand big cars are American".

The procedure for calculating the conclusion, shown in Equation (1.26), involves fuzzy addition ($\oplus$) and subtraction ($\ominus$) arithmetic operations.

$$Q \geq max(0, Q_1 \oplus Q_2 \ominus 1) \tag{1.26}$$

Antecedent Conjunction/Disjunction is a symmetric syllogism and it is assumed that $B = C, C = B, D = C, F = C$. The subject of the conclusion for the Antecedent Conjunction (AC) pattern is constituted by two conjuncted terms assuming that $E = A \cap B$ (schema shown in Table 1.37(a)); the subject of the conclusion for the Antecedent Disjunction (AD) pattern is constituted by two disjuncted terms assuming that $E = A \cup B$ (schema shown in Table 1.37(b)). Table 1.38 shows an example of each pattern and Figure 1.23 their graphical representation.

The AC pattern entails assuming that the items of evidence are conditionally independent [93]. This fact leads us to dismiss it from the perspective of syllogistic reasoning as this type of additional hypothesis is beyond the scope of the present research.

The AD pattern does not need additional constraints. The conclusion is $Q \leq (Q_1 \oplus Q_2)$, $\oplus$ being a fuzzy addition.

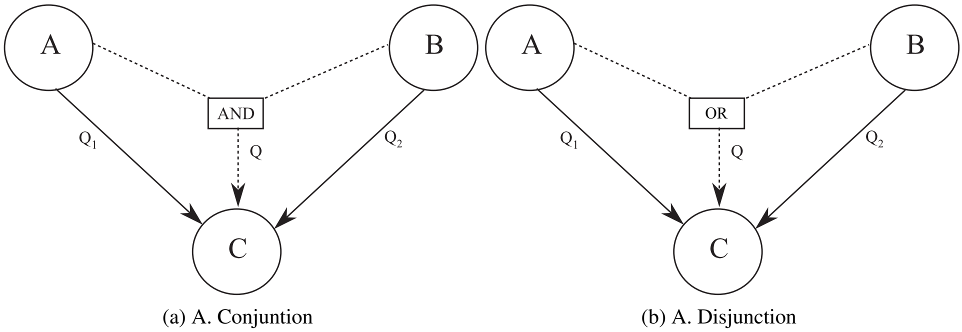

Figure 1.23: Graphic representation of A. Conjunction and A. Disjunction.

Having described Fuzzy Syllogistics, we select a number of critical reviews. In [32], E. Kerre and Y. Liu, shows the disadvantages of QEP. Firstly, different results can be obtained, depending on how the cardinality equations are written; in consequence, ambiguity arises in the result of a syllogism[11]. Moreover, inference schemes are not reversed either with respect to addition or multiplication. Finally, there is an incompleteness of fuzzy quantifiers and fuzzy numbers, and, therefore, a good chance of getting a discontinuous result [32].

1.1.5.3 Summary

Zadeh's model was the first that dealt with a genuine fuzzy syllogism; i.e., the vague quantifiers of the corresponding premises are represented through fuzzy sets. Furthermore, some new schemas of inference are proposed, both asymmetric and symmetric ones. We shall now go on to check the compatibility of Aristotelian Syllogistics with Zadeh's framework [50, 51].

Firstly, it is worth noting that no symmetric syllogisms are valid to manage Aristotelian syllogisms as they are not supported on the chaining of terms. Furthermore, the intersection/product scheme involves logic operations in the second premise and the conclusion that do not appear in the classical moods. Therefore, none of these patterns can be considered for this analysis. Thus, only the MC and MPR patterns are possible candidates.

[11] For instance, in the example of Table 1.31, if $most = \frac{3}{4}$, the quantifier of C is $\left(\frac{3}{4}\right)^2$. This functional relationship can also be written as $\frac{most^3}{most}$, but in the case of fuzzy arithmetic with fuzzy numbers the equivalence between $most^2$ and $\frac{most^3}{most}$ is not fulfilled.

Pattern	Major premise	Minor premise
Figure I	Subject	Predicate
Figure II	Predicate	Predicate
Figure III	Subject	Subject
Figure IV	Predicate	Subject
MC	Subject	Predicate
MPR	Subject	Predicate

Table 1.39: The position of the middle term in Aristotelian Figures and Zadeh's MC and MPR patterns

	AAA	EAE	AII	EIO	AAI	EAO
MC	No	No	No	No	No	No
MPR	No	No	Yes	No	No	No

Table 1.40: Behavior of Zadeh's patterns with respect to Figure I.

If we consider their capability to reproduce the four classical figures, only Figure I is compatible, as shown in Table 1.39, owing to the position of the middle term. Therefore, Figures II, III and IV cannot be modelled using this approach. This limits the scope of the model and hence only the six Aristotelian moods of Figure I can be checked (see Table 1.40). Our result is that only the AII mood is compatible with the MPR scheme (i.e., one mood out of twelve).

The problem in the MC pattern is in the constraint $B \subseteq A$ in the first premise (minor premise in Aristotelian terms), which is too restrictive. In this set of Aristotelian moods, the converse of the minor premise in linguistic terms is *some B are A*; that is, $A \cup B \neq \emptyset$; for instance, in the AII mood, $B \subseteq A$ cannot be inferred from the second premise, *I: some A are B*. This condition is less restrictive than Zadeh's and thus many of the possible inferences dismissed by Zadeh's pattern can be solved using the Aristotelian framework.

The MPR pattern is only valid for the AII mood, since the statement and its converse in the minor premise have the same quantifier ("some"). For the other moods, those that include A in the minor premise present the same problem as in the MC pattern. The solution of the EIO mood, which is compatible, presents a problem: the value obtained for Q is 0 instead of $[0, 1-\varepsilon)$ with $\varepsilon \in (0, 1]$, the usual representation for the quantifier "not all" in a proportional interpretation. The origin of this problem is that Zadeh's framework is not capable of managing moods involving decreasing quantifiers, such as "not all". In conclusion, Zadeh's approach to syllogistic reasoning is compatible with the AII mood of Aristotelian Syllogistics, but only for the MPR scheme.

On the other hand, as stated in [32], Fuzzy Syllogistics also shows relevant problems in the QEP, the core of its reasoning process, and from this point of view, it is a weak model. However, it introduces an interesting matter not considered until now: additional constraints or premises that are not necessarily explicit in the argument. For instance, MC and MPR models need assumptions that are not considered as premises in the reasoning schemas; what is more, if they are not included in the syllogism, it is not valid. In this case, we can think that Zadeh is assuming that they belong to the background or the context shared by the users and, therefore, they need to be clearly established to avoid fallacies in reasoning.

1.1.6 Generalized Intermediate Syllogistics

Another proposal that manages syllogistic arguments involving fuzzy quantifiers is the one developed by V. Nóvak and P. Murinová [42, 37] which aims to generalize Intermediate Syllogistics (see section 1.1.2) called, for that reason, Generalized Intermediate Syllogistics.

It is developed in the framework of fuzzy type theory (FTT), a higher order logic, where the intermediate quantifiers are introduced as special formulae using the formal theory of *trichotomous evaluative linguistic* (TEv) [41]. According to [42], this framework presents the following advantages:

1. The standard syntax of FTT is sufficient for managing intermediate quantifiers as they are only short forms for special formulas.

2. This theory is general enough to cover a broad class of intermediate quantifiers and provides a unique definition for them.

3. Syntactical properties of the quantifiers can be studied separately from their interpretation (semantics). Therefore, syllogisms can also be studied from a syntactical point of view, independently of their semantics.

We shall focus on a intuitive description of the main concepts of the framework, as well as on their main contributions to the study of syllogism. The formal development of this framework can be found in [37, 41, 42].

This section is organized as follows: section 1.1.6.1 describes how quantified statements are formalized; section 1.1.6.2 explains the new tool for interpreting quantifiers; section 1.1.6.3

describes a generalized version of the LSO; section 1.1.6.4 describes the main characteristics of the reasoning process proposed in this framework; and, finally, section 1.1.6.5 summarizes the strengths and the weakness of this proposal.

1.1.6.1 Theory of Generalized Intermediate Categorical Statements

Generalized Intermediate Categorical Statements are explicitly introduced as a generalization of Peterson's intermediate quantifiers (see section 1.1.2). Thus, they have the typical structure of quantified statements (*Q S are P*) and the same type of quantifiers (proportional ones). The most relevant contribution is the development of a formal theory for managing it.

Accordingly, quantified statements are formalized through FTT. Classical Type Theory is well known in mathematics and logic. It is a higher order logic built as an extension of first-order logic. It was proposed by A. Whitehead and B. Russell in their renowned book *Principia Mathematica* [87] published in three volumes between 1910 and 1913. The concept of *type* was introduced to overcome the *Russell paradox*[12], also known as *barber paradox*, and, in this way, defends the thesis of logicism; that is, mathematics is reducible to logic. Thus, *type* is a new kind of category that allows us to form a hierarchy and assigns each mathematical entity to a type. For instance, a *natural number* is a type. As in classical logic, the structure of its truth-values is $\{0,1\}$.

FTT is an extension of classical type theory to fuzzy logic. Therefore, it is still a higher order logic but the structure of the truth-values is $[0,1]$ [40]. It is relevant to note that, although the motivation of classical type theory is mathematics, the motivation of FTT is linguistics and the management of natural language.

Aristotelian quantifiers are also included in this framework, through their equivalent definition in first order logic; that is[13],

$$\begin{aligned}\text{all S are P} &\equiv \forall x(Sx \Rightarrow Px)\\ \text{some S are P} &\equiv \exists x(Sx \wedge Px)\end{aligned} \tag{1.27}$$

and their corresponding negations;

$$\begin{aligned}\text{no S are P} &\equiv \forall x(Sx \Rightarrow \neg Px)\\ \text{some S are not-P} &\equiv \exists x(Sx \wedge \neg Px)\end{aligned} \tag{1.28}$$

[12] Russell wrote it to G. Frege in a letter on June 16 [69] meanwhile he was working on *Principia Mathematica*.
[13] In type theory, $\Rightarrow$ is equivalent to $\rightarrow$.

Thus, any quantified statement in the standard form *Q S are P*, whose quantifier is modelled through TEv, is expressed as a special formula in the FTT attending to its position between the universal and the particular quantifiers.

1.1.6.2 Trichotomous Evaluative Linguistic (TEv)

Many of the expressions human beings use in their daily life are expressions such us “tall”, “very tall”, “medium”, etc. These expressions belong to the class of *evaluative linguistic expressions*, which “typically characterize a person’s behaviour or attitude in terms of the speaker’s subjective judgement” [31].

A subclass of evaluation linguistic expressions comprises *trichotomous evaluative expressions*, which are usually constituted by three terms: a nominal adjective, its antonym and a medium one between them. Its canonical form is of the type *small-medium-big*. They are frequently used to evaluate, judge, estimate, qualify, etc. different kind of situations. TEv expressions have the following properties:

– They are adjectives of manner; i.e, the are subjective and context dependent.

– They constitute a bounded ordered scale. Such a scale can consist of real measuring units (centimetres, litres, etc.) or some kind of abstract units (happiness, sadness, etc.).

– They are essentially vague.

As a vague part of natural language, they were dealt with by fuzzy logic. However, in most cases, they are interpreted by simple fuzzy sets in the universe of real numbers, using triangular or trapezoidal membership functions. Although fuzzy logic is sensitive to the context, this type of definition does not adequately manage its scalar character; i.e., the meaning of “small” is not only a question of centimetres or meters but what the definition of “big” is.

Nóvak [41] provides a formal mathematical model for dealing with this complex and vague meaning using FTT. He assumes the following hypothesis [41, p.2941]:

> Vagueness of the meaning of natural language expressions is a consequence of the indiscernibility phenomenon. Therefore, any extension of a natural language expression can always be characterized by the same specific indiscernibility relation.

Thus, the semantics of the indiscernibility TEv expressions is characterized in FTT through the use of *fuzzy equality*, an imprecise equality that characterizes the similarity between two objects. This idea has been widely studied in the fuzzy logic literature and its analysis is beyond the scope of this research.

Nóvak interprets intermediate quantifiers as a kind of TEv, the universal and the existential quantifiers being the extremes and any other intermediate quantifier between them, the medium element of the scale. The extremes, in this case, have crisp definitions and the medium corresponds with a fuzzy set defined as a membership function. Hence, any intermediate categorical statement is defined by one of the following formulas[14], where x and z denote any element of the universe of the discourse U and S and P the standard term-sets:

$$(Q * \forall_{Ev} x(S,P)) \equiv (\exists z)((z \subseteq S) \wedge Ev(\mu B)z \wedge (\forall x)(zx \Rightarrow Ax)) \tag{1.29}$$

$$(Q * \exists_{Ev} x(S,P)) \equiv (\exists z)((z \subseteq S) \wedge Ev(\mu B)z \wedge (\exists x)(zx \wedge Ax)) \tag{1.30}$$

As we can see, this definition consists of three parts:

$(\exists z)(z \subseteq S)$ It is the existential import, that is, the subject term of intermediate categorical statements cannot be an empty set. However, it is relevant to note that the existential import is not inherent to his definition but it is explicitly added; on the other hand, it also shows the left edge of the scale because it is equivalent to say that "there is at least one element in *S*".

$Ev(\mu B)z$ This represents the evaluative process that generates the fuzzy set corresponding to the meaning of the linguistic quantifier such as "most", "many", "few", etc. It corresponds with the medium member of the scale.

$(\forall x)(zx \Rightarrow Ax)/(\exists x)(zx \wedge Ax)$ It is the right-hand member of the scale.

It is relevant to note the syntactical character of this definition of the quantifiers. Although the fuzzy set that defines the meaning of the quantifier changes, the general formula is still valid. It should also be noted that the TEv definition substitutes the role of LSO in this framework, as it defines the meaning of the intermediate quantifiers according to the dependences with the universal and particular ones.

Regarding the quantifiers considered by Nóvak, he directly refers to the analysis provided by Peterson (see section 1.1.2; 5-quantity LSO).

[14] Here, we avoid the use of the specific formalism of FTT as it has not been adequately defined. At pointed out, the main aim of this section is to present the general idea of this approach rather than its technical details.

1.1.6.3 Generalized Logic Square of Opposition with Intermediate Quantifiers

In [36], P. Murinová and V. Novák develop a generalized version of the LSO proposed by P. L. Peterson and B. E. Thompson (see section 1.1.2.1, Figure 1.7) in the framework of FTT. We stated that, in all but Zadeh's Syllogistics, the role of an LSO is fundamental. In Interval Syllogistics, when the authors propose a linguistic version of their model, it is substituted by a biordered structure different from a LSO but which plays an analogous role. In conclusion, any syllogistic framework that manages linguistic quantifiers needs a version of the LSO.

P. Murinová and V. Novák assume that "intermediate quantifiers are classical general ($\forall$) or existential ($\exists$) quantifiers, but the universe of quantification is modified, and the modification can be imprecise" [36, p. 12]. Using a FTT that extends the TEv, they can manage them under a fuzzy framework. They obtain a high general and robust result as they only use syntactical proofs, which also involve the Aristotelian LSO as a particular case.

Regarding the problem of existential import, they presuppose that any term-set involved in a quantified statement are non-empty sets. They use first-order logic in the formulation of universal and include the existence of S by conjunction into the Aristotelian proposition; for instance, $A := (\forall x)(Sx \Rightarrow Px)\&(\exists x)x$, where & denotes the corresponding fuzzy conjunction. With this assumption, all the relationships of an LSO are proved. Figure 1.24 shows the square obtained, where a,e,p,b,t,d,k,g,i and o denote the truth value for each statement, not a fuzzy number associated to the quantifier, as P. Murinová and V. Novák only consider logic definitions of them. Thus, for example, if "almost all students are tall" with $p = 0.4$ then "most students are tall" is with $g = 0.45$ since $p \leq t$[15].

There are some relationships from Peterson and Thompson's 5-quantity LSO that are not included in this version; specifically, the relations between the statements K, O, G and I are omitted. The reason is that K and G have undefined meanings and the relationships between each one and with other quantifiers are much too complicated. The authors point out that dealing with these quantifiers will form part of subsequent investigations.

1.1.6.4 Theory of Generalized Intermediate Syllogism

The main innovation of this framework is the independence of syntax and semantics and how the role of the LSO is avoided using TEv. In this way, the validity of syllogisms is proved using syntactical proofs, which confers a great deal of robustness to this proposal.

[15] The proofs can be confronted in [36]

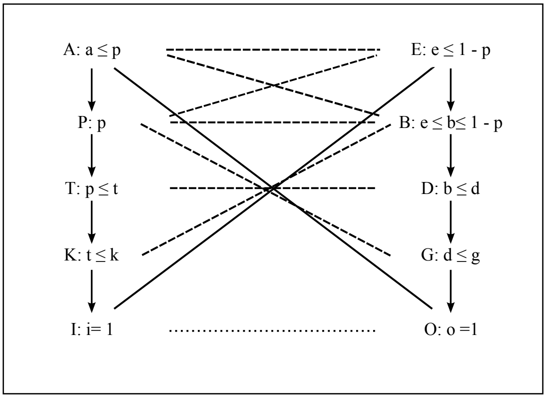

Figure 1.24: Generalized 5-Quantity LSO.

Nóvak distinguishes between valid implications and valid syllogisms [42, p.1240]. The former correspond to the concept of immediate inferences but they are exclusively based on the property of monotonicity; i.e., "wider" quantifiers entails "narrower" ones. The examples are well known, for instance, "almost all students are tall" implies "most students are tall"[16]. The latter correspond with the standard form of syllogism (three terms, two premises and a conclusion) but adopting a formulation in terms of propositional logic; i.e., P_1 stands for the MP, P_2 denotes the NP and C stands for the conclusion. Thus, a syllogism is *formally valid* if $P_1 \& P_2 \Rightarrow C$.

In [42, 37] all the valid intermediate syllogisms are proved and all the invalid ones are discharged. Hence, the compatibility between both approaches and the generalization to fuzzy logic of this framework respect to the previous one is proved.

1.1.6.5 Summary

Generalized Intermediate Syllogistics is a generalization of Intermediate Syllogistics in fuzzy logic. It is relevant to note that it is the only fuzzy approach that explicitly considers Aristotelian Syllogistics as a particular case. Furthermore, it also analyses syntax and semantics independently obtaining, by this way, such a strong model.

On the other hand, the TEv framework is a genuine fuzzy proposal for interpreting intermediate quantifiers without using an LSO. However, from the point of view of inference

[16] The proof is explained in [42, p.1240].

schemas, Nóvak only considers proportional quantifiers and arguments with the typical form: three terms, two premises and a conclusion.

1.2 Conditional Interpretation

Conditional interpretation assumes that the deep structure of quantified statements is conditional; i.e., an implication where a relationship between an antecedent and a consequent is established. For instance, “all natural numbers are real numbers” ≡ “if something is a natural number, then it is *necessarily* a real number”, where “necessarily” stands for the quantifier “all” and determines the power of the link between the antecedent part and the consequent one. This interpretation entails a different categorization of the elements that compound a quantified statement:

- Subject and predicate are not sets but propositions. So, instead of talking about about subject-term and predicate-term, we have to talk in terms of antecedent or sufficient condition (which corresponds with the subject-term) and consequent or necessary condition (which corresponds with the predicate-term) of the conditional. For instance, the subject-term *natural number* is transformed into the proposition “if something is a *natural number*”.

- Quantifier does not denote a quantity relationship between two terms but how strong is the link between the antecedent and the consequent of the implication. For instance, the quantifier “all” indicates a perfect link between *if-then*; i.e., “if something is a natural number then it is *necessarily* a real number”; other quantifiers such as “many human beings are Chinese” are equivalent to “if something is a human being then it is *quite probably* Chinese”.

- The universe of discourse is not restricted to a determined set rather it is unconditional. Thus, “if something is a natural number then...” refers to any element of the unconditional universe of discourse endowed with the property of being a natural number; not specifically the set whose elements are the natural numbers.

From these differences some additional conclusions can be inferred. Firstly, set-based interpretation allow us to interchange the subject-term and the predicate-term in a categorical statement and infer a new valid statement, but it cannot deal with singular terms (those that refers to a concrete being, not to sets); nevertheless, conditional reading does not allow us that direct conversion in the statements although it can manage propositions that involve singular terms [29] in a straightforward manner. For instance, from "all human beings are mortal" we can infer "some mortal beings are human beings"[17]; however, from a statement such as "Gaius Iulius Caesar is a human being" the same kind of inference cannot be done because *Gaius Iulius Caesar* refers to a concrete historical person and not to a set. Using the conditional interpretation, the equivalent statement to "all human beings are mortal" is "if something is a human being then it is mortal", where *something* refers to any element of the universe and *human being* and *mortal being* refers to the properties being *human* and *mortal* respectively. The interchange between both terms produces the statement "if something is mortal then it is a human being", that is not true and, then, the interchange is not acceptable; however, the conditional interpretation can deal perfectly with sentences that involve singular terms such as "Gaius Ilius Caesar"; i.e., "if something is Gaius Iulius Caesar then it is a human being".

Another corollary deals with the so-called problem of existential import [29]. The best way to illustrate this is through a question. Let us consider the statement "all gnomes are short". Is it true or false? First of all, we agree that gnomes do not exist. From a set-based interpretation and assuming the existential import, it can only be true if and only if there are gnomes; therefore, it is false. Without assuming the existential import, the statement is vacuously true but true, because *gnomes* denotes an empty set ($\emptyset$), that is a subset of any set. In this case, some inferences of the LSO are not valid, such as subalternation; for this reason, in general, all syllogistic approaches based on the set-based interpretation assume the existential import. However, for the conditional reading this is an irrelevant question because the inferences of the LSO is a type of inference not considered.

From the point of view of Linguistics, the set-based interpretation is compatible with the TGQ [86] and this allows us to introduce a new kind of linguistic quantifier, not traditionally considered in the framework of syllogistic reasoning, which increases the number of arguments that can be managed. Nevertheless, although proportional quantifiers have a clear conditional reading, the interpretation of other kinds of quantifiers as absolute ("twenty five", "three", etc.), exception ("all but three", etc.) or comparative ("double", etc.) is not so clear. For instance, the meaning of "all students but two are blond" is clear in the set-based reading, but how can it be expressed using the conditional reading? It can be something like *if x is a student then x is blond* [0.98]? Why not [0.99]?

[17] This kind of inference is detailed explained in section 1.1.1.

If we consider now the pragmatic dimension of language, there are also a number of additional considerations that must be discussed. Quantified statements, as a type of categorical statements, are defined as assertions; i.e., statements that can be true or false because they describe a particular fact or event of the reality. For instance, "all natural numbers are real numbers" describes a characteristic of natural numbers. This statement has neither uncertainty nor vagueness. However, its equivalent conditional statement, *if x is a natural number, then x is a real number*, is not an assertion; rather it describes a condition for the possibility of an event (choose a *x* that, in the case of having the property *natural number*, it also has the property of *real number*), with the corresponding uncertainty associated to it. For this reason, it is assumed that the conditional reading is a kind of probabilistic reading that can be equivalent to the set-based reading from an operative point of view, but they are not completely equivalent in other dimensions of natural language.

This section is organized as follows: section 1.2.1 describes the framework proposed by M. Oaksford and N. Chater in the field of psychology of reasoning; section 1.2.2 contains the framework proposed by M. Spies based on Dempster-Shafer theory and, finally, section 1.2.3 introduce briefly the proposal of syllogism by D. Schwartz.

1.2.1 Probabilistic Syllogistics

M. Oaksford and N. Chater [43] propose a framework to deal with syllogistic reasoning in terms of probability, also providing a set of "fast and frugal" heuristics instead of the usual logical and deductive framework. They assume the following hypothesis:

- In their daily life, human beings manage incomplete information, uncertainty and vague concepts all the time. Furthermore, our main way of communication, natural language, involves many of this kind of concepts, and this is not inconvenient for making inferences. Therefore, human ordinary reasoning must be managed using tools that are able to deal with imprecision and uncertainty.

- Probability and logic deal with disjoint sets of phenomena of human reasoning; i.e., there is no competition between them but they are complementary. However, if genuine logical tasks can be dealt with using a probabilistic approach, it is possible to conclude that a general theory of human reasoning must be a probabilistic framework.

- Aristotelian Syllogistics is the core area of logical deductive reasoning. Therefore, modelling syllogistic reasoning using a probabilistic framework is a significant contribution to a probabilistic general theory of human reasoning.

The main objectives to be achieved are:

- The authors work in the field of psychology of reasoning, not in logic (the former attempting to understand how human beings make inferences and the latter trying to discover the rules of the right reasoning, independently of how human beings execute them). As a consequence, the psychology of reasoning studies any reasoning process through which people execute an inference (valid or not) while logic focuses strictly on the validity of arguments.
- The authors' main aim is to describe the set of procedures that human beings use in the reasoning. In this way, they interpret reasoning in terms of heuristic strategies used by the reasoner to achieve a conclusion accepted as true, even though a subsequent logical analysis shows that it is invalid.
- The evidence of this framework is the psychological experiments with individuals analysing whether the conclusion predicted by the model and the conclusion achieved by the person match. Logical proofs of the arguments are not a part of the experiments.

We shall focus mainly on the interesting ideas from the point of view of this research and on the criticism made by B. Geurts [23] about the management of non-proportional quantifiers.

This section is organized as follows: section 1.2.1.1 describes how quantified statements are interpreted in terms of probability; section 1.2.1.2 describes the inference mechanism; section 1.2.1.3 explains the criticism formulated by B. Geurts of this probabilistic model's capabilities for representing other kinds of quantifiers, and, finally, section 1.2.1.4 summarizes the main characteristics of this framework.

1.2.1.1 Theory of Probabilistic Quantified Statements

Oaksford and Chater start with the following assumption regarding quantified statements [43, p. 219]:

> Quantified statements can be given intuitively natural meanings in terms of constraints on the conditional probability of the predicate given the subject.

Linguistic Expression	**Probabilistic Expression**
All S are P	$P(P\|S) = 1$
No S are P	$P(P\|S) = 0$
Some S are P	$P(P\|S) > 0$ and S P are not empty
Some S are not P	$P(P\|S) < 1$ and S *not* $-P$ are not empty
Most S are P	$1 - \varepsilon \leq P(P\|S) < 1, \varepsilon > 0$
Few S are P	$0 < P(P\|S) \leq \varepsilon, \varepsilon > 0$

Table 1.41: Probabilistic expression of "all", "none", "some", "some…not", "few" and "most".

Thus, a quantified statement such as *all S are P* means that the probability of P given S is 1; in probability notation, $P(P|S) = 1$. For intermediate quantifiers, such as "most", "many", etc. the value of the probabilities changes in the interval $[0,1]$. For instance, *most S are P* means that the probability of P given S is less than 1 but high; i.e., $1 - \varepsilon \leq P(P|S) < 1$, where $\varepsilon > 0$ although ε changes according to meaning of the linguistic quantifier. Thus, for *few S are P* the probability writing is $0 < P(P|S) \leq \varepsilon$, where ε is small as it represents the meaning of "few". Table 1.41 shows the probabilistic definition for the Aristotelian quantifiers and "most" and "few". It is relevant to note "all" does not include the existential import.

A new concept not explicitly used previously in the analysis of quantification propositions is introduced: the informativeness (I) of a statement (s), $I(s)$. It is taken from information theory developed by Shannon and Waver [73], where $I(s)$ is inversely related to its probability ($P(s)$):

$$I(s) = log_2 \left[\frac{1}{P(s)} \right] \tag{1.31}$$

Thus, given two quantified statements, B and Γ, if $B \rightarrow \Gamma$ then $I(\Gamma) < I(B)$. This provides a criterion for establishing a partial order among the quantified statements:

$$\begin{aligned}
&I(All) > I(Some)\\
&I(Most) > I(Some)\\
&I(Few) > I(Some)\\
&I(Most) > I(Some \ldots not)\\
&I(Few) > I(Some \ldots not)\\
&I(None) > I(Some \ldots not)
\end{aligned} \tag{1.32}$$

This partial order can be transformed into a total order with an additional assumption [43, p. 221]: the properties referred to in natural language typically apply to a small proportion of possible objects. Authors qualify it as a 'rarity' attending to the frequency of quantified statements that are true and, hence, informative. Many of the predicates that human beings manage are mutually exclusive, for instance *cat* and *dog*: no being can be a cat and a dog simultaneously; in other words, no dog is a cat. In general, "no" statements are true and not so informative because they describe information that is usually assumed as hidden premises in our arguments. It is true that at other times, "no" statements can be very unexpected and highly informative as they describe surprising relationships, but this is not the most common situation. "Some…not" sentences are even less informative than "no" statements as they can be predicated of almost all properties expressed in natural language; they include not only "no" statements but also "some" statements.

$$I(All) > I(Most) > I(Few) > I(Some) > I(None) >> I(Some \ldots not) \tag{1.33}$$

Expression (1.33) is the generalization of (1.32). It shows the total order I with this assumption, where $>$ stands for "more informative than" and $>>$ stands for "much more informative than".

1.2.1.2 Theory of Probabilistic Syllogism

Probabilistic Syllogistics preserves the canonical form of the syllogisms: arguments composed by three terms, two premises and a conclusion relating two end terms (NT and MT) via a middle term (DT). Hence, the premises express the conditional probabilities among NT, MT and DT; the conclusion, qualified as p-valid[18], can be calculated if the constraints expressed in the premises are sufficient for calculating the probability of MT given the NT. The general schema for the four Aristotelian Figures is shown in Table 1.42, where α and β stand for the probability defined by the quantifiers and $f(\alpha,\beta)$ denotes the corresponding procedure for calculating the conditional probability between the end terms. It is worth noting that this framework does not distinguish between MP and NP but between *max*-premise and *min*-premise, the first one being the most informative and the second one the least informative.

As it pointed out previously, Oaskford and Chater do not focus principally on determining how the f of $f(\alpha,\beta)$ must be, since it can be calculated using probabilistic logic [45], but

[18] As has already been pointed out, this probabilistic framework is focused on a notion of validity different to the set-based approaches. Hence, the authors develop a specific concept of validity applied to probabilistic statements known as probabilistically-validity of *p-validity*.

Figure I	**Figure II**
$P(MT\|DT) = \alpha$	$P(DT\|MT) = \alpha$
$P(DT\|NT) = \beta$	$P(DT\|NT) = \beta$
$P(MT\|NT) = f(\alpha,\beta)$	$P(MT\|NT) = f(\alpha,\beta)$
Figure III	**Figure IV**
$P(MT\|DT) = \alpha$	$P(DT\|MT) = \alpha$
$P(NT\|DT) = \beta$	$P(NT\|DT) = \beta$
$P(MT\|NT) = f(\alpha,\beta)$	$P(MT\|NT) = f(\alpha,\beta)$

Table 1.42: Inference schema of syllogism under a probabilistic approach.

on the human strategies for calculating the conclusion of the syllogism; i.e., the heuristics to achieve $h(\alpha,\beta)$, that can match or not with $f(\alpha,\beta)$.

The probability heuristics model tries to deal with the 'fast and frugal' way of human beings for inferring conclusions. It is divided into two steps. The first being to order the premises according to their informativeness following (1.33). Once the premises are ordered, the second step is to apply the heuristic method to obtain the conclusion. The authors distinguish five methods [43, p. 217-219]:

***min*-heuristic** The quantifier of the conclusion is the same one of the least informative premise. It is the most preferred method.

***P*-entailments** It means probabilistically-entail and address, more or less, the subalternation of LSO. For instance, if the conclusion of a syllogism is "all crows are black", then the p-entail is "some crow is black". It is the second most preferred conclusion.

***attachment*-heuristic** It deals with the order of the terms in the conclusion. It states that there is a tendency towards selection the subject of one of the premises as the subject of the conclusion.

***max*-heuristic** It is applied to choose the most informative conclusion. It predicts that there should be a linear order in the frequency of the *min*-heuristic response dependent on the *max*-premise such that "all" > "some" > "some…not" > "none".

Syllogism	Heuristic
All *DT* are *MT*	*max*-premise
Some *NT* are *DT*	*min*-premise
Some-type conclusion	by *min*-heuristic
Some *NT* and *MT*	by *attachment*-heuristic

Table 1.43: Inference schema of syllogism AII (Figure I) under a probabilistic approach.

***some. . . not*-heuristic** It is applied to avoid producing or accepting “some. . . not” conclusions, as they are the least informative. Therefore, syllogisms with “all”, “some”, or “none” conclusions are more acceptable that those with “some. . . not” conclusions.

The first three methods are oriented toward generating the conclusions of syllogisms, while the last two are procedures for determining the reliability or confidence of the conclusion achieved. The authors illustrate the procedure with the AII mood of Figure I shown in Table 1.43.

1.2.1.3 The Case of Non-Proportional Quantifiers

As has been shown until now, Oaksford and Chater introduce new quantifiers defined as probabilities, but considering exclusively proportional ones (i.e.; “most” and ”few”). Other quantifiers, such as absolute (“five”, “twenty-five”, etc.) or exception ones (“all but three”, “all but five”, etc.) are not explicitly dealt with.

B. Geurts [23] points out precisely this question as a weakness of this system, as TGQ involves more quantifiers than the proportional ones. The starting point is the difficulty in interpreting an absolute quantifier in a probabilistic interpretation of binary quantified statements; for instance, what is the probabilistic interpretation of “five students are tall”?

Answering this question, Oaksford and Chater dismiss the introduction of non-proportional quantifiers saying that:

1. TGQ is based on a second order logic. Second-order logics, in general, are not completely decidable and, therefore, there are propositions that cannot be proved or rejected. There is an exception with monadic second-order logic, which is decidable but it can only express propositions composed by a single argument such as “Sherlock Holmes is a consulting detective”. However, most of the quantified statements that human be-

ing manage are binary; i.e., “few people are consulting detectives” and some of them cannot be decidable.

2. A direct relationship between proportional and absolute quantifiers can be established if the size or the cardinality of the subject is known. For instance, let us consider a set of students with $|students| = 10$, “two students are tall” can be easily read as “some students are tall”, “few students are tall” or “almost all students are not tall”. For that reasoning, the effort of introducing absolute quantifiers in a probabilistic framework is not worthwhile.

3. The use of precise absolute quantifiers is very reduced in our daily use of natural language as they only appear in very specific contexts. Furthermore, precise assertions do not generate useful inferences if people can make them using proportional quantifiers.

4. Geurts’ critique of probabilistic heuristic model is supported by formal arguments. He does not offer the corresponding psychological experiments to justify his arguments.

To the aim of this research, the contra-arguments explained by Oaksford and Chater are relevant and must be analysed in detail. Hence;

1. It is true that no monadic second-order logic is not completely decidable. However, there are many statements that are decidable and, very likely, the formalization of most of quantified statements that we use in our daily life belong to this class. Hence, TGQ cannot be disregarded as an adequate tool for the analysis of syllogistic reasoning even though it is a second-order logic.

2. The direct relationship between proportional and absolute quantifiers demands additional information that may not be available in all cases. Furthermore, it is possible to build very simple and interesting arguments that combine both kind of quantifiers; for instance, “almost all students are tall and twenty-five students are blond; therefore, most blond students are tall”. It is not difficult to build a counter-example of this syllogism (the non-tall students are the twenty-five blond students), but assuming as previous conditions an independence between both sets and a uniform distribution of the elements, the conclusion is acceptable.

3. The model of syllogistic reasoning that we propose in this work tries to deal with common and uncommon situations. For this reason, although absolute and precise quanti-

fiers are less common than proportional ones in many contexts, we assume the possibility of managing them jointly as a challenge.

4. The research of the present work does not belong to the field of psychology of reasoning but to the field of logic and computer science. Thus, the validity of a syllogistic model must be supported in other kind of experiments and the obtained results should not be necessarily compared with psychological experiments.

1.2.1.4 Summary

Probabilistic Syllogistics interprets quantified statements in terms of conditionals propositions where the quantifiers denote the strength of the conditional in terms of percentages. "All" and "no" are the bounds (100% and 0%, respectively) and quantifiers such as "most", "few", etc. are percentages between them. Nevertheless, the interpretation of other quantifiers defined by TGQ, such as absolute or exception ones, is not so clear in terms of probabilities.

This system emphasizes the psychological and heuristic dimension of syllogism instead of the logical one. Aristotelian Syllogistics is deeply analysed and the compatibility between them its checked, although other possible inferences schemas are not considered. Nonetheless, an interesting contribution of this system is the relevance of the background of the individuals or the implicit premises in the arguments to achieve acceptable inferences.

Although the proposal of this memory follows the formal and logic vein, there two interesting ideas that will be fruitful:

– The human inference process can also be coherently interpreted as a heuristic process. The premises are constraints in a searching space of possible solutions, where the conclusion is one of the possible solutions (eventually the optimal one) and the reasoning rules are not deductive logical rules but strategies for checking the compatibility between the premises and the conclusion.

– With a number of additional assumptions, it is easy to establish an equivalence between set-based approaches and probabilistic ones, therefore, it is possible to develop a model of syllogistic reasoning that can also manage knowledge expressed by probabilities.

1.2.2 Support Logic Syllogistics

M. Spies [79] proposes another system of syllogistic reasoning that is known as Support Logic Syllogistics. It is based on a conditional interpretation of quantified statements where the quantifier indicates the degree of evidence or belief that supports the conditional. This approach is based on Dempster-Shafer's Theory of Evidence [15, 72], a discipline that belongs to the field of formal study of beliefs; i.e., epistemology. Compared with probability, the main point of theories of evidence is that the strength of a proposition depends on the knowledge of the agent rather than on objective evidence.

The degree of belief of any quantified statement is a value that can be divided into three jointly exhaustive parts. Thus, a speaker's knowledge or belief about a proposition Γ entails the following:

1. Usually, an agent has a degree of belief in the truthfulness of a proposition. For instance, considering the adventures of Sherlock Holmes, J. Watson's degree of belief that S. Holmes has died in *Reichenbachfall* equals 0.8. This is the part that speaks in favour of Γ (the proposition that express the belief of J. Watson); i.e. $Bel(\Gamma)$.

2. The same agent also has a degree of belief in the truthfulness of the opposite proposition of Γ, that is $W \backslash \Gamma$. For instance, J. Watson's degree of belief that S. Holmes has not died in *Reichenbachfall* equals 0.1. This is the part that speaks against Γ; i.e. $Bel(W \backslash \Gamma)$.

3. Finally, a part that neither speaks in favour of nor against Γ can also appear. Thus, considering the previous examples, where $Bel(\Gamma) = 0.8$ and $Bel(W \backslash \Gamma) = 0.1$, the addition is less than 1 (0.9) and, therefore, there is a neutral part, *Ignorance* (I) ($I(\Gamma) = 0.1$), that must be distributed among both parts. If the addition of $Bel(\Gamma)$ and $Bel(W \backslash \Gamma)$ equals 1, there is no ignorance.

Over these three concepts, a new concept known as "degree of plausibility" ($P(\Gamma)$) rises. It quantifies the part of the agent's knowledge compatible with the the proposition Γ; i.e., the addition of the degree of belief and ignorance in Γ obtaining $P(\Gamma) = 1 - Bel(W \backslash \Gamma)$.

For instance, J. Watson has asserted the statement Γ. The $P(\Gamma)$, J. Watson's belief, is $P(\Gamma) = 1 - 0.1$ ($P(\Gamma) = 1 - (Bel(W \backslash \Gamma))$); that is, $P(\Gamma) = 0.9$. This value can match $Bel(\Gamma)$ or not; in this case it is not since $I(\Gamma) = 0.1$. It is worth noting that in the case of a proposition Γ with $P(\Gamma) = 0$, then $P(\Gamma) = Bel(\Gamma)$; furthermore, $Bel(\Gamma) + Bel(W \backslash \Gamma) = 1$. Only in the cases where $I > 0$, $Bel(\Gamma) + Bel(W \backslash \Gamma) < 1$ and $P(\Gamma) + Bel(W \backslash \Gamma) > 1$.

This section is organized as follows: section 1.2.2.1 describes the structure of quantified statements expressed as conditional statements using Dempster-Shafer's Theory; section 1.2.2.2 describes the proposed syllogistic inference schemas and, finally, section 1.2.2.3 summarizes the main points of this framework emphasizing its strengths and weakness.

1.2.2.1 Theory of Support Logic Categorical Statement

Categorical statements in Support Logic follow the standard form of quantified statements expressed as conditional statements, i.e., *if x is A then x is B* $[a,b]$, where $[a,b]$ denotes the strength that links the antecedent with the consequent, *if x is A* the subject-term and *then x is B* the predicate-term. In this case, the strength is the degree of belief of the speaker in the truthfulness of the proposition and constitutes its *degree of support*. Precise quantifiers denote precise values of belief and fuzzy quantifiers fuzzy ones. For instance, "most basket players are tall" is equivalent to "if a human being is a basket player then he is tall with a suggestive evidence that is most". "Most" conveys a high quantity or a strong link between *basket players* and *tall people*, therefore, given that it is a fuzzy quantifier, an interpretation is the interval $most = [0.8, 1]$. Maximum degree of support is denoted by the quantifier "all" and minimum support is denoted by the quantifier "some not".

$Bel(\Gamma)$ and $P(\Gamma)$ constitute the *support interval* of propositions and it is represented trough the interval $[\alpha, b]$, where α stands for $Bel(\Gamma)$ and b stands for $P(\Gamma)$. Furthermore, it is relevant to note that $Bel(\Gamma) \leq P(\Gamma)$. Thus, each one of the values of the support interval has associated a measure of probability with the following meanings:

$Bel(\Gamma)$ It is the lower bound. It denotes a subjective probability; i.e., how is the relevant for the speaker the evidence that supports a particular event or fact.

$P(\Gamma)$ It is the upper bound. It denotes an objective probability; i.e., the plausibility of the event in relation with the other contemporary facts or events.

Thus, considering again the example "if a human being is a basket player then he is tall with a suggestive evidence that is *most*", we can assign to the quantifier "most" the support value $[0.8, 1]$, given a usual interpretation for this quantifier, obtaining "if a human being is a basket player then he is tall $[0.8, 1]$".

PR1	If q then r $[a,b]$
PR2	If p then q $[c,d]$
C	If p then r $[ac, 1-c(1-b)]$

Table 1.44: Chaining/Property Inheritance based on Support Logic.

PR1	If p then q $[a,b]$
PR2	If $p \wedge q$ then r $[c,d]$
C	If p then $r \wedge q$ $[b(1-a(1-d))]$

Table 1.45: Intersection/Product based on Support Logic.

1.2.2.2 Theory of Support Logic Syllogism

In order to build a full system for syllogistic reasoning, Spies defines a negation operator and a two-place connective:

Negation operator This is defined to negate the support values of a statement. Let us consider the statement $A[a,b]$; $Bel(Non-A) = 1-P(A)$ and $P(Non-A) = 1-Bel(A)$; the corresponding support interval is $[1-b, 1-a]$.

Two-place connective This is a combination for assessing the "joint" degrees of belief and plausibility of a set of statements that constitute the premises and the conclusion of a certain syllogism. Of course, all the statements involved in a syllogism must be defined by the corresponding support intervals.

Spies [79] deals with four of Zadeh's patterns (see 1.1.5.2), including asymmetric and symmetric ones. The notation used is explained as follows: p, q, r stands for propositions and the particles *if... then* denotes the typical form a conditional statement:

- **Chaining or Property Inheritance**, it is an asymmetric syllogism. Table 1.44 shows it schema.
- **Intersection/Product**, it is also an asymmetric syllogism. Table 1.45 shows its structure.
- **Antecedent Conjunction**, it is a symmetric syllogism. Two additional constraints must be assumed for reasoning with it: i) $p[1,1]$ and $q[1,1]$; ii) v and w are computed according to Dempster's rule applied to the granules with support pairs of $[a,b]$ and $[c,d]$, respectively. Table 1.46 shows its form.

PR1	If p then r $[a,b]$
PR2	If q then r $[c,d]$
C	If $p \wedge q$ then r $[v,w]$

Table 1.46: Antecedent Conjunction based on Support Logic.

PR1	If r then p $[a,b]$
PR2	If r then q $[c,d]$
C	If r then $p \wedge q$ $[ac,bd]$

Table 1.47: Consequent Conjunction based on Support Logic.

– **Consequent Conjunction**; this is a symmetric syllogism. Table 1.47 shows its schema.

1.2.2.3 Summary

Support Logic Syllogistics also adopts a conditional reading of quantified statements but measuring the strength between the antecedent and the consequent with Dempster-Shafer's theory of evidence. From an epistemological point of view, it is more intuitive than the probability theory as it evaluates the subjective evidence that supports a statement but, from an intersubjective point of view, its results are less intuitive as the individual motivations for a belief are not always transparent.

Attending to inference schemas, it is not a solid model. Although it deals with four reasoning patterns comprising by three terms, two premises and a conclusion, it is not compatible with Aristotelian Syllogistics as it only deals with Zadeh's schemas which, as we showed in section 1.1.5, are incompatible with classical moods. Furthermore, only the use of proportional quantifiers is considered and, while the results of the Probabilistic Syllogistics match with the set-based approaches, Support Logic Syllogistics can give way to different results.

1.2.3 Qualified Syllogistics

D. Schwartz [71] analyses the syllogism within his general framework of Dynamic System Reasoning (DRS), a global approach to the field of non monotonic reasoning. Schwartz defines it as [71, p. 103]:

Quantification	Usuality	Likelihood
all	always	certainly
almost all	almost always	almost certainly
most	usually	likely
many/about half	frequently/often	uncertain/about 50-50
few/some	ocasionally/seldom	unlikely
almost no	almost never/rarely	almost certainly not
no	never	certainly not

Table 1.48: Interrelations across several level of modifiers.

> A *qualified syllogism* is a classical Aristotelian syllogism that has been "qualified" through the use of fuzzy quantifiers, likelihood modifiers, and usuality modifiers.

It is worth noting that Schwartz's framework opens syllogism up to likelihood and usuality modifiers and not only consider quantifiers. Each one of them has different uses in natural language, depending on what it is intended to be said:

– Quantifiers: They are used for describing current affairs or a present state of the world. For instance, "today, most birds can fly".

– Usuality modifiers: They are used to show a regularity in past experiences. For instance, "usually birds fly".

– Likelihood modifiers: They are used to make predictions about the future or for describing our expectations about it. For instance, "if at some future time I randomly select a bird, it is likely that it will be able to fly".

The interrelationships across several levels of this type of modifier [71, p.118] are shown in Table 1.48.

These relationships also can be transferred to the field of reasoning, as "there is a natural connection between fuzzy quantification and fuzzy likelihood [...]. The foregoing two concepts are also connected with the concept of usuality" [71, pp.116-117] through the notion of statistical sampling.

This section is organized as follows: section 1.2.3.1 describes the form of qualified categorical statements; section 1.2.3.2 describes the main characteristics of his inference schemas and, finally, section 1.2.3.3 summarizes the strengths and weakness of this model.

PR1	Most men are vain
PR2	Socrates is a man
C	It is likely that Socrates is vain

Table 1.49: Example of qualified syllogism.

1.2.3.1 Theory of Qualified Categorical Statements

The structure of Qualified Categorical Statements does not include relevant additional contributions with respect to the conditional reading proposals described in sections 1.2.1 and 1.2.2. However, he also distinguishes between the two semantics explained in this chapter; i.e., conditional interpretation and set-based one:

- Bayesian semantics: It is supported on the Bayesian framework of probability and it is the most general one. Therefore, the notion of conditional probability is assumed as primitive and also a subjectivist theory of probability is adopted.

- Counting semantics: It implicitly refers to a domain of discourse. For instance, "most birds can fly" refers to a domain that includes a collection of birds. The subjectivist theory of probability and the notion of conditional probability are avoided and substituted by the standard definition of probability: events are represented as subsets of a universe of alternative possibilities. In the fuzzy case, $\Sigma - count$ is the most used proposal.

We classify this approach as conditional reading as Schwartz says that Bayesian semantics is the most general one.

1.2.3.2 Theory of Qualified Syllogism

Schwartz explicitly talks about Aristotelian syllogisms, but his proposal is based on a type of argument that is not properly a syllogism. Table 1.49 shows the typical example of qualified syllogism defined by Schwartz.

This argument does not follow the minimum characteristics of Aristotelian syllogisms as it involves a singular statement (PR2, "Socrates is a man") instead only categorical ones, since *Socrates* refers to an individual and not to a set. We can think that Schwartz assumes the equivalence between singular statements and A statements, but there is no reference to this question or to the existential import. As a consequence, there are not three terms involved in the syllogism but only two; i.e., *men* and *vain*. Finally, the inference pattern underlying this

PR1	If x is S, then x is P $[m]$
PR2	a is S
C	a is P[m]

Table 1.50: Canonical form of qualified syllogism.

example is based on *Modus Ponens* (see Table 1.50), although the explained modifiers ($[m]$) are incorporated.

1.2.3.3 Summary

Qualified Syllogistics is only a small part of a global approach to DRS. Its main contribution is that of establishing relationships of equivalence between quantifiers, usuality and likelihood modifiers. However, its analysis of quantifiers is reduced to a small set of labels as examples, without additional considerations about them. Furthermore, there is no clear difference between categorical and singular statements, something fundamental in the previous models of syllogism.

From the point of view of reasoning pattern, he directly considers syllogism as a particular case of *Modus Ponens*, which is modified introducing the modifiers previously defined.

1.3 Discussion: Set-Based Interpretation vs. Conditional Interpretation

Syllogistic reasoning can be briefly defined as reasoning about quantities (fuzzy or not) that are usually expressed as quantified statements. Thus, this chapter has sought to present a summary of the main approaches to syllogistic reasoning developed in the history of logic and lay the foundations for a further development.

We have described two different frameworks for managing this kind of inferences: the so-called set-based interpretation and the conditional one. The difference between them is in the analysis of quantified statements, although both schemes start from the same basic pattern *Q S are P*, where Q is the quantifier, S the subject and P the predicate.

Thus, the set-based approach establishes that S and P are sets of elements (known as term-sets) and the quantifier Q describes the relationship between both sets in terms of quantities (fuzzy or not); for instance, "most basket players are tall" means that a number of elements big enough of the set *basket players* belongs to the set of *tall people*. Since this is a relationship between sets, we are talking about a second order logic. Following this idea, we have presented six approaches which are the most relevant ones from our point of view. Aristotelian, Intermediate and Exception Syllogistics explores syllogistic reasoning considering crisp and vague quantifiers but out of a fuzzy framework while Interval, Fuzzy and Generalized Intermediate Syllogistics manage vague quantifiers using fuzzy logic.

Aristotelian Syllogistics is the first system in the history of logic and laid the foundations for the subsequent developments. He defines categorical statements as the only valid propositions for syllogistic reasoning, where the S and P are term-sets and Q the quantifier. He also defines the four quantifiers valid for making inferences ("all" and "no" are the universal ones and "some" and "some... not" are the particular or existential ones) and how they are linked between themselves (the LSO). The basic structure of the inference pattern is the chaining of two terms (the so called *NT* and *MT*) by a middle one (the so called *DT*); therefore, every syllogism comprises three terms, two premises and a conclusion. The combination of the position of the middle term in the premises (the so-called *Figures*) with the quantifiers of the LSO generates the 24 Aristotelian moods.

In twentieth century, coinciding with a recovery of Aristotelian thinking, an expansion of Aristotelian Syllogistics was proposed by introducing the concept of intermediate quantifier. This is defined as any quantifier between the universal and the particular ones such as "most", "few", "many", etc. Intermediate Syllogistics produces new LSOs with more quantifiers and, hence, new syllogistic moods appear. Furthermore, it also opens up the possibility of an algebraic method supported by Venn Diagrams to calculate the quantifier of the conclusion, thus complementing the classical proof method based on logical laws. This idea is consolidated in Exception Syllogistics, where a full algebraic method for syllogism is proposed. In [77] another algebraic method for syllogism appears, but only for the four Aristotelian quantifiers, and its results can be integrated in Exception Syllogistics. In addition to this method, Exception Syllogistics also introduces exception quantifiers (such as "all but three", "at most five", etc.), a relevant innovation with respect to the Intermediate Syllogistics (only focused on proportional quantifiers), and the combination thereof with proportional ones.

In the fuzzy field, Interval Syllogistics introduce the representation of quantifiers as intervals, an intermediate step to a full fuzzy one. In this case, linguistic quantifiers are numerically interpreted and the necessity of an LSO is avoided. On one hand, is a more precise proposal as it can manage precise and imprecise quantifiers simultaneously but on the other hand, it loses a degree of interpretability as the translation from natural language to percentages or intervals is not unambiguous. In this system, the quantifier conclusion is obtained through a calculation procedure where the premises are the constraints and the conclusion must be compatible with them. In this way, the concept of logic proof is not given up. Fuzzy Syllogistics follows this vein exactly, but using the following fuzzy definition for the quantifiers: each one has a fuzzy number associated to it and the quantifier of the conclusion is directly calculated using fuzzy arithmetic. Finally, Generalized Intermediate Syllogistics is a generalization of Intermediate Syllogistics to a fuzzy framework, but without relevant innovations in either in the quantifiers managed (although there is it in how they are defined) or in the syllogistic inferences schema.

In general, set-based approaches stand out in dealing with quantified statements as assertions; i.e., propositions that affirm or deny something although they involve fuzzy quantifiers. This is a relevant question from a pragmatic point of view, since an assertion is different of a rule or a condition. Furthermore, all these frameworks assimilate the logical form of the statements to their syntactical form, making their analysis easier. Non-fuzzy frameworks also involve relationships of opposition typical in natural language as logical rules (the links in the LSO), while fuzzy frameworks substitute them with the concepts, such as order or negation, which are also very typical in natural language. Another interesting innovation in fuzzy models is the transition from the syllogism based on logic proofs to a heuristic syllogism built using an algebraic method for calculating the quantifier of the conclusion in terms of the quantifiers of the premises.

Nevertheless, set-based approaches also have a number of weaknesses that must be considered. Firstly, the question of existential import of universal quantifiers, which has been under debate since the Middle Ages. There is not one single solution, although the most widely accepted one is to assume it as an additional restriction of the syllogism, saying that the subject-term of universal propositions is not empty. Secondly, introducing statements involving singular terms such as “Socrates” or “S. Holmes” into a syllogism is a problematic question, although from an intuitive point of view, inferences such as “all human beings are mortal, Socrates is a human being; therefore, Socrates is mortal” are very clear. The accepted solution is to transform singular statements into categorical statements qualifying the singular term with the universal quantifier “all” to emphasize that it is unique and transforming it into a categorical term; i.e., “all human beings are mortal, all Socrates is a human being; therefore, all Socrates is mortal”. However, from the point of view of natural language use, the statement is strange. Finally, other weak point of fuzzy approaches (except Generalized

Intermediate Syllogistics) is that their compatibility with the Aristotelian one is only partial, although Interval Syllogitiscs shows correct behaviour.

Conditional interpretation assumes that quantified statements can be rewritten as implications, where *S* is the antecedent, *P* the consequent and *Q* denotes the strength of the link between *S* and *P*. For instance, "most basket players are tall" is equivalent to "if *somebody is a basket player* then *somebody is tall* with *most* support". The main difference with respect to set-based interpretation is that properties and individuals are distinguished here; i.e., in terms of quantified statements, there is no difference between categorical and singular statements. This supplements one of the weak points of the other syllogistic frameworks. The three frameworks analysed following this interpretation differ in their semantics to quantified statements: Probabilistic Syllogistics assumes that a quantifier expresses probability; Support Logic Syllogistics assumes that it express belief; and Qualified Syllogistics assumes that it can express quantity, temporality or probability.

Probabilistic Syllogistics is based on the psychology of reasoning rather than on logic. This leads it to analyse syllogistic reasoning as a heuristic procedure in terms of Bayesian reasoning instead of as a logic one, substituting the concept of logic validity for reliability. It makes a deep analysis of Aristotelian Syllogistics and the compatibility between them is proved. Examples of intermediate quantifiers are also considered and the different psychological strategies for making inferences are explained. Support Logic Syllogistics substitutes probability theory for Dempster-Shafer's theory of evidence. Aristotelian Syllogistics is not compatible with this system as only Zadeh's inference schemas (Fuzzy Syllogistics) are considered. Finally, Qualified Syllogistics introduces the suggestive idea about the links or equivalences between quantifiers, usuality and likelihood modifiers, although syllogistic arguments are transformed into *Modus Ponens* substituting the Minor Premise of the syllogism for a singular statement.

In general, the approaches based on conditional interpretation deal with the combination of categorical statements and singular statements better than set-based approaches and facilitate the analysis of this kind of arguments, although they are not strictly syllogistic. The problematic question of existential import is also avoided as the statements managed are not assertions but rules in conditional form. Furthermore, Probabilistic Syllogistics uses Bayesian probability, which is well known and its results are very successful. Nevertheless, all these approaches only deal with proportional quantifiers; other quantifiers, such us absolute, exception or comparative, are explicitly rejected, as we have explained in Probabilistic Syllogistics.

On the other hand, the transformation of quantified statements into conditional propositions also entails transforming assertions into rules or conditions generating certain imprecision or vagueness not present in the original statement. For instance, “all students are tall” describes the fact the *all students are tall*[19] while “if somebody if a student then is sure that is tall” describes a hypothetical fact that is test when somebody in particular is chosen. Finally, the relationships of the LSO cannot be adequately explained using a probabilistic interpretation either, as the different kind of oppositions are not considered.

For all the reasons explained, the framework defined by set-based interpretation seems to be the most suitable for the purposes of our research: developing a proposal of syllogistic reasoning that is compatible with TGQ and flexible for dealing with different patterns of inference that can appear in natural language. We mainly focus on the development of a strategy for calculating the quantifier of the conclusion in terms of the quantifiers of the premises rather than a logic test to check its validity. In chapter 2 we go on to describe the model that we propose to achieve this aim.

[19] We are assume the Tarki’s Theory of Truth [80].

Chapter 2

Syllogistic Reasoning Using the Theory of Generalized Quantifiers

In the previous chapter, different frameworks for dealing with syllogistic reasoning were analysed. In this analysis, it is possible to see the evolution from the twenty-four Aristotelian moods, managing only four quantifiers and four inference patterns, to fuzzy approaches, which can manage vague quantifiers and more than a hundred inference patterns, where the chaining is not only based on a middle term but also in other logic operations, such us conjunction and disjunction. However, we have also concluded that there are a number of weak points that can be improved to increase natural language arguments that can be considered: i) other quantifiers that we manage into natural language must be incorporated into syllogistic reasoning; ii) the combination of different kind of quantifiers in the same syllogism; iii) dealing with complex arguments involving *m* terms, *n* premises and *p* conclusions; and, iv) managing crisp and fuzzy definitions of different kind of quantifiers.

As constraints, we assume that Aristotelian Syllogistics must be included as a particular case of this broad syllogistic framework and the compatibility with TGQ, the current linguistic theory for dealing with quantifiers in natural language. Furthermore, both challenges are not incompatible given the relationship between them explained by D. Westersthål [86].

As we have already stated, we consider the set-based interpretation to be the one best suited to our aims. The main reason for this choice is its compatibility with the TGQ [8]. It is the standard theory accepted by logicians and linguistics for dealing with quantifier expressions and analysing quantification in natural language combining the logic and the linguistics perspectives. Its key concept is the *generalized quantifier*, understood as a second order predicate that establishes a relationship between two classical sets. Assuming this theory opens up the possibility of going much further than the usual absolute/relative quantifiers. With regard to the reasoning process, a heuristic perspective is adopted rather than a logical one as it makes it easier to deal with arguments involving more than three terms and two premises.

This chapter is organized as follows: section 2.1 describes the structure of the statements that are used in our proposal of syllogistic framework; section 2.2 describes our proposal of reasoning schema and inference mechanism; section 2.3 shows some illustrative examples about how our framework works; section 2.4 explains the combination between absolute and proportional quantifiers; and, finally, section 2.5, summarizes the main characteristics of this framework and analyses its strengths and weakness.

2.1 Theory of Categorical Statements Using the Theory of Generalized Quantifiers

As we have already stated, the standard form of the quantified statements valid for syllogism according to set-based interpretation is *Q S are P*, where Q is the quantifier, S the subject-term and P the predicate-term. Depending on the particular syllogistics, Q could adopt different forms, although always binary (taking two sets as arguments); for instance, in Intermediate Syllogistics it is quantifiers such as "most", "few", etc. while in Exception Syllogistics it is terms such as "all but four", "at least two", etc. Our contribution in this topic is to use TGQ to define the form of Q. As examples, we describe the case of proportional, absolute and exception quantifiers in detail.

Considering the proportional interpretation of the quantifier "all", and denoting E as the referential universe and $\mathscr{P}(E)$ as the power set of E, the evaluation of the quantified statement "All Y_1 are Y_2" for $Y_1, Y_2 \in \mathscr{P}(E)$ can be modelled as:

$$
\begin{aligned}
& all : \mathscr{P}(E) \times \mathscr{P}(E) \rightarrow \{0,1\} \\
& \qquad\qquad (Y_1, Y_2) \rightarrow All(Y_1, Y_2) = \begin{cases} 0 : if & Y_1 \nsubseteq Y_2 \\ 1 : if & Y_1 \subseteq Y_2 \end{cases}
\end{aligned}
$$

For absolute quantifiers, evaluation of a statement like "Between 3 and 6 Y_1 are Y_2" ($Bet_{3,6}$) can be modelled as:

$$Bet_{3,6} : \mathscr{P}(E) \times \mathscr{P}(E) \to \{0,1\}$$

$$(Y_1,Y_2) \to Bet_{3,6}(Y_1,Y_2) = \begin{cases} 0 : if & |Y_1 \cap Y_2| \notin [3,6] \\ 1 : if & |Y_1 \cap Y_2| \in [3,6] \end{cases}$$

For quantifiers of exception, let us consider the sentence "All but three Y_1 are Y_2"; for $Y_1, Y_2 \in \mathscr{P}(E)$. Its evaluation can be modelled as:

$$\textit{All but three} : \mathscr{P}(E) \times \mathscr{P}(E) \to \{0,1\}$$

$$(Y_1,Y_2) \to \textit{All but three}(Y_1,Y_2) = \begin{cases} 0 : if & |Y_1 \cap \overline{Y_2}| \neq 3 \\ 1 : if & |Y_1 \cap \overline{Y_2}| = 3 \end{cases}$$

In general, for any quantified proposition whose quantifier is a binary one from TGQ; let $P = \{P_s,\, s = 1\ldots,S\}$ be the set of relevant properties defined in E. For instance, P_1 can denote the property "to be a student" and P_2 "to be tall". Let L_1 and L_2 be any Boolean combination of the properties in P. Thus, the typical form of our categorical statements is:

$$Q\ L_1 \text{ are } L_2 \tag{2.1}$$

where Q denotes a binary quantifier of TGQ, L_1 is the subject-term or restriction and L_2 the predicate-term or scope. It is worth noting that L_1 and L_2 can be any Boolean combination of P. Thus, we can incorporate to our proposal the inferences based on logic operators (the so-called symmetric syllogisms). We consider the following types of quantifiers:

- **Logical quantifiers (Q_{LQ})**: The four Aristotelian ones, "all", "no", "some", "not all" ("some... not").
- **Absolute Binary quantifiers (Q_{AB})**: Natural numbers with or without some type of modifier such us "around five", "more than twenty-five", etc.) following the general pattern Q_{AB} Y_1 *are* Y_2 (e.g., "around five students are tall").
- **Proportional Binary quantifiers (Q_{PB})**: Linguistic terms such us "most", "almost all", "few", etc. following the general pattern Q_{PB} Y_1 *are* Y_2 (e.g. "few students are tall").

– **Binary Quantifiers of Exception** (Q_{EB}): Proportional binary quantifiers with a natural number such us “all but three”, “all but three or four”, etc. following the general pattern Q_{EB} Y_1 *are* Y_2 (e.g., “all but five students are tall”).

– **Absolute Comparative Binary Quantifiers** (Q_{CB-ABS}): Natural numbers with modifiers of type *more, less,* such us (“three more... than”, “four less... than”, etc.) following the general pattern *There is/are* Q_{CB-ABS} Y_1 *than* Y_2 (e.g. “there are three more boys than girls”).

– **Proportional Comparative Binary Quantifiers** ($Q_{CB-PROP}$): Rational multiple or partitive numbers, such us “double”, “half”, etc. following the general pattern *There is/are* $Q_{CB-PROP}$ Y_1 *than* Y_2 (e.g., “there are double boys than girls”).

– **Similarity Quantifiers** (Q_S): Linguistic expressions that denote similarity (S) between two given sets “very similar”, “few similar”, etc. following the general pattern A *and* A' *are very/few/... similar*; where A denotes one of the sets of the comparison and A' denotes a set similar to A (e.g. “the audience of opera and ballet are very similar”).

Table 2.1 summarizes the definition of these quantifiers according to the TGQ.

2.2 Theory of Syllogism Using the Theory of Generalized Quantifiers

In the analysis conducted in chapter 1, we pointed out that some models, such as Aristotelian Syllogistics, focus on developing strategies to prove the logical validity of the syllogisms, while other models, such us Exception Syllogistics or Interval Syllogistics, focus on developing methods to calculate the conclusion in terms of the quantifiers of the premises. The fact that the former only consider one conclusion by argument and the latter consider several different possibilities justifies these different strategies. Furthermore, as pointed out by Probabilistic Syllogistics, this second strategy, which we qualify as heuristic, is also typical in human reasoning.

Thus, assuming a heuristic challenge and adhering mainly to Interval Syllogistics (see section 1.1.2), as it has shown the best behaviour for dealing with Aristotelian Syllogistics, we propose a transformation of the syllogistic reasoning into an equivalent optimization problem that consists of calculating the extreme quantifiers of the conclusion taking as restrictions the quantifiers of the premises.

Logical quantifiers

$$Q_{LQ-all}(Y_1 Y_2) = \begin{cases} 0: & Y_1 \not\subseteq Y_2 \\ 1: & Y_1 \subseteq Y_2 \end{cases}$$

$$Q_{LQ-no}(Y_1, Y_2) = \begin{cases} 0: & Y_1 \cap Y_2 \neq \emptyset \\ 1: & Y_1 \cap Y_2 = \emptyset \end{cases}$$

$$Q_{LQ-some}(Y_1, Y_2) = \begin{cases} 0: & Y_1 \cap Y_2 = \emptyset \\ 1: & Y_1 \cap Y_2 \neq \emptyset \end{cases}$$

$$Q_{LQ-not-all}(Y_1, Y_2) = \begin{cases} 0: & Y_1 \subseteq Y_2 \\ 1: & Y_1 \not\subseteq Y_2 \end{cases}$$

Absolute binary quantifiers

$$Q_{AB}(Y_1, Y_2) = \begin{cases} 0: & |Y_1 \cap Y_2| \neq N \\ 1: & |Y_1 \cap Y_2| = N \end{cases}$$

Proportional binary quantifiers

$$Q_{PB}(Y_1, Y_2) = \begin{cases} 0: & \frac{|Y_1 \cap Y_2|}{|Y_1|} \notin LT \\ 1: & \frac{|Y_1 \cap Y_2|}{|Y_1|} \in LT \\ 1: & |Y_1| = 0 \end{cases}$$

Exception binary quantifiers

$$Q_{EB}(Y_1, Y_2) = \begin{cases} 0: & |Y_1 \cap \overline{Y_2}| \notin N* \\ 1: & |Y_1 \cap \overline{Y_2}| \in N* \end{cases}$$

Comparative binary quantifiers

$$Q_{CB-ABS}(Y_1, Y_2) = \begin{cases} 0: & |Y_1| - |Y_2| \notin N* \\ 1: & |Y_1| - |Y_2| \in N \end{cases}$$

$$Q_{CB-PROP}(Y_1, Y_2) = \begin{cases} 0: & \frac{|Y_1|}{|Y_2|} \notin Q* \\ 1: & \frac{|Y_1|}{|Y_2|} \in Q* \end{cases}$$

Similarity quantifiers

$$Q_S(Y_1, Y_2) = \begin{cases} 0: & \frac{|Y_1 \cap Y_2|}{|Y_1 \cup Y_2|} < S, Y_1 \cup Y_2 \neq \emptyset, S \notin Q* \\ 1: & \frac{|Y_1 \cap Y_2|}{|Y_1 \cup Y_2|} \geq S, Y_1 \cup Y_2 \neq \emptyset, S \in Q* \\ 1: & Y_1 \cup Y_2 = \emptyset \end{cases}$$

Table 2.1: Definition of some binary quantifiers according to TGQ.

PR1:	$Q_1\ L_{1,1}$ are $L_{1,2}$
PR2:	$Q_2\ L_{2,1}$ are $L_{2,2}$
...	
PRN:	$Q_N\ L_{N,1}$ are $L_{N,2}$
C:	$Q_C\ L_{C,1}$ are $L_{C,2}$

Table 2.2: Inference scheme of broad syllogism.

(a) AAA mood.

PR1:	$Q_{all}\ L_{1,1}$ are $L_{1,2}$
PR2:	$Q_{all}\ L_{2,1}$ are $L_{1,1}$
C:	$Q_{all}\ L_{2,1}$ are $L_{1,2}$

(b) AII mood.

PR1:	$Q_{all}\ L_{1,1}$ are $L_{1,2}$
PR2:	$Q_{some}\ L_{1,1}$ are $L_{2,1}$
C:	$Q_{some}\ L_{2,1}$ are $L_{1,2}$

Table 2.3: Aristotelian moods expressed using broad inference pattern.

2.2.1 Inference Pattern

To achieve our challenge of developing an inference pattern flexible enough to deal with complex arguments composed of more than three terms and two premises, we have to deal with all the inference patterns developed by the different proposals of syllogisms.

So, let us now consider the form defined in the previous section to categorical statements, our proposal of broad inference pattern for N premises, $PR_n, n = 1,\dots,N$ and a conclusion, C, is described in Table 2.2; where Q_n, $n = 1,\dots,N$ are the linguistic quantifiers in the N premises; $L_{n,j}$, $n = 1,\dots,N$, $j = 1, 2$ denotes an arbitrary Boolean combination between the properties considered in the syllogism, Q_C stands for the quantifier of the conclusion and $L_{C,1}$ and $L_{C,2}$ stand for the subject-term and the predicate-term in the conclusion.

Let us mention a number of patterns in the analysed frameworks to syllogistic reasoning to show how this schema can be adapted.

From Aristotelian Syllogistics we take the AAA mood from Figure I and the AII mood from Figure III, the so-called axioms of this system according to J. Łukasiewicz [33]. Thus, since this framework only manages three terms, we have P_3; two premises, PR_2; a conclusion C and the four logic quantifiers, $Q_n \in$ {all, none, some, some...not}. Therefore, $L_{n,j}$ is with $n = 2$ and $j = 2$. Table 2.3 shows both schemas.

From Intermediate Syllogistics, we take, for instance, ETD from Figure II. As in Aristotelian Syllogistics, P_3, PR_2, C and $Q_n \in${all, almost all, most, many, no, few, most...not, many...not, some...not}. Therefore, $L_{n,j}$ is with $n = 2$ and $j = 2$. Table 2.4 shows its schema.

PR1:	Q_{none} $L_{1,1}$ are $L_{1,2}$
PR2:	Q_{most} $L_{2,1}$ are $L_{1,2}$
C:	Q_{most} $L_{2,1}$ are $not(L_{1,2})$

Table 2.4: ETD (Figure II) from Intermediate Syllogistics expressed using broad inference pattern.

PR1:	$Q_{at-least-all-but-0}$ $L_{1,1}$ are $L_{1,2}$
PR2:	$Q_{at-least-1}$ $L_{1,1}$ are $L_{2,1}$
C:	$Q_{at-least-1}$ $L_{2,1}$ are $L_{1,2}$

Table 2.5: AI- (Figure III) from Exception Syllogistics expressed using broad inference pattern.

PR1:	Q_1 $L_{1,1}$ are $L_{1,2}$
PR2:	Q_2 $L_{1,2}$ are $L_{1,1}$
PR3:	Q_3 $L_{1,2}$ are $L_{3,2}$
PR4:	Q_2 $L_{1,1}$ are $L_{3,2}$
C:	Q_C ($L_{1,1}$ and $L_{1,2}$) are $L_{3,2}$

Table 2.6: Pattern II from Interval Syllogistics expressed using broad inference pattern.

From Exception Syllogistics, we choose, for instance, the AI- schema of Figure III, that is, the corresponding to the AII schema of Aristotelian Syllogistics. In this case, P and PR are the same, but C can be multiple. The quantifiers are $Q_n \in$ {at least all but x, at most x, at least x, at least x…not} where x depends on the context and the cardinality of $L_{n,1}$. Table 2.5 shows the corresponding pattern, where C is the minimal conclusion.

All of these patterns are of asymmetric type; i.e., based on a chaining through a middle term. They only use negation as Boolean operation on the predicate of the statements. Fuzzy frameworks are the first ones to introduce symmetric syllogisms; i.e., those whose chaining is through logic operations such as conjunction or disjunction. This means that subject-terms and predicate-terms involve more Boolean operations than asymmetric frameworks. We shall now show some examples of how these inference patterns can be implemented in our proposal of broad reasoning pattern.

Hence, from Interval Syllogistics we took Pattern II, which involves a conjunction in the subject of the conclusion. Thus, P_3, PR_4, C and $Q_n = [\underline{q}, \overline{q}]$ with $[\underline{q}, \overline{q}] \in [0,1]$. Table 2.6 shows the corresponding schema.

PR1:	$Q_1\ L_{1,1}$ are $L_{1,2}$
PR2:	$Q_2\ (L_{1,1}\ and\ L_{1,2})$ are $L_{2,1}$
C:	$Q_C\ L_{1,1}$ are $((L_{1,2})\ and\ L_{2,1})$

Table 2.7: Intersection/product from Fuzzy Syllogistics expressed using broad inference pattern.

From Fuzzy Syllogistics, we took the most basic pattern defined by Zadeh, Intersection/product syllogism. In this case, P_3, $PR2$, C and Q_n are fuzzy numbers. The Boolean operator is a conjunction in the subject of $PR2$. Table 2.7 shows the corresponding schema.

In order to explain the inference mechanism to calculate C in terms of PR1 to PRN, we distinguish three steps: i) dividing the universe of discourse into disjoint sets; ii) defining sentences as systems of inequations; iii) selecting the optimization method that should be applied in each case in order to resolve the reasoning process.

2.2.2 Division of the Universe in Disjoint Sets

Following the ideas explained in sections 1.1.2 and 1.1.3, the referential universe E is partitioned into a new set of disjoint sets $PD = \{P'_1, \ldots, P'_K\}$ with $K = 2^S$ containing the elements in E that fulfil or not the S terms or properties defined in a syllogism in the following way:

$$\begin{aligned} P'_1 &= \overline{P_1} \cap \overline{P_2} \cap \ldots \overline{P_{S-1}} \cap \overline{P_S} \\ P'_2 &= \overline{P_1} \cap \overline{P_2} \cap \ldots \overline{P_{S-1}} \cap P_S \\ &\ldots \\ P'_K &= P_1 \cap P_2 \cap \ldots P_{S-1} \cap P_S \end{aligned} \tag{2.2}$$

where $P_s = \{e \in E,\ e\ fulfills\ P_s\}$ and $\overline{P_s} = \{e \in E,\ e\ does\ not\ fulfill\ P_s\}$. Therefore, $\bigcap_{k=1}^{K} P'_k = \emptyset$ and $\bigcup_{k=1}^{K} P'_k = E$.

Figure 2.1 shows an example for $S = 3$ properties, where the P'_k, $k = 1, \ldots, 8$ denote each one of the disjoint sets that are generated from the properties P_1, P_2 y P_3. Thus, for instance $P'_8 = P_1 \cap P_2 \cap P_3$, $P'_2 \cup P'_6 = \overline{P_2} \cap P_3$ or P_1=$P'_5 \cup P'_6 \cup P'_7 \cup P'_8$.

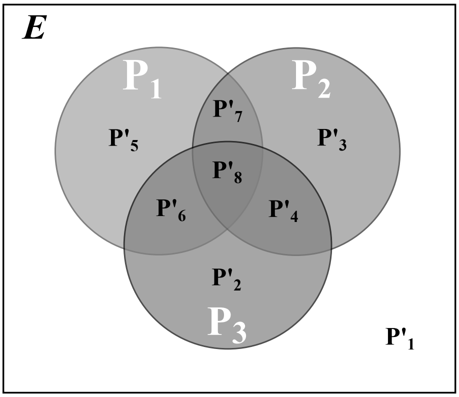

Figure 2.1: Division of the referential universe for the case of $S = 3$ properties and $8 = 2^3$ elements in the partition P'_k, $k = 1, \ldots, 8$. $P_1 = P'_5 \cup P'_6 \cup P'_7 \cup P'_8; P_2 = P'_3 \cup P'_4 \cup P'_7 \cup P'_8; P_3 = P'_2 \cup P'_4 \cup P'_6 \cup P'_8$.

2.2.3 Transformation into Inequations

The syllogistic reasoning problem is transformed into an equivalent optimization problem, disregarding the type of quantifiers involved; notwithstanding, we can deal at least with all types of quantifiers compatible with the linguistic TGQ (defined in Table 2.1).

We describe the transformation with quantifiers defined as crisp intervals ($Q = [a, b]$), since this is the basis for dealing with fuzzy quantifiers. For the sake of simplicity, we assume that the $L_{i,j}$ are atomic; i.e., they are not Boolean combinations of sets. This assumption does not limit the the generality of the solution that can be extended by simply and directly substituting the atomic parts by the Boolean combination of sets. For instance, for $S = 2$ we have the following combinations of disjoint sets:

$$\begin{aligned} P'_1 &= \overline{L_1} \cap \overline{L_2} \\ P'_2 &= L_1 \cap \overline{L_2} \\ P'_3 &= L_1 \cap L_2 \\ P'_4 &= \overline{L_1} \cap L_2 \end{aligned} \tag{2.3}$$

Where the usual notation L_1, L_2 for the terms in the propositions has been used instead of the one introduced in expression 2.2.

Logical definition		Equivalent inequation
$Q_{LQ-all}(Y_1Y_2) =$	$\begin{cases} 0: & Y_1 \not\subseteq Y_2 \\ 1: & Y_1 \subseteq Y_2 \end{cases}$	$x_2 = 0$
$Q_{LQ-no}(Y_1,Y_2) =$	$\begin{cases} 0: & Y_1 \cap Y_2 \neq \emptyset \\ 1: & Y_1 \cap Y_2 = \emptyset \end{cases}$	$x_3 = 0$
$Q_{LQ-some}(Y_1,Y_2) =$	$\begin{cases} 0: & Y_1 \cap Y_2 = \emptyset \\ 1: & Y_1 \cap Y_2 \neq \emptyset \end{cases}$	$x_3 > 0$
$Q_{LQ-not-all}(Y_1,Y_2) =$	$\begin{cases} 0: & Y_1 \subseteq Y_2 \\ 1: & Y_1 \not\subseteq Y_2 \end{cases}$	$x_2 > 0$
$Q_{AB}(Y_1,Y_2) =$	$\begin{cases} 0: & \lvert Y_1 \cap Y_2\rvert \notin [a,b] \\ 1: & \lvert Y_1 \cap Y_2\rvert \in [a,b] \end{cases}$	$x_3 \geq a; x_3 \leq b$
$Q_{PB}(Y_1,Y_2) =$	$\begin{cases} 0: & \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\rvert} < a \vee \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\rvert} > b \\ 1: & \lvert Y_1\rvert = 0 \\ 1: & \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\rvert} \geq a \wedge \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\rvert} \leq b \end{cases}$	$\frac{x_3}{x_2+x_3} \geq a;$ $\frac{x_3}{x_2+x_3} \leq b$
$Q_{EB}(Y_1,Y_2) =$	$\begin{cases} 0: & \lvert Y_1 \cap \overline{Y_2}\rvert \notin [a,b] \\ 1: & \lvert Y_1 \cap \overline{Y_2}\rvert \in [a,b] \end{cases}$	$x_2 \geq a; x_2 \leq b$
$Q_{CB-ABS}(Y_1,Y_2) =$	$\begin{cases} 0: & \lvert Y_1\rvert - \lvert Y_2\rvert \notin [a,b] \\ 1: & \lvert Y_1\rvert - \lvert Y_2\rvert \in [a,b] \end{cases}$	$(x_3+x_2)-(x_3+x_4) \geq a;$ $(x_3+x_2)-(x_3+x_4) \leq b$
$Q_{CB-PROP}(Y_1,Y_2) =$	$\begin{cases} 0: & \frac{\lvert Y_1\rvert}{\lvert Y_2\rvert} \notin [a,b] \\ 1: & \frac{\lvert Y_1\rvert}{\lvert Y_2\rvert} \in [a,b] \end{cases}$	$\frac{(x_3+x_2)}{(x_3+x_4)} \geq a; \frac{(x_3+x_2)}{(x_3+x_4)} \leq b$
$Q_S(Y_1,Y_2) =$	$\begin{cases} 0: & \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\cup Y_2\rvert} < a, Y_1 \cup Y_2 \neq \emptyset \\ 1: & \frac{\lvert Y_1\cap Y_2\rvert}{\lvert Y_1\cup Y_2\rvert} \geq a, Y_1 \cup Y_2 \neq \emptyset \\ 1: & Y_1 \cup Y_2 = \emptyset \end{cases}$	$\frac{x_3}{x_2+x_3+x_4} \geq a; \frac{x_3}{x_2+x_3+x_4} \leq b$

Table 2.8: Definitions of interval crisp Generalized Quantifiers $Q = [a,b]$

Denoting $x_k = |P'_k|$, $k = 1,\ldots,K$ we therefore have,

$$\begin{aligned} x_1 &= \left|\overline{L_1} \cap \overline{L_2}\right| \\ x_2 &= \left|L_1 \cap \overline{L_2}\right| \\ x_3 &= \left|L_1 \cap L_2\right| \\ x_4 &= \left|\overline{L_1} \cap L_2\right| \end{aligned} \tag{2.4}$$

Now, the evaluation of any quantified proposition is equivalent to solving an inequation where the previously defined cardinality values $x_k, k = 1,\ldots,K$ are involved as variables. Table 2.8 (right) summarizes the inequations corresponding to each quantifier.

Let us consider the following example, "between three and six Y_1 are Y_2". It entails the absolute binary quantifier defined as an interval; that is $Q_{AB} = [3,6]$. According to the definitions shown in Table 2.8, this means that $x_3 \geq 3$ and $x_3 \leq 6$ (see (2.3) and (2.4)).

2.2.4 Definition and Resolution of the Equivalent Optimization Problem

Once the inequations corresponding to the statements involved in a syllogism have been defined, we are in a position to approach the reasoning problem as an equivalent mathematical optimization problem. The fundamental idea of this transformation is described in [17, 18] for binary proportional quantifiers and consists of defining the set of inequations associated to the premises in the syllogism and solving it by applying of the appropriate resolution method (Simplex, fractional programming, etc.)[9]. However, our approach is more general since we can also deal with all types of quantifiers and also any Boolean combination in the restriction and scope of the quantified statements (premises and conclusions).

In order to correctly apply the resolution method, we need to add three additional constraints to the set of inequations obtained from the syllogistic argument. The first one guarantees that there are no sets with a negative number of elements; i.e., the existential import is assumed as a previous condition;

$$x_k \geq 0, \forall k = 1, \ldots, K = 2^S \tag{2.5}$$

The other two constraints are only necessary if the quantifier of the conclusion is a proportional one. In this case, since the function to be optimized is a rational one, it should be guaranteed that;

- there are no 0 in the denominator of the involved fractions in order to avoid any indefinition in the results; i.e., $L_{n,1} \neq \emptyset$; thus if we denote by $P_1'^{n,1} \cdots P_r'^{n,1}$ the disjoint parts of $L_{n,1}$, and by $x_r^{n,1}$ the cardinalities of the disjoint sets, the following must hold:

$$x_1^{n,1} + \cdots + x_r^{n,1} > 0, \forall n = 1, ..., N \tag{2.6}$$

- the sum of the cardinalities must equal the cardinality of the referential universe:

$$\sum_{k=1}^{K} x_k = |E| \tag{2.7}$$

The conclusion of the syllogistic argument is the statement Q_C $L_{C,1}$ *are* $L_{C,2}$ (as indicated in 2.2). For the case of quantifiers of the absolute type ("no", "some", "some... not", absolute, exception and comparative absolute), $Q_C = [a, b]$ the expressions that have to be optimized are the following ones:

$$a = minimize\ x_{m,n_1} + \ldots + x_{m,n_I} \tag{2.8}$$

$$b = maximize\ x_{m,n_1} + \ldots + x_{m,n_I} \tag{2.9}$$

with $x_{m,n_i} \in \{x_k, k = 1, ..., K\} \, \forall i = 1, ..., I$ subject to the following restrictions: the premises of the syllogisms (from *PR*1 to *PRN* as in explained in section 2.2.3) and the restriction stated in (2.5). This kind of syllogism can be solved using SIMPLEX directly.

For the case of proportional quantifiers in the conclusion, techniques from fractional programming must be used. These are the quantifiers "all", proportional, comparative proportional and similarity. The expressions to be optimized are:

$$a = minimize \, \frac{x_{m,n_1} + \ldots + x_{m,n_I}}{x_{m,d_1} + \ldots + x_{m,d_J}} \tag{2.10}$$

$$b = maximize \, \frac{x_{m,d_1} + \ldots + x_{m,n_I}}{x_{m,d_1} + \ldots + x_{m,d_J}} \tag{2.11}$$

with $x_{m,n_i}, x_{m,d_j} \in \{x_k, k = 1, ..., K\} \, \forall i = 1, ..., I, \forall i = j, ..., J$, and subject to the following restrictions: premises of the syllogism and the three additional constraints (2.5), (2.6) and (2.7).

Another contribution of our proposal to syllogistic reasoning is to deal with the combination of absolute type quantifiers with proportional type ones in the premises of the same argument. As we have stated, this idea was pointed to by N. Pfeifer [64], but only combining exception and proportional quantifiers in the same statement and with the same form for all the premises. We propose to go a step further; each premise can have its particular form and different quantifiers in different propositions.

All Aristotelian moods except AAA of Figure I are good examples of this form of syllogism. According to our formalization of quantifiers, "all" has a proportional interpretation as we assume that it refers to 100% meanwhile "no", "some" and "some…not" have an absolute reading; i.e., "zero elements", "at least one element" and "at least one element…not". For instance, AII of Figure III involves "all" in the Major Premise and "some" in the Minor Premise and in the conclusion.

On the other hand, at pointed out in Exception Syllogistics (see section 1.1.3), the reasoning with absolute quantifiers entails certain information on the cardinality of the subject-term of the premises. Given that our framework is flexible enough to introduce more than two premises, this information can be introduced in a syllogism through additional premises with the form *Q S are S*, where the subject-term and the predicate-term are the same and $Q = Q_{AB}$.

From the point of view of resolution procedure, certain problems with constraint (2.7) can appear as, in proportional terms, it is normalized to 1 or to 100% and an absolute quantifier higher than them could generate a problem in the inequations. We propose using a very high value (i.e., 10^8, 10^9, 10^{10},…) as the normalization value to avoid this problem.

The optimization procedure depends on the types of quantifiers that are being used. In the following sections we illustrate the behaviour of this proposal with different examples of syllogisms.

2.3 Some Examples of Syllogisms Using the Theory of Generalized Quantifiers

We stated that this framework was based on an interval interpretation of quantifiers, but it can effectively deal with three different interpretations of the them:

- Crisp interval quantifiers: a crisp quantifier Q defined as an interval $[a,b]$. The case of precise quantifiers also can be managed as the particular case when $a = b$. For instance, "fifteen students are tall" the absolute quantifier "fifteen" is defined as $Q_{AB} = [15,15]$.
- Fuzzy quantifiers approximated as pairs of intervals: a fuzzy quantifier Q defined as a pair of intervals $\{KER_Q, SUP_Q\}$, where $KER_Q = [b,c]$ represents the kernel and $SUP_Q = [a,d]$ the support of Q. For instance, "around fifteen students are tall", the absolute fuzzy quantifier "around fifteen" is defined as $Q_{AB} = \{[14,16],[13,17]\}$.
- Fuzzy quantifiers Q represented in the usual trapezoidal form with parameters $[a,b,c,d]$. For instance, "around fifteen students are tall", "around fifteen" is defined as $Q_{AB} = [13,14,16,17]$.

In the following sections, we describe the behaviour of our approach by using some examples of fuzzy syllogisms.

2.3.1 Syllogisms with Crisp Interval Quantifiers

This is the simplest approach and the basis for solving the other two. The particular optimization technique to be applied depends on the type of quantifier in the conclusion. Absolute, exception and absolute comparative cases can be calculated using the SIMPLEX optimization method, since the interval to be optimized depends on linear operations. In the case of proportional, comparative proportional and similarity quantifiers, linear fractional-programming techniques must be used [9].

$PR1$:	All animals but two are dogs
$PR2$:	All animals but two are cats
$PR3$:	All animals but two are parrots
C :	There are Q_C animals

Table 2.9: Syllogism with quantifiers of exception.

2.3.1.1 Example 1: Dogs, Cats and Parrots

As a first example, we consider the following problem[1]:

> *Dogs, cats and parrots*. How many animals do I have in my home, if all but two are dogs, all but two are cats and all but two are parrots?

As we can see, a combination of exception crisp quantifiers are involved. In order to cast the wording into the typical form of a syllogism, the first step is to identify the number of terms involved. In this case, for the sake of simplicity and coherence with Figure 2.1, let us consider $E = \textit{animals in my home}$ and three terms $dogs = P_1$, $cats = P_2$ and $parrots = P_3$.

Table 2.9 shows an initial formalization of the problem considering only the explicit information of the wording. The corresponding quantifiers in $PR1$, $PR2$ and $PR3$ are Binary of Exception (Q_{EB}) as defined in Table 2.8 (left) with $[a,b] = [2,2]$.

According to Table 2.8 (right), each one of the premises generates the following system of inequations:

$$\begin{aligned}
PR1 :& x_1 + x_2 + x_3 + x_4 = 2;\\
PR2 :& x_1 + x_2 + x_5 + x_6 = 2;\\
PR3 :& x_1 + x_3 + x_5 + x_7 = 2;\\
C :& x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = a;
\end{aligned} \tag{2.12}$$

where $Q_C = [a,b]$ is the quantifier in the conclusion. Applying the SIMPLEX method, we obtain, nevertheless, $Q_C = [0, inf)$; i.e., an undefined result. The origin of this result is that the set of premises does not contain all the necessary information that human beings manage to solve this problem. It is necessary to incorporate additional premises with the implicit or contextual information. In this case, we have added five premises ($PR4 - PR8$) thus producing the extended syllogism shown in Table 2.10.

[1] Problem taken from http://platea.pntic.mec.es/jescuder/mentales.htm

$PR1$:	All animals but two are dogs
$PR2$:	All animals but two are cats
$PR3$:	All animals but two are parrots
$PR4$:	No dog, cat or parrot is not an animal
$PR5$:	No animal is not a dog, a cat or a parrot
$PR6$:	No dog is a cat or a parrot
$PR7$:	No cat is a dog or a parrot
$PR8$:	No parrot is a dog or a cat
C :	There are Q_C animals

Table 2.10: Syllogism with quantifiers of exception, logic and additional premises.

The corresponding quantifiers of the additional premises $PR4-PR8$ are logical quantifiers $Q_4 = Q_5 = Q_6 = Q_7 = Q_8 = no$, labelled as Q_{LQ-no} in Table 2.8 (left). According to Table 2.8 (right), the corresponding system of inequations is:

$$\begin{aligned} PR4 :& x_1 = 0; \\ PR5 :& x_1 = 0; \\ PR6 :& x_6 + x_7 + x_8 = 0; \\ PR7 :& x_4 + x_7 + x_8 = 0; \\ PR8 :& x_4 + x_6 + x_8 = 0; \\ C :& x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = a; \end{aligned} \tag{2.13}$$

Applying the Simplex method, we obtain $Q_C = [3,3]$; that is, “There are *three* animals”; that is, I have a dog, a cat and a parrot in my home.

On the other hand, it is relevant to note that without the premises $PR4$ and $PR5$, we obtain $Q_C = [2,3]$; i.e., without these constraints we are assuming the possibility of the existence of animals in my home that are not dogs, cats or parrots. Nevertheless, in the wording of the problem it is implicit that the only animals that can be at home are dogs, cats and parrots and for that reason they are included.

The example shows a relevant aspect to manage in problems expressed in natural language: contextual and implicit information must be incorporated to the syllogism. Those syllogistic patterns limited to predefined patterns or with a small number of premises cannot manage this kind of problem.

$PR1$:	At least 70% of sixth course students passed Physics
$PR2$:	At least 75% of sixth course students passed Mathematics
$PR3$:	At least 90% of sixth course students passed Philosophy
$PR4$:	At least 85% of sixth course students passed Foreign Language
C :	Q_C sixth course students passed Physics, Mathematics, Philosophy and Foreign Language

Table 2.11: Syllogism with five terms and five premises.

2.3.1.2 Example 2: Sixth Course Students

The following example is mainly focused on the use of interval quantifiers. The exercise is extracted from the Spanish Mathematical Olympiad 1969-70[2]:

> In the tests of sixth course in a secondary school, at least 70% students passed the subject of Physics; at least 75% students passed Mathematics; at least 90% passed Philosophy and at least 85% passed Foreign Language. How many students, at least, did pass these subjects?

In this case, five terms must be considered in the universe E: *students in the sixth course* $= P_1$, *students that passed Physics* $= P_2$, *students that passed Mathematics* $= P_3$, *students that passed Philosophy* $= P_4$ and *students that passed Foreign Language* $= P_5$. Table 2.11 shows a possible formalization, where the corresponding quantifiers are: $Q_1 = [0.7, 1]$ of $PR1$; $Q_2 = [0.75, 1]$ of $PR2$; $Q_3 = [0.9, 1]$ of $PR3$; $Q_4 = [0.85, 1]$ of $PR4$; and $Q_C = [a, b]$ the quantifier of the conclusion.

In this case, we avoid the details regarding the corresponding set of inequations of each premise. According to section 2.2.2, $2^5 = 32$ disjoint sets must be generated and the corresponding inequations. The quantifier of the conclusion is the binary proportional type and it must be calculated by applying *linear fractional programming* obtaining $Q_C = [0.2, 1]$; i.e. "at least 20% of sixth course students passed Physics, Mathematics, Philosophy and Foreign Language", which corresponds to the expected result.

[2] *http://platea.pntic.mec.es/csanchez/olimp_1963-2004/OME2004.pdf*; p. 29

2.3.2 Syllogisms with Fuzzy Quantifiers Approximated as Pairs of Intervals

In this method, each fuzzy quantifier Q in the syllogism is defined as $Q = \{KER_Q, SUP_Q\}$, where $KER_Q = [b,c]$ corresponds with the kernel and $SUP_Q = [a,d]$ corresponds with the support. In this model, the solution is obtained from two systems of inequations, the first of them taking KER_Q as crisp interval definitions for all the statements in the syllogism and the second one taking SUP_Q. The corresponding solutions define KER_{Q_C} and SUP_{Q_C} for the quantifier Q_C, respectively. We should note that this is an exact and very simple approach for cases where the quantifiers are trapezoids. For the cases of non-trapezoidal quantifiers this produces an approximate solution. In the cases where non-normalized quantifiers may appear (either in the definition of the premises or due to the non-existence of solutions for the inequations system) this approach cannot produce results and therefore solutions should be obtained by using the approach described in section 2.3.3.

For the example in Table 2.11 we have the following quantifiers: $Q_1 = \{[0.8, 0.9], [0.7, 1]\}$; $Q_2 = \{[0.8, 0.85], [0.75, 0.9]\}$; $Q_3 = \{[0.92, 1], [0.9, 1]\}$; $Q_4 = \{[0.9, 0.95], [0.85, 1]\}$; and $Q_C = \{[b,c], [a,d]\}$ the quantifier of the conclusion.

Applying the procedure described in section 2.3.1 for each system, we obtain $KER_{Q_C} = [0.42, 1]$ for the system defined by taking the *Kernels* of all the quantifiers in the premises and $SUP_{Q_C} = [0.20, 1]$ for the system defined by taking the *Supports* of all the quantifiers in the premises. The result is therefore $Q_C = \{[0.42, 1], [0.2, 1]\}$ which is a result that entails the definition for Q_C = "At least 20% by mainly from 42%" obtained in the previous section. Figure 2.2 shows the graphical representation of the quantifier Q_C, where the interval $[0.2, 1]$ denotes the support of the fuzzy set and the interval $[0.42, 1]$ denotes its kernel.

As we can see, by using this approach we can only calculate the two represented intervals. Any other value between them must be interpolated. The main problem of this approach is for fuzzy quantifiers that are not represented as trapezoidal functions or those that are non-normalized. In these cases, we must use fuzzy quantifiers.

2.3.3 Syllogisms with Fuzzy Quantifiers

The use of fuzzy quantifiers supposes a generalization of the reasoning procedure described in the previous sections. Each fuzzy quantifier is managed through a number of $\alpha-$cuts, which are crisp intervals defined on the universe of discourse of the quantifier. Therefore, an

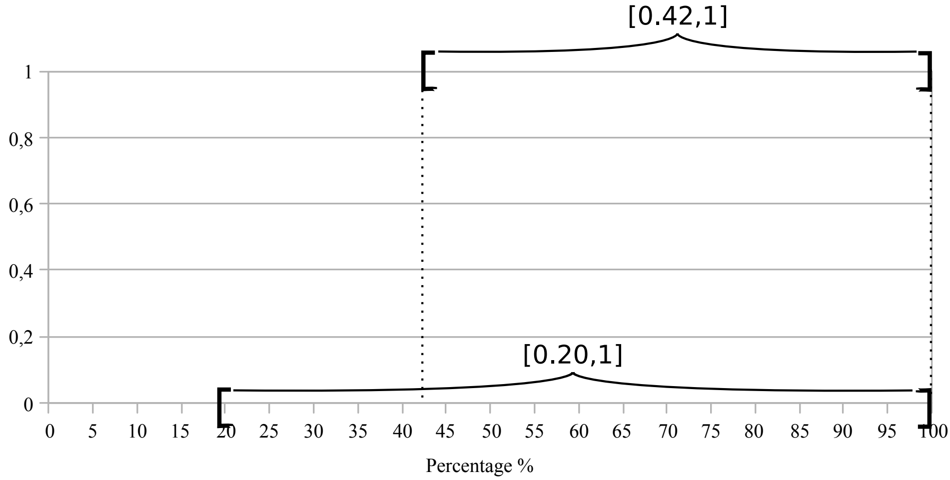

Figure 2.2: Graphical representation of the solution obtained for Q_C for the example in Table 2.11, $SUP_{Q_C} = [0.2, 1]$ and $KER_{Q_C} = [0.42, 1]$.

inequation system similar to the ones in section 2.2.3 is obtained for each $\alpha-$cut considered. The reasoning procedure consists of applying the method described in section 2.2.4 for each of these inequation systems. Each of the solutions obtained for all these systems define the corresponding $\alpha-$cut for the quantifier Q_C in the conclusion. By using this general approach, it is possible to manage any generalized quantifier (as the ones shown in Table 2.1) for three situations that are not considered in the previous models:

– trapezoidal quantifiers where non-normalized results are obtained in any of the solutions, since for this case KER_{Q_C} cannot be calculated and therefore the previous approach leads to indefinition.

– quantifiers defined with non-trapezoidal functions (in this case, this is an approximate solution where the level of approximation can be defined as we wish) and

– typical linguistic fuzzy quantifiers like “most”, “many”, “all but around three”, etc.

In order to better illustrate this approach we shall consider the five types of examples described in the subsequent sections.

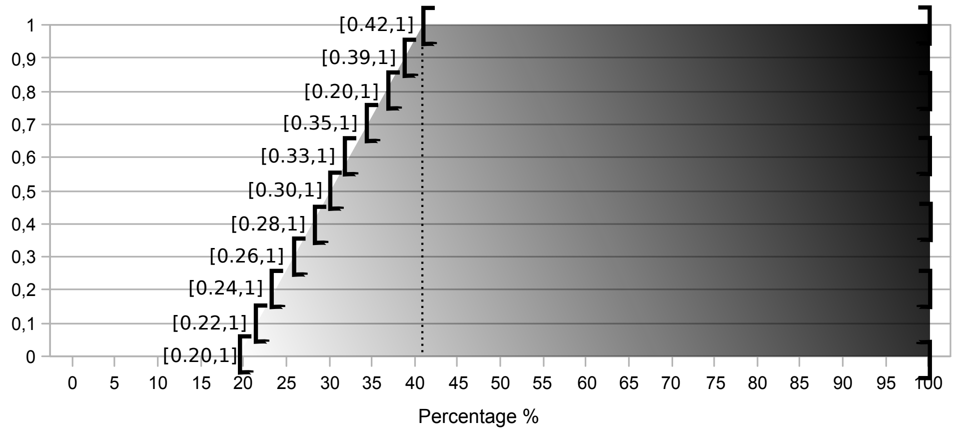

Figure 2.3: Graphical representation of Q_C with 11 $\alpha-$cuts.

2.3.3.1 Example 1 (Fuzzy Extension): Students of Sixth Course

This first example is the one in Table 2.11 which is useful for showing the consistency of this approach with the one described in the previous section. We shall consider fuzzy trapezoidal quantifiers that comprise the definitions for the quantifiers in Table 2.11. Using the usual notation for trapezoidal fuzzy membership functions where $Q = [a,b,c,d]$ ($KER_Q = [b,c], SUP_Q = [a,d]$) we have: $Q_1 = [0.7, 0.8, 0.9, 1]$; $Q_2 = [0.75, 0.8, 0.85, 0.9]$; $Q_3 = [0.9, 0.92, 1, 1]$; $Q_4 = [0.85, 0.9, 0.95, 1]$; and $Q_C = [a,b,c,d]$ the quantifier of the conclusion. Applying the corresponding optimization method, we obtain $Q_C = [0.2, 0.42, 1, 1]$, which is consistent with the results in the previous sections. Figure 2.3 shows a graphical representation of Q_C for the case considered involving eleven α-cuts. As we can see, the eleven α-cut intervals allow us to better approximate a fuzzy definition for the conclusion quantifier Q_C. It is worth noting that the extreme cases, $\alpha-$cut$= 0$ and $\alpha-$cut$= 1$, correspond to the *support* and *kernel* cases in the approximation described in the previous section 2.3.2 (see Figure 2.2).

2.3.3.2 Example 2 (Non-normalized Fuzzy Extension): Students of Sixth Course

We now illustrate a syllogism that produces a non-normalized fuzzy set as result. We consider again the example of Table 2.11 but adding the following premise:

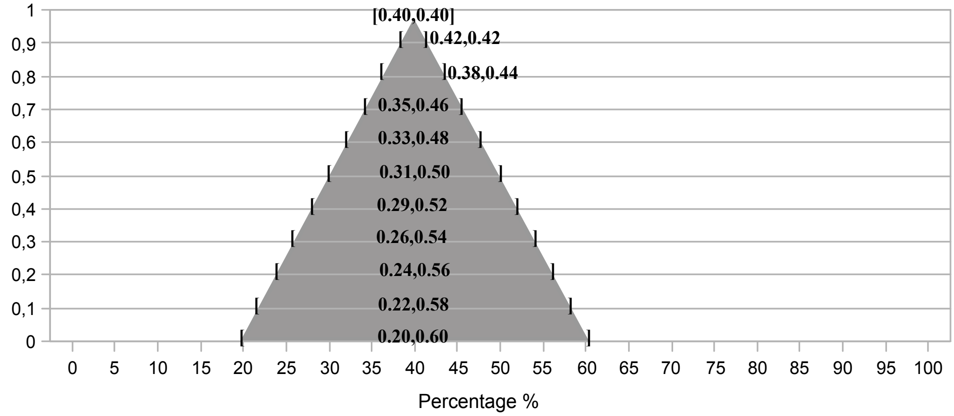

Figure 2.4: Graphical representation of the solution obtained for Q_C for the example in Table 2.11 with the additional premise $PR4$.

$PR5$: Between 40% and 90% sixth course students did not pass Physics or Mathematics or Philosophy or Foreign Language

Its definition is $Q_5 = [0.4, 0.6, 0.8, 0.9]$ and the corresponding system of inequations is added to the previous system. Results are shown in Figure 2.4. For $\alpha-$cuts higher than 0.95 the system has no solution; i.e., the fuzzy set that defines the quantifier of the conclusion is a non-normalized fuzzy set. Therefore, this syllogism cannot be solved by managing fuzzy quantifiers approximated as pairs of intervals (section 2.3.2).

2.3.3.3 Example 3 (Non-trapezoidal Result): Wine Warehouse

In this case, we show a syllogism using the class of proportional quantifiers known as Regular Increasing Monotone (RIM) quantifiers [89], that were proposed within the framework of quantifier guided aggregation of combination of criteria. This is a way of defining proportional quantifiers (such as "few", "many", "most", etc.) where the linguistic term is interpreted as a fuzzy subset Q of the $[0,1]$ interval. Its basic definition is shown in equation 2.14

$$Q^{\alpha}(p) = p^{\alpha}; \alpha > 0 \tag{2.14}$$

where Q^{α} denotes a linguistic quantifier and $p \in [0,1]$. For instance, for Q^2 (labelled as "most"), $Q^2(0.95) = 1$ means that saying 95% completely fulfills the meaning conveyed by

$PR1:$	Q^1 bottles of red wine are sold in the United Kingdom
$PR2:$	Q^1 bottles of red wine sold in the United Kingdom are bought by J. Moriarty
$C:$	Q_C bottles of red wine are sold in the United Kingdom and bought by J. Moriarty

Table 2.12: Zadeh's Intersection/product syllogism.

"most" and $Q^2(0.6) = 0.75$ means that saying 60% fulfils with degree 0.75 the meaning conveyed by "most". Other usual RIM quantifiers are $Q^{0.5}$, usually labelled [34] as "a few" and Q^1 (identity quantifier, also labelled as "a half"). RIM quantifiers are also associated with the semantics of "the greater the proportion of... the better" owing to its increasing monotonic behaviour [89]. Under this interpretation, the semantics should be labelled accordingly to the actual value of α as "linear" ($\alpha = 1$), "quadratical" ($\alpha = 2$), "sub-linear" ($\alpha < 1$), etc.

So, let us consider the following example using RIM quantifiers: in a wine warehouse that sells red wine, white wine and other products derived from grapes, we have the following sentences:

Q^1 bottles of red wine are sold in the United Kingdom.
Q^1 bottles of red wine sold in the United Kingdom are bought by J. Moriarty.

Both propositions involve the identity quantifier Q^1 ($\alpha = 1$), that can either be interpreted as "a half" or as "the greater... the better" (linear, $\alpha = 1$).

If we apply the Zadeh's [93] Intersection/product pattern, we can infer the following statement: "Q_C bottles of red wine are sold in United Kingdom and bought by J. Moriarty". So, the corresponding syllogism is shown in Table 2.12.

As the result of the corresponding optimization method, we obtain the quantifier Q_C shown in Figure 2.5, such that $Q_C = Q^{0.5}$. Its linguistic interpretations are "a few" [34] "a few bottles of red wine sold in the United Kingdom are bought by J. Moriarty" or "the greater the proportion of bottles of red wine that are sold in the United Kingdom and bought by J. Moriarty the better" (sublinear, $\alpha = 0.5$).

In case this example was solved using the $KER-SUP$ linear interpolation approach a much worse approximation (with bigger error) would be obtained for the inferred sublinear quantifier $Q_C = Q^{0.5}$ as it is shown in Figure 2.5.

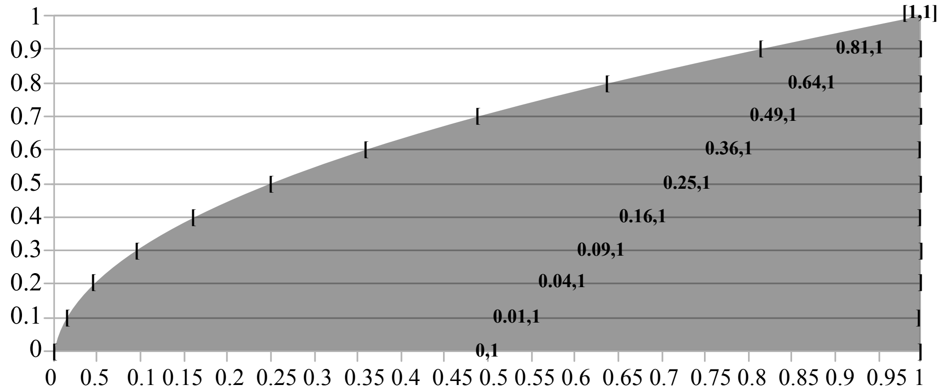

Figure 2.5: Q_c of Example 2.12.

$PR1$:	Many sales are from red wine
$PR2$:	Few sales are from white wine
C :	Q_C sales are not from red wine or white wine

Table 2.13: Zadeh's Consequent Disjunction syllogism with different conclusion.

2.3.3.4 Example 4 (proportional fuzzy quantifiers): Account of a Wine Warehouse

In this case, we show a syllogism involving fuzzy quantifiers of the type "many", "most", etc. Let us consider the account of the wine warehouse, that is not so good, and they only manage the following data:

Many sales are of red wine. A few sales are of white wine.

Taking both statements as premises in the universe of *products sold by wine warehouse* = E, we have a syllogism with three terms: *sales*=P_1, *red wine*=P_2 and *white wine*=P_3. With these two premises we can use, for example, consequent disjunction pattern of Zadeh [93], the conclusion of which takes the form Q_C P_1 *are* P_2 *or* P_3 with P_1 being the subject of the premises and P_2 and P_3 the corresponding predicates. However, in this example, we change the conclusion for the following one "Q_C sales are not of red wine or white wine", because this allows them to know how many sales are of other products. Table 2.13 shows the complete syllogism.

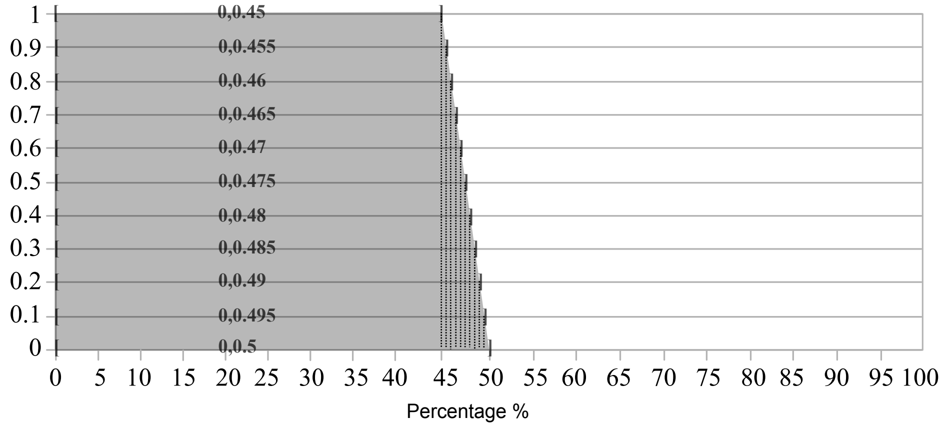

Figure 2.6: Graphical representation of the solution obtained for Q_C for the example in Table 2.13.

The first step to perform in the inference process is to assign the corresponding trapezoidal function to each one of the linguistic quantifiers of the premises. On the basis of [76], we assign the following trapezoids: *many* $= [0.5, 0.55, 0.65, 0.7]$ and $few = [0.1, 0.15, 0.2, 0.25]$. Next, we apply the reasoning procedure obtaining the following result: $Q_C = [0, 0, 0.45, 0.5]$; this can be associated with the following semantics: "at most half"; therefore, the conclusion can be stated as: "at most half sales are not of red wine or white wine". In Figure 2.6 we show the corresponding fuzzy set. It is relevant to note that the obtained Q_C of this example is a decreasing quantifier, a type of quantifier that some classical models of fuzzy syllogism cannot manage [93, 57].

The example shown in Table 2.13 shares the same aspect of P. Peterson's Intermediate Syllogisms [63], but with a different conclusion, as Peterson's schemas only include simple conclusions, without involving logical operations such as disjunction. Table 2.14 shows the syllogism of Table 2.13, but with different terms for a more clear example (*people*, *white hats* and *red ties*) and according to BKO (B meant that $PR1$ involves the quantifier "few"; K that $PR2$ involves the quantifier "many", and O that C involves the quantifier "some... not") Peterson's schema of Figure III. Using the same previous definitions for the "few" and "many" quantifiers, we obtain the following result: $Q_C = [0, 0, 0.98, 0.98]$, which is consistent with the usual definition for the quantifier *some... not*$= [0, 0, 1-\varepsilon, 1-\varepsilon]$.

$PR1$:	Few people are wearing white hats
$PR2$:	Many people are wearing red ties
C :	Q_C people that are wearing red ties do not wear white hats

Table 2.14: Peterson's BKO syllogism.

$PR1$:	All but around fifteen boxes of wine are for J. Moriarty
$PR2$:	Around four boxes of the boxes that are not for J. Moriarty are for J. Watson
C :	Q_C boxes are not for J. Moriarty neither for J. Watson

Table 2.15: Zadeh's Intersection/product syllogism.

2.3.3.5 Example 5 (Exception Fuzzy Quantifiers): Wine Warehouse

The final example shows a syllogism combining fuzzy quantifiers that have not been previously considered in the literature as exception quantifiers (i.e., "all but around four", etc.) combined with absolute ones (i.e., "around four", etc.):

> All but around fifteen boxes of wine are for J. Moriarty. Around four boxes of the boxes that are not for J. Moriarty are for J. Watson.

We can identify the following three terms in the statements: *boxes of wine* $= P_1$, *boxes for J. Moriarty* $= P_2$ and *boxes for J. Watson* $= P_3$. Given this information, we can infer how many boxes are not for J. Moriarty neither for J. Watson; that is, "Q_C boxes are not for J. Moriarty neither for J. Watson". In Table 2.15 is afforded the complete argument.

The quantifier of $PR1$ is a fuzzy quantifier of exception, as it indicates that there are *around fifteen* boxes of wine that are not for J. Moriarty. For instance, the associated trapezoidal function to this number can be *around 15*$= [13,14,16,17]$. The quantifier of $PR2$ is another fuzzy quantifier, but in this case an absolute one. This kind of quantifier has been analysed in the literature from the very beginning (Zadeh's distinction between absolute/proportional fuzzy quantifiers [90]), but it has not been considered from the point of view of its use in reasoning. In this case, we assign *around four*$= [3,4,4,5]$.

After applying the reasoning process, we obtain the absolute quantifier $Q_C = [8,10,12,14]$ for the conclusion, with the usual associated meaning of "around eleven" (since value 11 is in the middle of the trapezoid's kernel). Therefore, in this example the conclusion states that "around eleven boxes are not for J. Moriarty neither for J. Watson".

On the other hand, it is worth noting that this model is consistent with a fuzzy arithmetic approach to this problem; since from $PR1$ we have that "around fifteen' boxes are not for J. Moriarty, and from $PR2$ we have that "around four" of these are not for J. Watson; therefore, "around fifteen" $\ominus$ "around four" = "around eleven", which is the number of boxes that are not for J. Moriarty neither for J. Watson.

2.4 Combination of Absolute and Proportional Quantifiers

In the analyses of the most relevant syllogistic patterns of the literature presented in chapter 1 we concluded that only in Exception Syllogistics was the possibility of combination proportional and absolute quantifiers (exception ones) addressed. However, this proposal only considered the combination of quantifiers in the same statement for determining the value of the subject-term and, therefore, the inference mechanism did not have any relevant innovation.

However, in natural language it is not difficult to think of in arguments involving statements with different type of quantifiers. For instance, "there are around forty students in this classroom, most of them are blond; therefore, no more than five or six are not blond". Although this is an example of fuzzy syllogism and the quantifier of the conclusion ("no more than five or six") may be a matter of debate, nobody would say that this is a fully invalid inference. Therefore, the combination of absolute and proportional quantifiers in a same syllogism can be easily exemplified in natural language and, therefore, it constitutes a challenge for any new proposal of syllogistic reasoning system.

The difference between an absolute and proportional reading of a linguistic quantifier is a topic that has ben deeply analysed in linguistics. In [44, 21] and interesting analysis of this question is described. Both works analyse the case of the quantifier "many" and match in their definitions for the absolute and proportional readings. So, let us consider a proposition *many S are P*:

- Absolute (or cardinal) reading: $|S \cap P| > n$; n a real number.
- Proportional reading: $\frac{|S \cap P|}{|S|} \geq k$; k a fraction or percentage.

Given that "many" is a vague quantifier, the values of n and k are context-dependent. Other quantifiers, such as "few", "many", etc. also have both possibilities of reading and in their definitions $>$ and $\geq$ can change for $<$ and $\leq$.

$PR1:$	All kings are powerful
$PR2:$	Some kings are single
$C:$	Q_C powerful people are single

Table 2.16: AII mood of Figure III.

Exception Syllogistics (section 1.1.3), the only syllogistic system that deals with arguments with absolute quantifiers, pointed out that a value for S must be assigned to a right inference and if it is unknown, it should be estimated. Therefore, this idea must be incorporated to our model. It is flexible enough and there is no *a priori* problem in building inequations systems that involve simultaneously fractional and linear equations as they can be solved using *fractional programming* techniques. From a computational point of view, the key point is to assign a large enough size to the universe of discourse. Thus, any more or less large value for n but also expressed in terms of percentage to satisfy k can be used in the syllogism.

To illustrate this idea and, for the sake of simplicity, we shall consider the AII mood (Figure III) to show how the combination between the proportional and absolute quantifiers work; since A involves the proportional quantifier "all" and I the absolute one "some". This is shown in Table 2.16.

The conclusion of this argument is also an I statement. Hence, let us consider $Powerful = P_1$, $King = P_2$ and $Single = P_3$ (see Figure 2.1). To preserve the Aristotelian meaning, we shall use crisp definition for these quantifiers according to Table 2.8 (left); i.e., $all = 100\%$ and $some \geq 1$. According to Table 2.8 (right), the corresponding system of inequations is:

$$\begin{aligned} PR1 :& \frac{x_7 + x_8}{x_5 + x_6 + x_7 + x_8} = 10^{100}; \\ PR2 :& x_4 + x_8 \geq 1; \\ C :& x_6 + x_8 \geq a; \end{aligned} \tag{2.15}$$

As can be observed, the right-hand member of the equation PR1 is a large enough value but preserving the proportional meaning of "all", as it still denotes the inclusion of the subject-term in the predicate-term, and it avoids conflicts with PR2. Applying the fractional programming techniques, we obtain $Q_C \geq 1$, the right result; i.e., "some powerful people are single".

2.5 Summary

In this chapter we have described our proposal for a syllogistics that focused on improving the weaknesses of the models analysed in the previous chapter; i.e.: i) non-fuzzy approaches that involve vague quantifiers have a logic foundation rather a heuristic one, generating a relevant amount of new patterns but are too rigid; ii) fuzzy approaches propose heuristic procedures for calculating the conclusion of the argument in terms of the quantifiers of the conclusion, but they are not fully compatible with Aristotelian Syllogistics; iii) many of the new proposed patterns are practically limited to the form of three terms, two premises and a conclusion. The remaining ones only consider some additional premises but the same number of patterns; and, iv) only schemas with absolute (such as exception ones) or proportional quantifiers (such as proportional ones) are considered. There is no model that proposes how to combine them although this is a common situation in natural language.

Thus, our approach consists of formulating a new framework of syllogistic reasoning, firstly, compatible with TGQ, the standard theory of quantification in natural language. Hence, preserving the standard form of quantified statements (*Q S are P*), we introduce new quantifiers into syllogism, such as comparative or exception ones, under the same approach. Furthermore, this framework also allows us to use crisp, interval and fuzzy definitions of quantifiers without changing the procedure.

On the other hand, we assume the heuristic strategy, emphasized by Probabilistic Syllogistics, and adopt an algebraic theory of reasoning, as in Interval Syllogistics, but supported on Venn diagrams, as Interval and Exception Syllogistics. Thus, we transform a reasoning problem into a mathematical optimization problem, where the quantifier of the conclusion is calculated in terms of the quantifiers of the premises. Hence, from each premise n inequations are generated, depending on the definition of the quantifier (crisp, interval or fuzzy), which are the constraints for optimizing the conclusion, which is also expressed as inequations. The techniques used to deal with with operations are Simplex and fractional programming.

We also consider a fundamental challenge to our proposal to be compatible with a linguistic version of syllogism, where the quantifiers can be ordered into an LSO or another structure, such as in Interval Syllogistics, where each quantifier have associated a crisp value, an interval or a fuzzy set and the result of the inference procedure is approximated to the closer definition.

This model is partially implemented in a software library called SEREA (Syllogistic Epistemic REAsoner) that is capable of managing the reasoning scheme described in this paper. The current version of SEREA is freely downloadable from our website [2].

Having analysed the case of vagueness in quantifiers, the case of vagueness in terms also deserves our attention, as this is also a very common phenomenon in natural language. So, how can a syllogism that involves vague terms as middle term be qualified? Can a valid or acceptable syllogism be obtained or only a fallacy? In the next chapter, we shall explore this idea. This leads us to a new form of syllogism where the problem is not the calculation of the quantifier of the conclusion in terms of the quantifiers of the premises, but the reliability of or confidence in the conclusion given the approximate character of the chaining.

CHAPTER 3

APPROXIMATE CHAINING SYLLOGISM

In the two previous chapters, we described different approaches to syllogism where vagueness only appears in the quantifiers. However, as stated at the end of chapter 2, vagueness can also appear in the terms because they are natural language words. As have previously stated, syllogism is a form of chaining reasoning and, therefore, vagueness in the middle term is the most interesting case for producing a new form of reasoning pattern not considered until now and for reconsidering some arguments usually qualified as fallacies; the so-called *fallacy of four terms* and the *fallacy of ambiguous middle term* [12]. The fallacy of four terms appears when the syllogism involves two different terms in the middle term (DT); for instance, "all *students* are tall, all dogs are *pets*; therefore, all dogs are tall" (AAA-Figure I) is fallacious because *students* and *pets* are different. The fallacy of the ambiguous middle term appears when different meanings of the middle term are used in the different premises; for instance, "all *revolutionary* people have brilliant ideas, all anarchist are *revolutionary*; therefore, all anarchist have brilliant ideas" (AAA-Figure I) is fallacious as in the first premise *revolutionary* refers to those people that defend the abolition of government, while in the second premise it refers to people whose have innovative ideas.

Nevertheless, in certain cases, these arguments are not fallacies but approximate inferences. Let us consider the example shown in Table 3.1. The DT is played by two different terms, *disaster* and *calamity*, and, therefore, it falls into the fallacy of four terms. Nevertheless, from a pragmatic or everyday reasoning point of view, it is an approximate but acceptable inference as the speakers know that the terms *disaster* and *calamity* are similar enough to preserve a chaining that is sufficiently strong. In this case, any weakness point in the chaining

PR1:	All earthquakes are a *disaster*
PR2:	No *calamity* is desirable
C:	No earthquake is desirable

Table 3.1: Syllogism with vagueness in the DT.

affects the reliability of the syllogism, which should be qualified in terms of acceptability rather than validity.

In this chapter we shall focus on this type of syllogism. Our motivation is because, although the chaining can be constituted by two different terms, they can have a relationship of similarity or a family resemblance between them that is sufficiently strong to render the argument acceptable for the speakers. To qualify the reliability of the argument, we propose to evaluate the conclusion in terms of truth value or degree of confidence, which depends on the strength of the chaining between the terms. We assume the following statement as an axiom: *the greater the similarity between the terms, the higher the degree of confidence in the conclusion*.

The best way, from our point of view, to measure this eventual vague link is in terms of similarity, as this is a well defined concept both in maths and fuzzy logic. We identify three possible situations: i) similarity in terms of fuzzy inclusion, ii) similarity in terms of fuzzy overlapping, and, iii) similarity in terms of synonymy. The first two have been studied in depth in fuzzy logic and we only make a brief description of them; case iii), although a very typical and intuitive semantic relationship in natural language, is less studied in this field.

This chapter is organized as follows: in section 3.1 different proposals for dealing with similarity between terms are analysed; in section 3.2 we elaborate a measure of synonymy between two terms; in section 3.3, we propose a basic inference pattern for this type of syllogisms including some examples; and, finally, in section 3.4, we summarize the main contributions of this chapter.

3.1 Similarity between Terms: Fuzzy Inclusion, Fuzzy Overlapping and Synonymy

We address the syllogism with fuzzy terms (which we shall call Approximate Chaining Syllogism (ACS)) using the set-based interpretation of quantified statements, since it is the one used in our proposal of syllogistic system (see chapter 2). As we said above, we identify

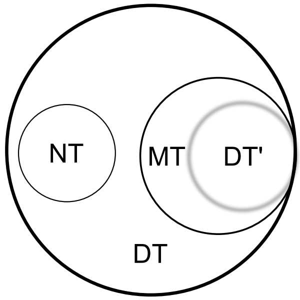

Figure 3.1: Counter example of ACS when DT' is included in DT.

three possible scenarios with three different underlying ideas of similarity: i) fuzzy inclusion, ii) fuzzy overlapping and, iii) synonymy. In the subsequent sections, each one them will be analysed.

3.1.1 Similarity as Fuzzy Inclusion

Fuzzy inclusion is a type of operation well known in fuzzy logic. It denotes the degree to which a fuzzy set is a subset of another fuzzy set such as, for instance, the inclusion of the fuzzy set *slippery roads* in the fuzzy set of *dangerous roads*. There are different definitions; for instance, let A and B be two (fuzzy) sets in a universe U with $B \subseteq A$;

$$\mu_B(u_i) \leq \mu_A(u_i) \; ; \forall u_i \in U \tag{3.1}$$

or

$$\mu_B(u_i) \leq \mu_A(u_i) \; ; B \subseteq A \equiv B \rightarrow A \tag{3.2}$$

For the ACS, fuzzy inclusion is valid in a particular case; i.e., when the middle term of the Minor Premise is a (fuzzy) subset of the middle term of the Major Premise (see Figure 3.2). In the opposite direction, when the middle term of the Major Premise is a (fuzzy) subset of the middle term of the Minor Premise, is not valid as a simple and clear counter example can be generated, as Figure 3.1 shows, where DT denotes the middle term of the Minor Premise and DT' the middle term of the Major one, NT the minor term and MT the major term.

For the right direction of the inclusion, a counter example can also be generated, but, in this case, it is necessary to assume as a precondition a substantial difference in the size of the sets or a degree of fuzzy inclusion very low, as can be deduced observing Figures 3.1 and 3.2.

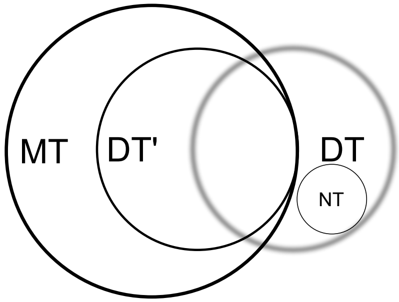

Figure 3.2: Counter example of ACS when DT is included in DT'.

$PR1$	Some people are *intelligent*
$PR2$	No *wise person* ignores $E = mc^2$
C	Some people do not ignore $E = mc^2$

Table 3.2: Syllogism with approximate chaining.

$PR1$	Some people are *intelligent*
$PR2$	No *wise person* ignores $E = mc^2$
$PR3$	Almost all wise people are intelligent
C	Some people do not ignore $E = mc^2$

Table 3.3: Syllogism where the additional premise for *Spanish* and *European*.

Table 3.2 shows an example of syllogistic argument involving the terms *intelligent* and *wise* in the chaining. The users assume that *wise people* $\subseteq$ *intelligent* and, thus, they can be linked trough a fuzzy inclusion.

Other possibility to make explicit the inclusion of DT in DT' is to introduce an additional premise in the syllogism with the form "almost all wise people are intelligent" and generate a poly-syllogism [12], as it is shown in Table 3.3.

3.1.2 Similarity as Overlapping

Another possible relationship of similarity between two terms A and B is the so-called fuzzy overlapping. This idea of chaining was initially proposed in Mamdani's reasoning style [10] and subsequently extended to other type of fuzzy reasoning [11]. Although Mamdani's reasoning style is different to syllogism, the underlying idea can be applied to it. For a better

Set of premises	If *X is A*, then *Y is B* If *Y is B'*, then *Z is C*
Observation	*X is A'*
Conclusion	*Z is C'*

Table 3.4: Mamdani's Approximated Chaining schema.

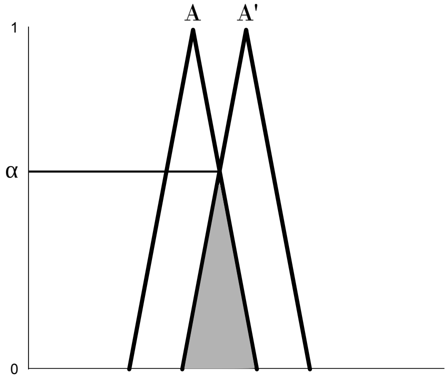

Figure 3.3: α indicates the overlapping between A and A'.

explanation of this idea, we shall explain the concept of fuzzy overlapping using a Mamdani's argument and other applying fuzzy overlapping to the syllogism.

Arguments in Mamdani's reasoning involves a *set of rules*, an *observation* and a *conclusion*. The rules take the form "If *X is A* then *Y is B*". The chaining is through the antecedent and the consequent parts: there is an intermediate variable in the consequent part of a rule (or some rules) and in the antecedent part of other(s) [10]. When a relaxation in the chaining is assumed, the pattern shown in Table 3.4 is produced. Thus, the result of the consequent of the first premise (*Y is B*) must be reassessed when is the antecedent of the second premise (*Y is B'*). One way to do so is to measure the overlapping between B and B' (see Figure 3.3) obtaining the degree of fulfilment (α) associated to it.

This idea can be easily adapted to ACS, whether the middle term comprises two terms where one of them involves a linguistic modifier of the other, or whether they are different labels in the same linguistic partition. Let us consider the example shown in Table 3.5, where *wet* (DT) and *extremely wet* (DT') are the middle terms. If a linguistic partition for evaluating humidity is considered, *wet* and *extremely wet* are two labels with certain degree of overlap-

Some spring days are *wet*
All *extremely wet* days are cloudy
Some spring days are cloudy

Table 3.5: Syllogism with overlapping between DT and DT'.

All earthquakes are a *disaster*
No *calamity* is desirable
No earthquake is desirable

Table 3.6: Syllogism with DT and DT' synonyms.

ping, which affect to the degree of confidence of the conclusion: the higher the degree of overlapping, the higher the confidence on the conclusion.

3.1.3 Similarity as Synonymy

The last type of relationship of similarity that we propose is synonymy, a relationship that denotes resemblance among the meanings of words. It is a very common and intuitive for speakers in natural language but the modelling thereof is not obvious in terms of fuzzy sets. Let us consider again the example of the introduction (Table 3.6), where the middle terms are *disaster* (DT) and *calamity* (DT').

We cannot properly say that *disaster* and *calamity* have a inclusion relationship as neither is *disaster* a subset of *calamity* nor *calamity* a subset of *disaster*. Nor are they labels on a linguistic partition on the universe of *bad events*, where neither of them can be built through a linguistic modifier or a shifting over the other; for instance, *calamity* are not, for instance, *serious disasters* or *more negative* than disasters. Hence, neither of the two previous cases can be applied. However, as we have already stated, this is an acceptable syllogism as the conclusion follows reasonably from the premises given that the chaining is established through the strong link between *disaster* and *calamity*.

In the next section, we shall analyse in depth the concept of synonymy and will propose a measure to calculate a degree of synonymy between two terms.

(a) Syllogism AAA - Figure I.

PR1:	All human beings are mortal
PR2:	All Greeks are human beings
C:	All Greeks are mortal

(b) Syllogism AAA - Figure I with DT interchanged.

PR1:	All featherless bipeds are mortal
PR2:	All Greeks are featherless bipeds
C:	All Greeks are mortal

Table 3.7: Examples of syllogisms with the terms interchanged.

3.2 A Proposal for a Measurement of Synonymy between Two Terms

We address synonymy between terms attending to the aim of this research: evaluating the validity of a syllogism where the middle term is played by two synonymous terms. This substitution can be performed in two senses: i) changing all the occurrences of the DT by a synonym; ii) changing only some occurrences of the DT for the synonymous term.

For case i), let us consider the argument of Table 3.7(a). The role of DT is played by the term *human being*; if we change this term for *featherless bipeds*, Plato's definition for *human beings*, we obtain a new syllogism (see Table 3.7(b)) equivalent to the previous one because both terms are completely equivalent.

This case is not the most relevant one as if we substitute all the occurrences of the DT by a synonym, we obtain an equivalent syllogism preserving the precise chaining.

Case ii) entails, in fact, a modification in the chaining and we shall focus on it. So, let us consider the example shown in Table 3.6. The role of DT is played by the terms *disaster* and *calamity*. To preserve the validity of the argument, the terms must be or totally equivalent or, at least, they must have a very similar meaning. In this context, they satisfy the condition or having a certain degree of similarity and, hence, the argument is acceptable.

Nevertheless, the definition of synonymy and what type of relationship it denotes is still a question with multiple dimensions. In the literature, we can find two main approaches to synonymy [74]:

1. Synonymy as a theoretical challenge: It is a theoretical problem and, therefore, its main aim is to obtain a clear definition of it. Thus, synonymy is understood as a precise relationship and its criterion is identical meaning.

2. Synonymy as a phenomenon of natural language: It is studied from the use of the speakers; therefore, it is a experimental challenge. Thus, synonymy is understood a

relationship of similarity, not identity, whose criterion is similar meaning according to the context of utterance.

Defenders of the first perspective can be found in the field of philosophy of language. For instance, W. O. Quine advocates this point of view. Thus, if we consider modelling synonymy as a crisp relation, we must identify meanings with equivalence classes of synonymous expressions [66]; in [67, p. 32], he said:

> A truth that can be turned into a logical truth by substituting synonyms for synonyms. Statements may be said simply to be cognitively synonymous when their biconditional (the result of joining them by "if and only if") is analytic.

If we assume this definition of synonymy, any syllogism involving two synonyms in the middle term is a full valid syllogism as the terms are equivalent from this perspective.

The second perspective also has numerous defenders, starting from Gabriel Girard [24]. He made an approach to this idea stating that:

> In order to obtain property, one does not have to be demanding with words; one does not have to imagine all that so-called synonyms are so with all the rigorousness of perfect resemblance; since this only consists of a principal idea which they all enunciate, rather each one is made different in its own way by an accessorial idea which gives it its own singular character. The similarity bought about by the principal idea thus makes the words synonyms; the difference that stems from the particular idea, which accompanies the general one, means that they are not perfectly so, and that they can be distinguished in the same manner as the different shades of the same colour.

Currently, although both perspectives are valid, the most widely accepted is the second one; i.e., a relationship of similarity that must be studied from its empirical use. We adopt the framework proposed in [74, 75], where a pragmatical study of synonymy using a *thesaurus*[1] is proposed as it gathers the use of synonyms of the speakers of a language. Furthermore, two additional reflections about synonymy are assumed:

1. It has a gradual character. Thus, although two expressions can be synonym to a certain degree, this degree may not be high enough for making them interchangeable. So, for

[1] *Thesaurus* is a book that lists words in groups of synonyms and related concepts.

instance, *car* and *automobile* are interchangeable in an informal conversation between two people about buying a new car; they are not interchangeable in a report about traffic accidents because *automobile* not only refers to car accidents but also motorbike or lorry accidents.

2. The polysemy of the terms. Sometimes, two synonymous terms cannot be interchangeable as one of them is polysemic and the meaning in a particular statement does not correspond with the synonym. For instance, the polysemic word *table* is synonymous of *desk* in one of its meanings; however, it cannot be substituted by this one in the statement "the data in the *table* of your last paper is wrong".

As we have stated previously, an experimental criterion of synonymy is based on what words the speakers of a language actually consider as synonyms. Thesaurus dictionaries are the documents that gather this information. Thus, as criterion of synonymy, we adopt the following: two words are synonyms if both words appear as synonyms in a thesaurus dictionary. Nevertheless, a couple of additional conditions must be taken into account:

1. Natural languages are systems in constant evolution and, accordingly, synonymy, that is characterized by its graduability and its context-dependence, also changes constantly. Hence, the thesaurus must be improved according to the changes in the corresponding language.

2. It is reasonable to think that synonymy is a symmetric property (if a is synonymous with b then b is synonymous with a), but this is not always the case in a thesaurus dictionary. The usual origin of this situation is that the relationship between the terms is not exactly synonymy but relationships of different nature such as hyponymy, hiperonymy, meronymy, etc., which are not symmetric.

3. With respect to transitive property, if a is synonymous with b and b is synonymous with c, a is synonymous with c, this may not be the case. For instance, we can say that *car* is synonymous of *automobile* and *automobile* is a synonym of *motorbike* but *car* is not synonymous with *motorbike*. From the point of view of reasoning, this question must be taken into account for the case of polyllogisms (syllogisms with more than two premises), as the chaining can be broken by an inadequate transitivity between the terms.

Thus, we shall adopt an experimental definition of synonymy based on a thesaurus. This leads us to interpreting the synonymy between two terms as a question of degree. Therefore, the next step is to define a procedure to measure or calculate the similarity, resemblance or synonymy between two terms. This question will be addressed in the next section.

3.2.1 Calculation of the Degree of Synonymy

Following [74, 75], we adopt a framework where the degree of synonymy between two terms is expressed as x, where $x \in [0,1]$ according to the information retrieved from a thesaurus dictionary. For instance, for two given terms a and b, it is possible to calculate their degree of synonymy x. The elaboration of this measure is divided into two steps: i) selection of a thesaurus dictionary; ii) model for calculating degrees of synonym.

We adopt an on-line thesaurus dictionary as they usually are in constant evolution and incorporate the changes in the uses of the speakers in their language better (one of our prerequisites) and consulting them is also very easy in general. On the other hand, they usually involve both conceptual-semantic and lexical relations, generating a wide set of relationships among terms, which satisfies the other two prerequisites that we have established: to distinguish among the subtypes of synonymy (hyponymy, hiperonymy, etc.) that are not symmetric and to consider the transitivity problem. We shall use WordNet [35] to illustrate our examples.

WordNet is a “large lexical database of English. It includes nouns, verbs, adjectives and adverbs that are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept” [1]. Thus, each concept provides two types of information: *synset* and *gloss*. Synset is the class terms that have synonymy relationship (total or partial) of a particular concept; each one of the synonyms that compound a synset is called a *variant*. Gloss is a distinctive definition for characterizing a synset. Figure 3.4 shows a screen capture using the term *write*.

It is a polysemic word and, hence, it has various glosses that appear in brackets:

- produce a literary work
- communicate or express by writing
- write music
- mark or trace on a surface
- record data on a computer

Figure 3.4: The term *write* on WordNet.

- write or name the letters that comprise the conventionally accepted form of (a word or part of a word)
- create code, write a computer program

The glosses are not defined using an univocal way but they can adopt different forms. In [26, p. 64-65] nine different patterns for defining glosses are identified.

Regarding the synsets associated to each concept, different categories are distinguished. The most relevant are hyponymy, meronymy and hyperonymy:

- Hyponymy: The specific term used to designate a member of a class; X is a hyponym of Y if X is a (kind of) Y [1]. In the case of verbs, it is known as troponymy; i.e., a verb that indicates more precisely the manner of doing something by replacing a verb of a more generalized meaning. For instance, *draw* defined as write a legal document or paper is a troponym of *write* under its gloss “produce a literary work”.
- Meronymy: The name of a constituent part of, the substance of, or a member of something. X is a meronym of Y if X is a part of Y [1]. It is typical in nouns. For instance, the part-whole relation holds between Synsets like *chair* and $\{back, backrest\}$, $\{seat\}$ and $\{leg\}$. Parts are inherited from their superordinates: if a chair has legs, then an armchair has legs as well.

– Hyperonymy: The generic term used to designate a whole class of specific instances. Y is a hyperonym of X if X is a (kind of) Y [1]. For instance, *create verbally* is direct hyperonym of *write* of its first gloss. It is the reverse of hyponymy.

Nevertheless, it is worth noting that synonyms are grouped into unordered sets; i.e., the synonyms are not ordered attending to their degree of synonym or similarity respect to the main concept. Furthermore, nor are they classified according to their linguistic register and hence, vulgar and cultured terms are not distinguished. Finally, the synonymy is not only between terms but also between terms and phrasal verbs; for instance, *close up* and *keep mum*. This case is dealt with in a very irregular way in WordNet [26].

Galnet [26] is a version of WordNet adapted to the Galician language. It is integrated into the EuroWordNet project, that is a version of WordNet improved with an interlingual index (ILI) to European languages. It is a tool that allow us to blind the synsets of the WordNet of a particular language with the synsets of other languages. It is based on the use of ontologies and hierarchies of domains, avoiding, by this way, false friends that are common in the translation of languages. This neither includes a organization of synsets based on the degree of similarity but it assumes as objective a best analysis of phrasal verbs. It is under development but the preliminaries results, for body parts and substances, are available[2].

Once the thesaurus has been chosen, the next step is to establish the way to measure the synonymy between two terms because, as we said before, this value is not provided by the dictionary. We assume the thesis that the degree of synonymy between two given terms can be calculated using similarity measures [78]. So, let a and b be two terms for calculating its degree of synonym and $A = Synset(a)$ and $B = Synset(b)$. In [20], we can find several definitions of similarity measures that can be used for our aim:

– Jaccard coefficient: The percentage of synonyms shared by a and b and their union.

$$\textsc{Syn}_{JC}(a,b) = \frac{|A \cap B|}{|A \cup B|} \tag{3.3}$$

– Dice coefficient: The proportion between the number of synonyms shared by a and b and the arithmetic mean of $|A|$ and $|B|$.

$$\textsc{Syn}_{DC}(a,b) = \frac{|A \cap B|}{\frac{|A|+|B|}{2}} = \frac{2 \cdot |A \cap B|}{|A| + |B|} \tag{3.4}$$

[2]adimen.si.ehu.es/cgi-bin/wei/public/wei.consult.perl

– Cosine coefficient: This is the same as Dice coefficient and with the arithmetic mean substituted by a geometric mean.

$$\text{SYN}_{CC}(a,b) = \frac{|A \cap B|}{\sqrt{|A| \cdot |B|}} \tag{3.5}$$

– Mutual similarity coefficient: This is also the same as Dice coefficient and with the harmonic mean of number of elements of A and B.

$$\text{SYN}_{MSC}(a,b) = \frac{|A \cap B|}{\frac{2}{\frac{1}{|A|}+\frac{1}{|B|}}} = \frac{\frac{|A \cap B|}{|A|} + \frac{|A \cap B|}{|B|}}{2} \tag{3.6}$$

– Overlap coefficient: In this case, the denominator is the *min* of the cardinality of $|A|$ and $|B|$.

$$\text{SYN}_{OC}(a,b) = \frac{|A \cap B|}{min(|A|,|B|)} \tag{3.7}$$

All these measures are symmetric (the degree of synonymy between a and b is the same that between b and a). They can be ordered attending to their results[3]:

$$\text{SYN}_{JC} \leq \text{SYN}_{DC} \leq \text{SYN}_{CC} \leq \text{SYN}_{MSC} \leq \text{SYN}_{OC} \tag{3.8}$$

We have pointed out that synonymy is, in general, a symmetric relationship although it may not be symmetric in some particular cases; i.e., if one of the synonymous terms play the role of prototype and the other not. For instance, the synonymy between *window* and *dormer* is not symmetric if we consider *window* as the prototype (the term that involves any kind of window) and *dormer* as a particular type because all *dormers* are windows but not all *windows* are dormers.

A non-symmetric model of distance measure is the so-called *ratio model* [84]. This measure calculates the quotient between the cardinality of the set of synonymous that are shared by A and B divided by the addition of the cardinalities of the three sets of synonyms: i) the set of synonyms shared by a and b ($A \cap B$); ii) the set of synonyms of a that are not synonymous of b ($A - B$) with a weight $\alpha \geq 0$; iii) the set of synonyms of b that are not synonymous of a ($B - A$) with a weight $\beta \geq 0$. The result is the following measure:

$$\text{SYN}_{RM}(a,b) = \frac{|A \cap B|}{|A \cap B| + \alpha|A - B| + \beta|B - A|} \tag{3.9}$$

[3] The proof for these inequalities can be seen in [20, p. 182]

Values for α and β	Obtained Measure
$\alpha = \beta = 1$	SYN_{JC}
$\alpha = \beta = 1/2$	SYN_{DC}

Table 3.8: Equivalences between *Ratio Model* and *Dice* and *Jaccard* coefficients.

This measure generalizes some of the previous ones modifying the values for α and β. Table 3.8 show them.

Although ratio model is a non-symmetric measure, it is symmetric whether $|A| = |B|$ or whether $\alpha = \beta$. For this reason, this measure seems to be the best one for describing the real behaviour of synonyms in natural language. However, it entails certain difficulties because the cardinalities of A and B depend on the thesaurus dictionary and, therefore, the symmetry or non-symmetry depends on the weights α and β. The main question is how these weights can be estimated.

A solution for providing a higher degree of synonymy to $\text{SYN}_{RM}(a,b)$ than $\text{SYN}_{RM}(b,a)$ if it is the case that a is synonymous of b in the thesaurus but b is not synonymous of a. Hence, the weight of β must be double the weight of α; i.e., $\beta = 2\alpha$. The obtaining measure is the following one [84]:

$$\text{SYN}_{NSRM}(a,b) = \frac{|A \cap B|}{|A \cap B| + |A - B| + 2 \cdot |B - A|} \tag{3.10}$$

3.3 Basic Inference Pattern

In chapter 2 we presented a proposal of syllogistic reasoning involving fuzzy quantifiers and a way to deal with them. As presuppositions, we assumed fully true statements and crisp chaining among the terms. Taking this model as a starting point, our next step is to propose an extension to deal with syllogisms where the chaining is not crisp but fuzzy, understanding this as those cases where the middle term is played by two similar terms, that can be linked by a fuzzy inclusion, by a fuzzy overlapping or by synonymy. We only focus on asymmetric syllogisms, i.e., those which chaining is through a middle term. The chaining of symmetric syllogisms is based on logic operators, and, therefore, they are out of the aim of this proposal.

The procedure for calculating the value for the quantifier of the conclusion Q_C is independent of the strength of the chaining in the middle terms. As we explained, quantifiers only refer to quantity relationships between the terms independently of their semantics. Therefore, any modification in the chaining of the middle term cannot be transferred to the quantifier as

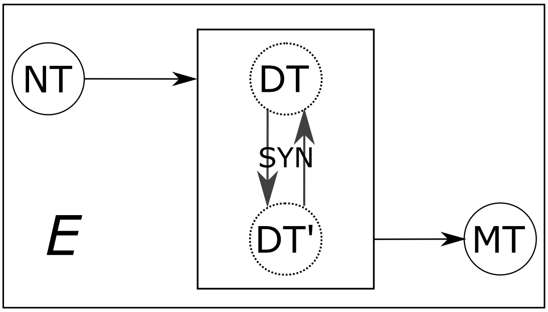

Figure 3.5: Graphic schema of ACS of Figure I.

it is a semantic relationship independent of the cardinality of the terms. Thus, we propose that the conclusion of this type of syllogism must be qualified or weighted by a degree of confidence γ. This degree should depend on the degree of synonymy between the involved terms, given that the higher the degree of synonymy between them, the higher the degree of confidence on the conclusion. Therefore, we should have $\gamma \in [0,1]$, with $\gamma = 1$ when the terms are completely equivalent or identical (classical syllogism), and $\gamma = 0$ when there is no synonymy relationship between them (fallacious syllogism). Taking this into account, we propose γ to be defined as the degree of synonymy between the terms DT and DT'. As a consequence, different quantitative values can be obtained for different synonymy measures (expressions (3.3)-(3.7)). The most adequate one should be chosen for each particular use.

Firstly, we shall consider a syllogism in its basic form; i.e., the one comprising three terms, two premises and a conclusion. The four positions that the middle term can have correspond with the four Aristotelian figures.

Figure 3.5 shows the graphic representation of Figure I with approximate terms. In a context E, a syllogistic argument involves an NT and an MT (crisp or fuzzy sets; circle with solid lines) and two terms (DT and DT'), which are fuzzy sets (circles with dotted lines) related by a SYN relationship that denotes similarity between them.

Figure 3.6 shows the graphic representation of Figure II. In this case, DT and DT' are the predicates in both premises.

Figure III is represented in Figure 3.7. In this case, DT and DT' are the subjects in both premises.

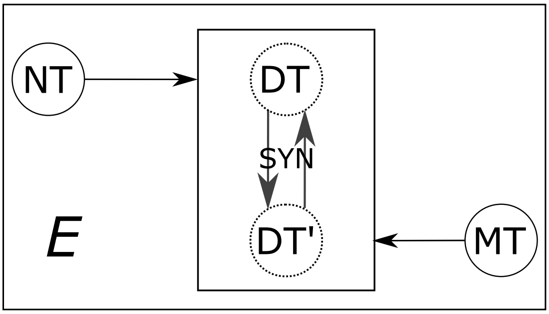

Figure 3.6: Graphic schema of ACS of Figure II.

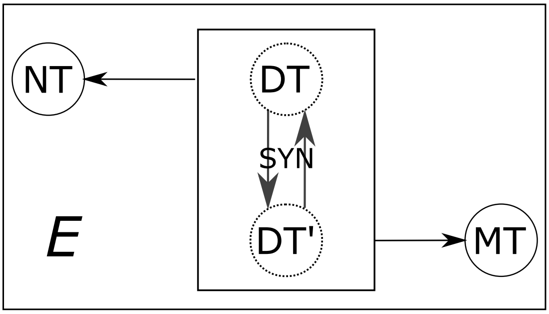

Figure 3.7: Graphic schema of ACS of Figure III.

Finally, Figure 3.8 shows the representation of Figure IV. This is the converse of Figure I; i.e., the middle term is the predicate in the first premise and the subject in the second one, but preserving the same form in the conclusion.

Table 3.9 summarizes the linguistic schemas of the four Aristotelian Figures with synonyms in the middle term, where Q_i stands for the respective quantifiers of the argument, and DT and DT' stands for the similar terms in the role of middle term. It is worth noting that the modification of the position of the middle terms in the premises does not entail any modification in our underlying thesis (the higher the degree of similarity between terms the higher the degree of confidence in the conclusion), as this modification is only relevant for calculating the quantifier of the conclusion.

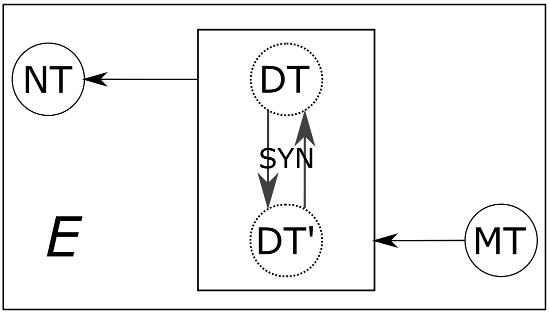

Figure 3.8: Graphic schema of ACS of Figure IV.

Figure I	Figure II	Figure III	Figure IV
Q_1 DTs are MTs	Q_1 MTs are DTs	Q_1 DTs are MTs	Q_1 MTs are DTs
Q_2 NTs are DT's	Q_2 NTs are DT's	Q_2 DT's are NTs	Q_2 DT's are NTs
Q NTs are MTs γ	Q NTs are MTs γ	Q NTs are MTs γ	Q NTs are MTs γ

Table 3.9: The fuzzy chaining version of the four Aristotelian Figures.

3.3.1 A Brief Analysis of the Measures of Synonymy based on Examples

In section 3.2.1, we pointed out that the selection of the measure of synonym depends on the context in which the syllogism is used. In this section, we shall introduce some examples of different situations to discern some rules or indications for the selection of one measure among others.

In [12, p. 275], the argument of Table 3.10 is proposed as an example of argument where the synonymy between the DTs terms, *wealthy* and *rich*, is perfect. According to WordNet, their Synsets are the following:

– SYNSET(wealthy)={affluent, flush, loaded, moneyed, wealthy, substantial, rich}

– SYNSET(rich)={rich, affluent, flush, loaded, moneyed, wealthy, substantial}

In this case, for any measure of distance, the result is 1 because SYNSET(wealthy)= SYNSET(rich) and, hence $\gamma = 1$. In fact, any of the measures pointed out is valid when synonymy is total. However, this is not the most frequent case in natural language, but it may

All lawyers are *rich people*
No *wealthy people* are vagrants
No lawyers are vagrants

Table 3.10: Syllogism with degree 1 of synonymy in DT.

All *denizens* are native
Some citizens are *resident*
Some citizens are native

Table 3.11: Syllogism with a high degree of synonym in DT.

be interesting for scientific fields, where the specific vocabulary is usually univocal rather polysemic.

Table 3.11 shows a new example of syllogism of Figure I where the role of DT is taken by the terms *denizen* and *resident*. We consider an informal conversation between two friends talking about democracy and politics.

Firstly, we analyse the number of glosses of each term and if they share some:

– Denizen: i) person who inhabits a particular place; ii) a plant or animal naturalized in a region.

– Resident: i) someone who lives at a particular place for a prolonged period or who was born there; ii) a physician (especially an intern) who lives in a hospital and cares for hospitalized patients under the supervision of the medical staff of the hospital.

In this case, we can see that they share a gloss i) partially, what is more, gloss i) of *denizen* is a subclass of gloss i) of *resident*. The other two glosses of both concepts are completely different (one is about plants and animals while the other is about medicine) and it is not excessive to assume that this facilitates selection of gloss i) to our context. In this case, we can draw up a shortlist of synonymy measures between those with the higher values; for instance, $\textsc{Syn}_{OC}$ and $\textsc{Syn}_{MSC}$.

So, to apply the measures, we have to choose the synonyms that appears in gloss i). Given the relationship between the glosses of both terms explained above, in *denizen* we focus on synonyms that fall under the label "direct hyponym" and in *resident* under the label of "direct hyperonym". It is worth noting that, at this level, *resident* only has 1 gloss while *denizen* has 30. This can lead us to select the low level measure as there is a big difference between them;

however, if we analyse these glosses in detail, $23/30$ are related with concrete geographical areas such as Europe, Australia, etc. and, hence, they are too specific to be considered in this context. Therefore, the two previously selected measures $\textsc{Syn}_{OC}$ and $\textsc{Syn}_{MSC}$ will be applied. Thus, the selected Synsets are the following:

– SYNSET(denizen)={inhabitant, habitant, dweller, denizen, indweller, cottager, cottage dweller, liver, marcher, resident, occupant, occupier, tellurian, earthling, earthman, worldling}

– SYNSET(resident)={resident, occupant, occupier, inhabitant, habitant, dweller, denizen, indweller}

Applying the $\textsc{Syn}_{OC}$, we obtain the highest degree of synonymy between both terms, which is shown in 3.11:

$$\begin{aligned}\textsc{Syn}_{OC}(DE,RE) &= \frac{|Synset(DE)\cap Synset(RE)|}{min(|Synset(DE),Synset(RE)|} = \\ &= \frac{8}{8} = 1\end{aligned} \tag{3.11}$$

i.e., the degree of confidence of the conclusion is $\gamma = 1$. For the Mutual similarity coefficient, the other selected, we obtain;

$$\begin{aligned}\textsc{Syn}_{MSC}(DE,RE) &= \frac{\frac{|Synset(DE)\cap Synset(RE)|}{|Synset(DE)} + \frac{|Synset(DE)\cap Synset(RE)|}{|Synset(RE)}}{2} = \\ &= \frac{\frac{16}{2}+\frac{8}{8}}{2} = 0.75\end{aligned} \tag{3.12}$$

that is, $\gamma = 0.75$.

As we can see, there is a significant difference in the degree of confidence of the conclusion attending to the measure of synonymy that is used. $\textsc{Syn}_{OC}$ emphasizes shared terms with respect to no shared ones since it is sensitive to the smallest Synset in comparison. As can be observed, if the smallest Synset is included in the largest one (although there are many others different synonyms) the result is 1 because fits it with $min(Synset(a),Synset(b))$. This is an interesting result as it adequately represents fuzzy inclusion. Syn_{MSC} is more restrictive as it weights the Synsets of both terms; although the result obtained is sufficiently high.

Hence, this example shows us that the selection of the Synsets to apply the measure must be performed also attending to the number of glosses that the terms share. The criteria for

All *dishes* are beautiful
Some *porringers* are artworks
Some beautiful things are artworks

Table 3.12: Syllogism with a low degree of synonym in DT.

selecting them is necessarily the context in which the syllogism is uttered, but the variability in the pattern definition of glosses [26] seems to make this step in one difficulty to automatize. How to solve this question is out of the aims of this research, but it is an open question to a further development of this version of syllogism.

Table 3.12 shows another example of Figure III, where the middle term is the subject in both premises. The role of DT is played by the terms *dishes* and *porringers*. The context is the organization of an old house.

The term *dish* has six glosses;

– a piece of dishware normally used as a container for holding or serving food

– a particular item of prepared food

– the quantity that a dish will hold

– a very attractive or seductive looking woman

– directional antenna consisting of a parabolic reflector for microwave or radio frequency radiation

– an activity that you like or at which you are superior

and *porringer* only one;

– a shallow metal bowl (usually with a handle)

As we can see, there is a meaningful difference respect to the number of glosses in both terms; *dish* has many more than *porringer*. Furthermore, *dish* involves three glosses related with objects in the kitchen and food although in the case of *porringer* this reference is omitted. For that reason, we propose selecting the measures with the lower values as the similarity between both terms seems to be accidental rather structural. Thus, we shall use $\textsc{Syn}_{JC}$ and $\textsc{Syn}_{DC}$.

Next, we shall show the Synsets of both concepts. *Porringer* only has one direct hypernym while *dish* has ten direct hyponyms:

- SYNSET(porringer)={porringer, bowl}
- SYNSET(dish)={bowl, butter dish, casserole, coquile, gravy boat, gravy holder, sauceboat, boat, Petri dish, ramekin, ramequin, serving dish, sugar bowl, watch glass}

Again, we calculate the the value for SYN_{JC} (3.13) and SYN_{DC} (3.14):

$$\begin{aligned} \text{SYN}_{JC}(PO,DI) &= \frac{|Synset(PO) \cap Synset(DI)|}{|Synset(PO) \cup Synset(DI)|} = \\ &= \frac{1}{15} = 0.066 \end{aligned} \tag{3.13}$$

$$\begin{aligned} \text{SYN}_{DC}(VO,CI) &= \frac{2 \cdot |Synset(PO) \cap Synset(SI)|}{|Synset(PO)| + Synset(DI)|} = \\ &= \frac{2}{16} = 0.125 \end{aligned} \tag{3.14}$$

Thus, for SYN_{JC}, $\gamma = 0.66$ and for SYN_{DC}, $\gamma = 0.125$. These low results are consistent with our intuition about the syllogism, as the two terms are not so close. Although the conclusion of the syllogism is true, it is difficult to accept it as an inference from the premises and, hence, the syllogism has a low degree of reliability.

With these examples we try to show how the context is fundamental for selecting the Synsets that should be considered to calculate the similarity between the terms. It determines which glosses are relevant and, from these, we can make a shortlist of similarity measures to calculate the corresponding γ for the conclusion. However, from the point of view of this implementation, the different patterns used to define the glosses make it difficult to deal with them.

3.3.2 Syllogisms with Weighted Quantified Statements

The fact that the conclusion of fuzzy chained syllogisms is a quantified statement with the form $Q\ A$ *are* $B\,[\gamma]$, where γ stands for the degree of fulfilment or confidence of the statement, is a major change in the habitual form of the statements involved in a syllogism, since they are usually assumed to be (fully) true and the usual concern is what types of information can be inferred from them [88]. To reason with this new form of quantified statements, which we

PR1:	$Q_1\ L_{1,1}$ are $L_{1,2}[\gamma_1]$
PR2:	$Q_2\ L_{2,1}$ are $L_{2,2}[\gamma_2]$
...	
PRN:	$Q_N\ L_{N,1}$ are $L_{N,2}[\gamma_n]$
C:	$Q_C\ L_{C,1}$ are $L_{C,2}[\gamma_C]$

Table 3.13: Broad reasoning schema for weighted quantified statements.

call *Weighted Quantified Statements* (WQS), it is necessary to incorporate a new procedure to manage the degree of confidence; hence it is independent of the calculation of the quantifier.

Table 3.13 shows our proposal of syllogistic schema compatible with TGQ developed in chapter 2. The value for Q is calculated according to the usual procedure, independently of the value of γ, as the degree of confidence of the conclusion does not modify the value of the quantifier. On the other hand, γ_C is calculated as $min(\gamma_1, \gamma_2)$ as the premises that constitute the syllogism are connected by the logical operation *AND*. This manner of dealing with the degrees of confidence in the premises is the same as in Mamdani's and other fuzzy reasoning models and uses involving rules with degrees of confidence or certainty factors [14].

The aggregation of the premises is based on *and* and, therefore, γ_C must be calculated according to (3.15).

$$\gamma_C = \bigwedge_{n=1}^{N} \gamma_n \tag{3.15}$$

3.4 Summary

In this chapter we introduce a new form of approximate syllogism involving vagueness not only in the quantifiers but also in the terms. This inference schema had already been considered in classical crisp syllogism but as a kind of invalid syllogism: the so-called fallacy of four terms. This appears when the role of middle term is played by two different terms generating, in this way, a non-existent chaining between the extreme terms.

Nevertheless, the terms also can be fuzzy and, although the chaining is not perfect, it can be strong enough to generate an acceptable argument in the context of everyday reasoning. In this case, we assume that the conclusion must be qualified through a degree of confidence γ. It depends directly on the degree of similarity of the middle terms: the higher the degree of similarity, the higher the degree of confidence.

We explore in depth the similarity in terms of synonymy, a typical semantic relationship between terms. We assume that the higher the number of synonyms that two terms share, the higher the degree of similarity between them. Using the on line tools, such as WordNet or its equivalent in Galician language Galnet, we develop a procedure for measuring the synonymy of two terms:

1. Select the gloss(es) of the terms according to the context of the syllogism.

2. Make a shortlist of measures of similarity to be applied depending on the number of glosses that they share: the higher the number of univocal glosses, the selection the higher measures of similarity.

3. Select the adequate synset to each concept.

4. Calculate the degree of synonymy of the terms according to the selected measure and move this value directly to γ.

It is worth noting that this proposal is compatible with our model of syllogistic approach compatible with TGQ (see chapter 2), since the calculation of the quantifier of the conclusion is independent of its degree of confidence. Furthermore, and following similar models in the literature if the premises involved in the syllogism are also qualified through a value γ, the degree of confidence of the conclusion is the result of applying the *min* operator over the γ_i of the premises because the premises of the syllogism are linked by *and* operators.

Thus, incorporating vague terms to syllogism is a new extension thereof which helps to deal with more fragments of natural language, as human beings manage a huge number of vague terms, not only in the quantifiers but also in the other parts of the premises. Nevertheless, the context-dependence and the irregular definition of glosses makes it difficult to address the automation of this procedure, and the treatment thereof will be proposed as future work.

On the other hand, obtaining quantified statements qualified by a degree of fulfilment leads us to consider the Generalized LSO proposed by P. Murinová and V. Nóvak [36] to develop a fuzzy LSO and, therefore, obtain a new set of fuzzy immediate inferences. This possibility is proposed as future work.

Chapter 4

Use Cases: Probabilistic Reasoning Using Syllogisms

In section 1.3 we explored the equivalence between probabilities and proportional quantified statements; particularly, the use of binary proportional quantified statements as a linguistic way of expressing conditional probabilities; in section 1.2.1, this relationship was analysed from the point of view of psychology of reasoning [43] and the syllogistic inference pattern was substituted by probabilistic reasoning. Nevertheless, as pointed out therein, this proposal entails a number of drawbacks, such as its incapability for managing quantifiers other than proportional ones. In this chapter, we shall adopt the opposite perspective; i.e., to use the syllogism as a tool for managing conditional probabilities and probabilistic reasoning using natural language.

This same idea has already been partially explored in the literature. In [3, 16], in the framework of Interval Syllogistics, a combination between Pattern I (see section 1.1.4) and a Generalized version of Bayes' Theorem (GBT), characterized by comprising only conditional probabilities, is proposed. The result is a hybrid inference schema between syllogism and Bayesian reasoning. However, this is limited to deductive inferences through the chaining of the terms involved in the argument. In [27], another proposal for dealing with fuzzy probabilities is developed using linguistic terms. In this case, fuzzy logic is used to generalize the space of classical probability theory ($[0,1]$) with linguistic labels. This framework was applied to Bayesian Networks (BNs), a probabilistic tool for managing uncertainty that allows also to make inferences, and a linguistic version of them was developed. Nevertheless, the use of linguistic terms was only devoted to label the probability values, not to the building of the whole BN (e.g., the arcs).

In this chapter, we take this a step further proposing the use of natural language (in particular, fuzzy quantified statements) not only for expressing the probabilities of a BN but also for building the entire the reasoning procedure. Syllogisms, given the link pointed out between conditional probabilities and proportional quantifiers statements, and using the model that we have developed, offer us a framework that involves all the necessary tools for this objective: i) precise and vague quantifiers can be managed; ii) the chaining between several terms can be expressed; iii) our proposal is not reduced to deductive syllogisms but also has a heuristic character; and, iv) there is no restriction in the number of premises and terms that can be used. On the other hand, we shall also check the compatibility between our proposal and the GBT to verify whether we can deal with it, given that we use Interval Syllogistics as the fundamental base of our model.

If we adopt the perspective of users, our syllogistic model also shows some interesting aspects to be considered with respect to the probabilistic approaches habitually used to deal with situations or problems involving uncertainty or vagueness. We can substitute precise numerical percentages with approximate linguistic quantifiers changing only the quantifier of the corresponding premise. In addition, the entire syllogism is built using the same type of propositions (binary quantified statements, which everybody is familiar with) and to make it easier the modelling for non-specialized users; the proper use of probability, in contrast, usually requires certain level of specialized knowledge.

Our main objective is to illustrate how a computational model of syllogism can be used in other fields than the traditional one (elaborate simple arguments involving only three terms and a few quantifiers) which require the use of natural language argumentation and probabilities. We shall illustrate this idea using examples taken from medical diagnoses and legal arguments. We shall use precise cases as, adopting a common idea in fuzzy logic, they must be a particular case of fuzzy models; if they are wrong, the fuzzy generalization is not valid. In addition, as explained in chapter 2, the step in our model from precise values to fuzzy ones is simple.

This chapter is organized as follows: section 4.1 describes the main characteristics of the combination Pattern I with GBT developed in [3, 16] and some inferences using our model of syllogism; section 4.2 introduces BNs and how their different basic types of topological patterns can be modelled using syllogisms; an example of medical diagnosis is explained;

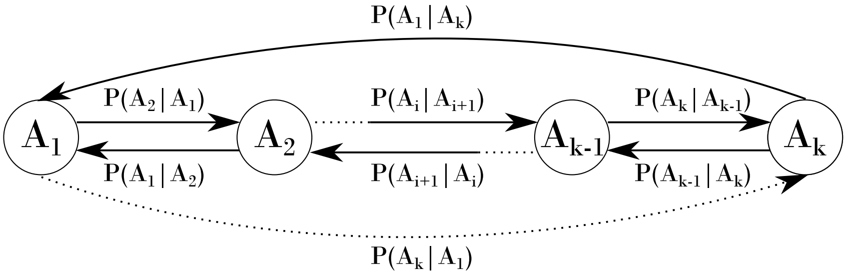

Figure 4.1: The chaining of terms described by GBT.

in section 4.3 we illustrate the use of syllogisms in legal argument to avoid probabilistic mistakes; both examples are developed to illustrate through real cases how syllogisms would be built; section 4.4 describes how to build a set of linguistic labels like Interval Syllogistics proposes (see 1.1.4.2); and, finally, section 4.5 summarizes our main conclusions about these questions and the strengths and drawbacks of our proposal.

4.1 Generalized Bayes' Theorem

An initial proposal for the combination of syllogistic and probability reasoning can be found in [3, 16]. It is developed in Interval Syllogistics (see 1.1.4) and consists of the combination of Pattern I with the Generalized Bayes' Theorem (GBT), a version of Bayes' theorem using only conditional probabilities. It addresses arguments involving more than three terms, although limited to an unidirectional chaining among the terms or nodes of the argument. Figure 4.1 illustrates the basic structure of GBT, where $A_1, A_2, \ldots, A_k$ are the nodes of the probabilistic network or the terms of the syllogism. For instance, for $k = 4$, for calculating $P(A_4|A_1)$, we need to know $P(A_1|A_4)$, $P(A_1|A_2)$, $P(A_2|A_1)$, $P(A_2|A_3)$, $P(A_3|A_2)$, $P(A_3|A_4)$ and $P(A_4|A_3)$.

The mathematical expression of this theorem is shown in (4.1)[1].

$$\forall A_1, \ldots, A_k, P(A_1|A_k) = P(A_k|A_1) \prod_{i=1,k-1} \frac{P(A_i|A_1)}{P(A_{i+1}|A_i)} \tag{4.1}$$

GBT only uses positive conditional probabilities as the authors assume that the a priori probability values that classic Bayes' theorem needs are unknown. For that reason this version is qualified as *Generalized*.

[1]The proof can be checked in [3, pp. 28-29].

	student	**sport**	**single**	**young**	**children**
student	$[1.000, 1.000]$	$[0.700, 0.900]$	$[0.000, 1.000]$	$[0.850, 0.950]$	$[0.000, 1.000]$
sport	$[0.400, 0.600]$	$[1.000, 1.000]$	$[0.800, 0.850]$	$[0.900, 1.000]$	$[0.000, 1.000]$
single	$[0.000, 1.000]$	$[0.700, 0.900]$	$[1.000, 1.000]$	$[0.600, 0.800]$	$[0.050, 0.100]$
young	$[0.250, 0.350]$	$[0.800, 0.900]$	$[0.900, 1.000]$	$[1.000, 1.000]$	$[0.000, 0.050]$
children	$[0.000, 1.000]$	$[0.000, 1.000]$	$[0.000, 0.050]$	$[0.000, 0.050]$	$[1.000, 1.000]$

Table 4.1: Results of the students' study about the people in the campus.

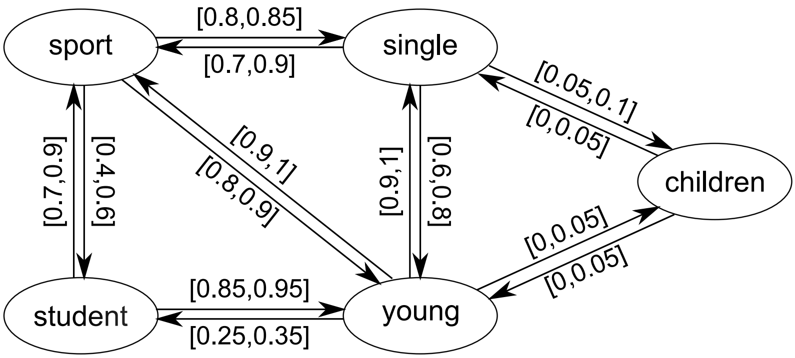

Figure 4.2: Example of network for reasoning with Pattern I and GBT.

To illustrate how this proposal works, we shall analyse the example used in [16]. Let us consider the people in a university campus. A group of students in sociology make a small study analysing the different links among those who are students, single, young, practice sports and have children. To ensure the most reliable values, different students ask the same questions and obtain slightly different results. To preserve these differences, they are represented by means of intervals. Table 4.1 shows the results obtained.

Once the table is completed, they discover that some information is missing. This is represented through the interval $[0.000, 1.000]$, which means indetermination. Some examples of this are the conditional probabilities between *students* and *singles* or between *have children* and *practice sport*. Figure 4.2 shows the graphical representation of the values in Table 4.1.

To obtain the missing values, the framework of Pattern I with GBT can be applied. As previously stated, GBT does not substitute Pattern I; rather both reasoning schemas are complementary. Pattern I can only manage arguments with three term-sets while GBT needs as initial values the information of arcs that can only be generated using Pattern I. The procedure for combining both inferences schemas is explained using pseudo-code as follows [3, p. 30]:

while *Probability intervals can be further improved* **do**
  repeat
    the probability intervals cannot be further improved;
  until *Apply Pattern I to generate the missing arcs*;
  repeat
    improve the arcs generated by Pattern I;
  until *Apply GBT*;
end

Algorithm 1: Pseudo-code for Pattern I and GBT.

	student	**sport**	**single**	**young**	**children**
student	$[1.000, 1.000]$	$[0.900, 0.900]$	$[0.607, 1.000]$	$[0.850, 0.850]$	$[0.000, 0.271]$
sport	$[0.400, 0.400]$	$[1.000, 1.000]$	$[0.850, 0.850]$	$[0.900, 0.958]$	$[0.000, 0.154]$
single	$[0.222, 0.366]$	$[0.700, 0.700]$	$[1.000, 1.000]$	$[0.800, 0.800]$	$[0.050, 0.100]$
young	$[0.350, 0.350]$	$[0.834, 0.888]$	$[0.900, 0.900]$	$[1.000, 1.000]$	$[0.000, 0.050]$
children	$[0.000, 0.099]$	$[0.000, 0.127]$	$[0.000, 0.050]$	$[0.000, 0.044]$	$[1.000, 1.000]$

Table 4.2: *Saturated* network of Figure 4.2.

Once the model is applied, the *saturated* network described in Table 4.2 is obtained. As we can see, not only are the values for the missing nodes, such as *students* and *single*, improved, but also certain values of the initial data, such as *students* and *sport* (from $[0.7, 0.9]$ to $[0.9, 0.9]$).

Next, we shall apply our model of syllogistic reasoning to illustrate, with some examples, how this problem can also be addressed using fuzzy quantified statements and syllogisms. The first step is to build the premises of the argument, which must include all the available information; i.e., Table 4.1. The form of the premises is the standard one (Q S are P)[2], where S are the nodes of the lines and P the nodes of the columns. This is shown in Table 4.3.

The conclusion has the same form of the premises. For each arch to be calculated, the conclusion has to be modified accordingly. Table 4.4 shows the results obtained, which are identical to the ones in Table 4.2; keeping very small differences due to the approximated numerical procedure.

Most of these values ($19/25$) can be calculated in one step using the argument of Table 4.3 and changing the corresponding terms in the conclusion. Nevertheless, some of them ($6/25$) cannot be calculated using this procedure, since there are some incongruities between

[2] To facilitate the expression of the argument and avoid expressions such us "Q students are players of sports", the verb *to be* is substituted in some cases by non-copulative verbs although preserving the meaning of the sentence.

PR1	Between 70% and 90% students play sports
PR2	Between 85% and 95% students are young
PR3	Between 40% and 60% sport players are students
PR4	Between 80% and 85% sport players are single
PR5	Over 90% of the sport players are young
PR6	Between 70% and 90% single people are sport players
PR7	Between 60% and 80% single people are young
PR8	Between 5% and 10% single people have children
PR9	Between 25% and 35% young people are student
PR10	Between 80% and 90% young people are sport players
PR11	Over 90% of the young people are single
PR12	Less than 5% of the young people have children
PR13	Less than 5% of the people with children are young
PR14	Less than 5% of the people with children are single
C	*Q S* are *P*

Table 4.3: Syllogistic argument for Figure 4.2.

	student	**sport**	**single**	**young**	**children**
student	$[1.000, 1.000]$	$[0.899, 0.900]$	$[0.607, 1.000]$	$[0.849, 0.849]$	$[0.271, 1.000]$
sport	$[0.399, 0.399]$	$[1.000, 1.000]$	$[0.849, 0.849]$	$[0.900, 0.958]$	$[0.000, 0.153]$
single	$[0.222, 0.366]$	$[0.699, 0.700]$	$[1.000, 1.000]$	$[0.799, 0.800]$	$[0.049, 0.099]$
young	$[0.350, 0.350]$	$[0.834, 0.887]$	$[0.899, 0.900]$	$[1.000, 1.000]$	$[0.000, 0.050]$
children	$[0.000, 0.099]$	$[0.000, 0.126]$	$[0.000, 0.049]$	$[0.000, 0.044]$	$[1.000, 1.000]$

Table 4.4: Results for Figure 4.2 applying syllogisms.

the premises, owing to the fact that our method is global whilst Pattern I with GBT is local (only arguments with three nodes are considered). We shall illustrate this problem with an example. The arc that links *sport players* with *have children* is not compatible with the premises "$[0, 0.05]$ people with children are single" and "$[0, 0.05]$ people with children are young". Once this premise has been eliminated, the obtained result, $Q = [0, 0.1539]$, is again the exact result.

In summary, we have illustrated how our proposal of syllogistic system can deal with arguments involving only conditional probabilities without restrictions in the number of terms, premises and in the position in the chaining terms. This entails a simplification of the algorithm designed for using the GBT, as the division into two steps is not necessary. Furthermore, we obtain more precise results in one single execution in most of the cases, although in some

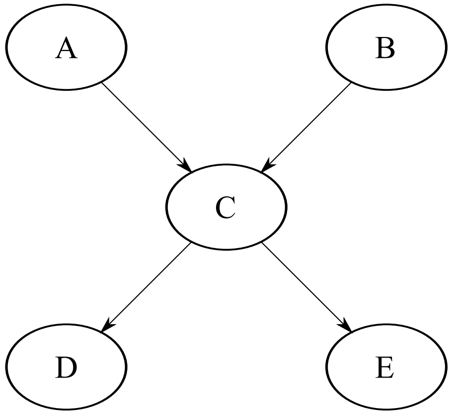

Figure 4.3: Example of BN for causality.

of them more executions are needed to obtain the precise results. In any case, the number of steps is less than or equal to those used by the algorithm for GBT.

4.2 Bayesian Networks

BNs are a knowledge representation tool for dealing with domains with uncertainty, codifying the information in a graphical structure using probabilities and allowing also to make different types of inferences [30, p. 29]. They are composed of nodes, that represent a set of random variables from the domain, and arcs, which link pairs of nodes, showing the direct dependencies between them. Figure 4.3 shows an example of a BN, where we can see that nodes A and B are independent events, as they are not connected between themselves while C, D and E are dependent, as there are arcs that link C with D and E. The only constraint about these graphs is that they are *directed acyclic graphs* (DAGS) (you cannot return to a node simply by following directed arcs).

Nodes are characterized by being mutually exclusively and exhaustively defined; i.e., each variable can only take one single value at a time. The most typical types of nodes are: Boolean (for instance, {True, False}), ordered (for instance, {small, medium big}) and integral (for instance, $\{1, 2, \ldots, 120\}$). The structure or topology of the network captures the links between the variables. Table 4.5 shows the values that the nodes can take.

Node name	**Type**	**Values**
A	Binary	{a,b}
B	Boolean	{True,False}
C	Boolean	{True,False}
D	Ternary	{c,d,e}
E	Integral	{1,2,3,4,5}

Table 4.5: Nodes and values for the BN of Figure 4.3.

C	**P(D=c|C)**	**P(D=d|C)**	**P(D=e|C)**
T	0.3	0.5	0.2
F	0.7	0.1	0.2

Table 4.6: Nodes and values for the node D of Figure 4.3.

One of the most common uses of BNs is for modelling causality. In these cases, the topology of the network denotes causality relationships, where some nodes are causes, others are effects and the arcs indicate the direction from causes to effects. Considering again the network of Figure 4.3, we can say, for instance, that A is a cause of C or E is an effect of C.

This information used to build a BN is usually expressed trough the so-called Conditional Probability Tables (CPT). Table 4.6 shows an example of CPT for the nodes C and D, where the numerical values are always conditional probabilities. It is worth noting the sum of the values of each row must be 1 to satisfy the constraint of exhaustiveness.

Regarding reasoning with BNs, they basically allow us to make two types of inferences: i) inferences without evidence, ii) inferences with evidence. Case i) only deals with inferences following the direction of the arcs (i.e.; from causes to effects in causal BNs) and consists of in the simple propagation of the values introduced in the BN (*a priori* probabilities and CPTs). For instance, in the BN of Figure 4.3, let us $P(A = a) = 0.9$ and $P(B = True) = 0.3$, a case i) inference is calculating the probability of C; i.e., $P(C = True) = 0.0116$.

Case ii) involves the so-called evidence. These are the new data that modify the value of any node of the network. In this case, BNs offer many more possibilities of inferences. This process, known as conditioning (also called probability propagation, inference or belief updating), flows through the arcs, although it is not limited to their directions. Basically, the reasoning process consists of calculating the new probability distribution for a set of query nodes given the new observations (or evidence) incorporated into the BN. For instance, in the

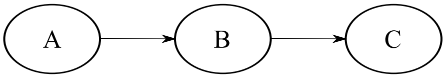

Figure 4.4: Causal Chain network.

BN of Figure 4.3, a case ii) inference is $P(D=c|C=T)=0.3$, which the evidence is $(C=T)$. Three types of evidences are usually considered:

- Specific evidence: A node takes a particular value; i.e., $B=True$.
- Negative evidence: The evidence says that the node cannot take one of possible values although it can take any of the remaining ones; i.e., $D=not_c$.
- Virtual or likelihood evidence: The evidence is incorporated to the BN but it is not fully reliable.

4.2.1 Topological Types of Bayesian Networks

Next, we shall describe the specific characteristics of BNs when they are used for representing causality (the parent node is a cause and the child node is an effect). These are the so-called topological patterns, which define how the variables are connected and how the evidences in one of them modify the values of the others. There are three basic patterns: causal chain, common causes, and common effects [30, p. 40-41].

4.2.1.1 Causal Bayesian Networks: Causal Chain

Consider a chain between three nodes A,B and C, where A causes B and this C, as shown in Figure 4.4. This type of chain gives rise to conditional independence as $P(C|B)=P(C|A,B)$. In other words, the probability of C depends on the probability of A if and only if there is no evidences of B; in this is the case, the probability of A does not affect C. This idea can be summarized in the following rule:

Rule 1 The evidence is transmitted from an extreme to the another in a causal chain, save that some intermediate node is instantiated.

The network can be extended to any n number of intermediate nodes between the extreme ones.

Step 1 Identifying the number of nodes of the network. They are the terms of the syllogism.

Step 2 Expressing the *a priori* probability of the parent nodes as "*Q U are X*", where *Q* denotes the quantifier that expresses the a priori probability (which can either be precise, imprecise or fuzzy); *U* the universe of the discourse or reference; and, *X* the corresponding node.

Step 3 Expressing the conditional probabilities tables (CPT) as "*Q Y are Z*", where *Q* denotes the quantifier that express the conditional probability (which can also be precise, imprecise or fuzzy); *Y* denotes the occurrence of the given node or parent node; and, *Z* the node under consideration or child node; in TGQ, *Y* denotes the scope and *Z* the restriction of the quantifier. These statements, together with the previous ones, are the premises of the syllogism.

Step 4 Expressing the probability that is aimed to be calculated as the conclusion of the syllogism. It follows the same schema; i.e., "*Q X are Z*", where *Q* denotes the quantifier calculated by our model; *X* denotes the occurrence of a given node or the universe of the discourse; and, *Z* the node under consideration.

Table 4.7: Procedure for building a syllogism for Causal Chain pattern.

To build the syllogism corresponding to this network, the procedure in Table 4.7 must be followed. We shall illustrate this procedure with the network in Figure 4.4: **step 1**, there are three nodes (*A, B* and *C*) and, hence, three terms; **step 2**, $(P(A) = 0.3) \equiv$ "*30% elements of Us are As*"; **step 3**, $P(B|A) = 0.05) \equiv$ "*5% As are Bs*"; **step 4**, $P(C|A) = 0.20) \equiv$ "*20% As are Cs*".

4.2.1.2 Causal Bayesian Networks: Common Causes

Consider a network where a node B has two effects, A and C. It is represented by a network of inverted v-structure; as shown in Figure 4.5. A and C are linked through a common cause B; therefore, evidence of any of the children nodes modifies the remaining ones; i.e., an evidence about A modifies the probability value of C through B. If B is instantiated, this is the only one that determines the value of the other children. This idea can be summarized in the following rule:

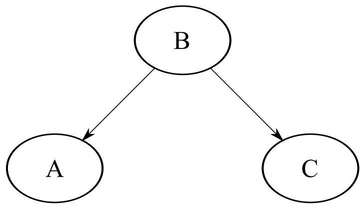

Figure 4.5: Common Causes network.

Rule 2 The evidence is transmitted between the children in a common cause network, except if that the parent node is instantiated.

The network can be extended to any n number of children nodes.

The manner of building the syllogism is the same of causal chain pattern (see Table 4.7) and we shall illustrate it with the network in Figure 4.5: **step 1**, the syllogism involves three terms A, B and C; **step 2**, the *a priori* probabilities are of a single node, $(P(C) = 0.3) \equiv$ "*30% elements of Us are Cs*"; **step 3**, all the conditional probabilities between B and A and B and C must be also included; for instance, "*10% Cs are As*", "*60% Cs are Bs*", etc.; **step 4**, the conclusion, e.g. "*5% Us are As*" or "*30% Cs are As*".

4.2.1.3 Causal Bayesian Networks: Common Effects

Consider a network where a node B (effect) has two causes, A and C. It is represented by a network v-structure; as shown in Figure 4.6. The parent nodes, A and C, are independent, which means that evidence of any of them does not modify the value of the other. Nevertheless, they become conditionally dependent if the common effect is instantiated. The following rule summarizes this idea:

Rule 3 The evidence is not transmitted between the parents in a common effect network, except if the child node is instantiated.

The network can be extended to any n number of parent nodes.

The strategy for establishing the syllogism is the same of the previous models, although an additional consideration is needed. The procedure is shown in Table 4.8 and we shall illustrate it with the network in Figure 4.6: **step 1**, the three terms involved, A, B and C; **step 2**, the *a priori* probabilities of the parent nodes, "*20% Us are As*" and "*30% elements of Us are Cs*", and their indepdence, $P(A,C) = P(A) \cdot P(C)$; i.e., $(P(A) \cdot P(C) = 0.2 \cdot 0.3 = 0.06) \equiv$ "*6% Us*

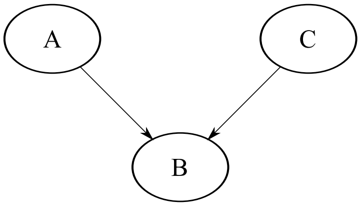

Figure 4.6: Common Effects network.

Step 1 Identifying the terms of the syllogism from the nodes of the network.
Step 2 The *a priori* probabilities of the parent nodes must be included. In addition, the independence between these nodes must also be explicitly included as their joint probability according to the independence rule. This value is expressed through a conjunction in the predicate of the premise.
Step 3 The conditional probabilities between the parent and child nodes of the CPTs.
Step 4 The conclusion preserves the standard form; i.e., "*Q X are Z*".

Table 4.8: Procedure for building a syllogism for Common Effects pattern.

are As and Cs"; **step 3**, "*60% As are Bs*", "*70% Cs are Bs*", etc.; **step 4**, "*10% Us are Cs*", etc.

Once we have defined the structure and the different forms that the networks can adopt, the next step is to describe the reasoning procedure. BNs allows different types of inferences, which are analysed in the following section.

4.2.2 Types of Inferences with Bayesian Networks

The use of BNs for reasoning enables us to execute different types of inferences. Depending on their direction, three basic types are defined [30, p. 34]:

– Predictive: It goes from causes to effects; i.e., given evidence about a cause, the new probabilities of each effect is calculated. It follows the direction to the arcs.

– Diagnostic: It goes from effects to causes; i.e., given evidence about an effect, the probability of each cause is calculated. It follows the opposite direction to the arcs; it is the opposite of the previous one.

– Intercausal (explaining away): It is the reasoning about mutual causes of a common effect. Thus, given two causes initially independent, if we have new evidence about one of the causes and the effect, the probability of the other cause is modified. In other words, two causes initially disconnected become dependent by the common effect updated by new evidence.

Diagnostic and predictive inferences can be executed in all networks described in section 4.2.1 while Intercausal inference can only be applied in common causes pattern, since it is devoted to analysing the dependency generated between two previously independent nodes. There is an additional pattern of inference that can be mentioned, although we do not consider it as a basic one as it is a combination between two of the basic ones: diagnostic and predictive. It is the so-called *combined* inference.

Figure 4.7, taken from [30, p. 34], illustrates these four types of reasoning using a BN with five binary nodes with an X-structure (a combination between *inverted v* and *v* structures), which results from the combination of common causes and common effects patterns. The dotted ellipses denote the nodes with the new evidence and the grey ellipses the queries that must be solved.

To illustrate how each one of these types of inference works and, how they can be modelled using syllogisms, we shall use an example of medical diagnosis of lung cancer also took from Korb's book [30, p. 30].

4.2.3 BNs in Medical Diagnosis: Lung Cancer

Let us suppose the fictitious hospital Princeton-Plainsboro, where the following situation happens:

> A patient has been suffering from shortness of breath (called *dyspnea*) and visits the doctor, worried that he has lung cancer. The doctor knows that other diseases, such as tuberculosis and bronchitis, are possible causes, as well as lung cancer. She also knows that other relevant information includes whether or not the patient is a smoker (increasing the chances of cancer and bronchitis) and what sort of

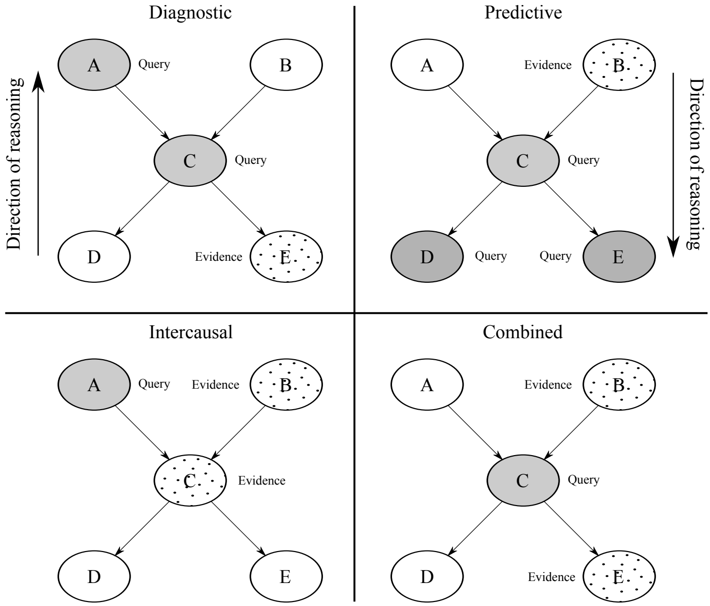

Figure 4.7: Types of Bayesian reasoning.

Node name	Type	Values
Pollution	Binary	{Low,High}
Smoker	Boolean	{True,False}
Cancer	Boolean	{True,False}
Dyspnoea	Boolean	{True,False}
X-ray	Binary	{Positive,Negative}

Table 4.9: Types and values of the nodes of BN for the lung cancer problem.

air pollution he has been exposed to. A positive X-ray would indicate either tuberculosis or lung cancer.

With the information included in this wording it is possible to identify the variables and values described in Table 4.9 and the BN described in Figure 4.8, where the corresponding CTP are also stated.

In the following subsections, we shall analyse the different types of inference and how they must be expressed as syllogisms.

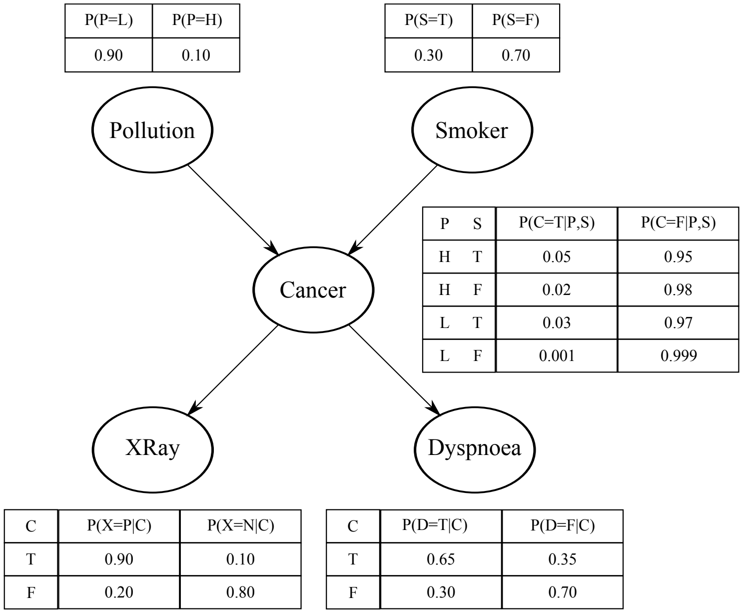

Figure 4.8: A BN for the lung cancer problem.

4.2.3.1 Predictive Reasoning

The oncologist, J. Wilson, wants to know the *a priori* probability that the patient has lung cancer, X-Ray positive and dyspnoea. From the point of view of the inference to be executed in the BN, this is a simple propagation without additional evidence using any of the explained patterns (see section 4.2.1).

The elaboration of a syllogism comprises several steps. As previously stated, the premises must include the *a priori* probabilities of the parent nodes over a particular universe and the CPTs. In this case, given that the network has an X-structure with five binary nodes, there are two premises for the *a priori* probabilities of the root node, one for including the independence between them (which is calculated as the product of the *a priori* probability of each node), and eight for including all the conditional probabilities of the CPTs. In the example of Figure 4.8,

PR1	90% of people live with LP
PR2	30% of people are S
PR3	3% of peopole live under HP and are S
PR4	5% of people that live under HP and are S have C
PR5	2% of people that live under HP and are not S have C
PR6	3% of people that live under LP and are S have C
PR7	0.1% of people that live under LP and are not S have C
PR8	90% of people with C have positive X
PR9	20% of people without C have positive X
PR10	65% of people with C have D
PR11	30% of people without C have D
C	*Q S* are *P*

Table 4.10: Syllogistic argument about lung cancer.

the nodes *Pollution* and *Smoker* are the root nodes; independent, which entails[3]:

$$P(PH\&ST|E) = P(PH) \cdot P(ST) = 0.1 \cdot 0.3 = 0.03 \tag{4.2}$$

$$P(PH\&SF|E) = P(PH) \cdot P(SF) = 0.1 \cdot 0.7 = 0.07 \tag{4.3}$$

$$P(PL\&ST|E) = P(PL) \cdot P(ST) = 0.9 \cdot 0.3 = 0.27 \tag{4.4}$$

$$P(PL\&SF|E) = P(PL) \cdot P(SF) = 0.9 \cdot 0.7 = 0.63 \tag{4.5}$$

Table 4.10 shows the syllogistic argument corresponding the Figure 4.8. PR1 and PR2 denotes the *a priori* probability of the independent events *live with low pollution air* and *be smoker* and PR3 the joint probability of both; and PR4 to PR11 is the information of the CPTs. It is worth noting that in this example, since all variables are binary or Boolean, it is only necessary to include half of the information of the CPTs (with one of the values) because the other can be directly derived from these by the exhaustiveness of the network.

The conclusion comprises the information that is aimed to obtain; i.e. the query of the BN (see Figure 4.7). In general, the subject term is the evidence and the predicate term is the query node. The calculation of having cancer is the direct application of common causes pattern, while the calculation of positive X-Ray and Dyspnoea entails the combination of common causes pattern with causal chain one.

[3]LP stands for "low pollution", HP stands for "high pollution", S stands for "smoker", C stands for "cancer", X stands for "X-Ray" and D stands for "dyspnoea".

Conclusion	Syllogism	BN
Q people have cancer	0.011630	1.16%
Q people have positive X-Ray	0.208141	20.8%
Q people have Dyspnoea	0.304074	30.4%

Table 4.11: Conclusions for simple propagation.

Conclusion	Additional Premise	Syllogism	BN
Q S have C		0.032000	3.20%
Q S have D	30.7% D are S	0.031116	31.1%
Q people live with HP air have D	89.2% of people with D live with LP	0.310152	31.0%
Q people C have X		0.900000	90.0%

Table 4.12: Conclusions for predictive reasoning with evidence.

Given that in this case our queries are about *a priori* probabilities, the subject term of the conclusion is the universe of reference. Table 4.11 summarizes the conclusions obtained with our model compared with the obtained from the BN[4].

Another possibility in the predictive reasoning is the inference with new evidences about the nodes; for instance, the probability of have a *positive X-Ray* given that the patient is *smoker*. This is expressed in syllogistic terms as "Q smokers have positive X-Ray", according to the said above.

In this case, an additional consideration must be taken into account. If the new evidence is about a parent node of the query node taken; no additional premises are required; for instance, for the conclusion "Q smokers have cancer" it suffices to consider the basic set of premises. In other case, the inverse statement of the conclusion must be added as an additional premise; for instance, in order to conclude "Q smokers have dyspnoea" the statement "Q patients with dyspnoea are smokers" must be added as an additional premise. If this information is not incorporated into the argument, the quantifier of the conclusion will be more imprecise, although consistent in any case with the precise result. It is worth noting that through an iterative procedure (such us the one explained in section 4.1) it is possible to specify, iteration by iteration, the obtained result. Table 4.12 shows some results with the additional premises that must be added with each conclusion, if it is the case.

[4]The BN was modelled with NETICA, a graphical tool for modelling BNs.

Conclusion	Additional Premise	Syllogism	BN
Q D have C		0.024861	2.49%
Q D live with HP air	31.0% D are PH	0.101950	10.2%
Q people with X positive have D	21.7% D have positive X	0.304000	30.4%
Q people with X negative are S	77.8% S have negative X	0.295075	29.5%

Table 4.13: Conclusions for diagnostic reasoning.

4.2.3.2 Diagnostic Reasoning

Traditionally, syllogism was devoted exclusively to deductions which means, in terms of BNs, predictive inferences. However, as we have pointed out in chapter 2, our model has also a heuristic character and, hence, other type of inferences can be managed. In this section, we shall illustrate its behaviour respect to diagnostic reasoning.

Diagnostic inferences always require evidence; therefore, the inverse one from the conclusion must be added as an additional premise in almost all cases. Only when the query node (the predicate-term of the conclusion) is the parent of the evidence (the subject-term) is this additional premise not necessary. As in the predictive reasoning with evidences, if these premises are not included, the final results obtained are more imprecise.

Table 4.13 illustrates some examples. As can be observed, the results of the syllogism coincide with the BN. Diagnostic inference can be applied in any of the inference patterns explained in section 4.2.1.

4.2.3.3 Intercausal Reasoning

Intercausal reasoning or explaining away is mainly applied in inferences that use a common causes pattern. For instance, in our example, given the evidence of being *smoker* and having *cancer*, calculate the probability of being *exposed to low pollution air*.

This type of inference requires using at least two items of evidence instead of a single one. This condition can be easily expressed in the syllogism. Thus, given that two items of evidence are simultaneously required, the linguistic way for expressing it is through an *and*: i.e., a conjunction in logic terms. Since our model can manage this type of Boolean operations, both items of evidence must appear in the subject of the conclusions linked by an *and*. In the example that we are analysing, it is only possible between the nodes *pollution*, *smoker* and *cancer*. Given that the nodes under consideration are immediate, no additional premises are demanded. Table 4.14 shows four conclusions using intercausal pattern.

Conclusion	Additional Premise	Syllogism	BN
Q PH with C are S		0.517241	51.7%
Q PH without C are S		0.293512	29.4%
Q S without C are PL		0.901860	90.2%
Q S with C are PH		0.156250	15.6%

Table 4.14: Conclusions for combined reasoning.

4.2.3.4 Combined Reasoning

Combined reasoning is defined as the combination of predictive and diagnostic reasoning. As we have explained above, we do not consider this model as a basic one.

From the point of view of syllogism, this type of reasoning can present certain problems with the inclusion of the inverse of the conclusion as an additional premise as two directions of the inference are executed at the same time. With the current definition for syllogisms that we manage, inconsistencies between the premise and the conclusion appear in some cases.

Given that the other three basic reasoning schemas can be effectively expressed in terms of syllogisms, the development of this type of inference will be addressed as future work.

4.3 Probabilistic Reasoning in Legal Argumentation

Not all probabilistic reasoning is straightforward to be modelled and executed through BNs. We shall focus on legal argumentation, as it is a discipline in which the arguments in natural language are fundamental and can determine the guilt or innocence of a person.

We shall illustrate the usefulness of our model of syllogistic reasoning in this field analysing the so-called *Prosecutor's fallacy* [83]. This consists of a bad use of probabilities in statistical reasoning and it is typically used by the prosecution, to argue the guilt of the accused. Several cases of this type of argument has been documented [68], such us the Sally Clark case[5] or the well-known case of People vs. Collins [30, pp.18-19], which we shall consider in what follows:

> In 1964 an interracial couple was convicted of robbery in Los Angeles, largely on the grounds that they matched a highly improbable profile, a profile which fit witness reports. In particular, the two robbers were reported to be a particular set

[5] A document from the Royal Statistical Society about this case can be consulted in http://www.rss.org.uk/uploadedfiles/documentlibrary/744.pdf.

Characteristics	**Probabilities**
A man with a moustache	$1/4$
Who was black and had a beard	$1/10$
And a woman with a ponytail	$1/10$
Who was blonde	$1/3$
The couple was interracial	$1/1000$
And were driving a yellow car	$1/10$

Table 4.15: Characteristics and probabilities of the suspicious couple.

> of characteristics and the prosecution suggested that these had the probabilities shown in Table 4.15.
>
> The prosecutor calls an instructor of mathematics who, assuming that all these events are mutually independent, applies the "product rule" obtaining that the couple were innocent was no more than $1/12000000$. The jury was persuaded and the couple convicted.

Now, we shall express this case using a syllogism and applying our model to estimate the probability of the couple being guilty. Our universe the population of Los Angeles and the information of Table 4.15 the used one to build the premises of the argument. The conclusion is the frequency of a couple like the one described by the witness in the set of couples in Los Angeles; i.e., "Q people are a black man with a moustache and beard and a blonde woman with a ponytail and an interracial couple driving and yellow car". Table 4.16 shows the argument which result is $Q = [0,1]$; i.e., inconclusive. It shows that the set of premises used to build this syllogism is not enough to achieve a relevant conclusion.

It is obvious that these premises are not independent but that there are some clear links between them. For instance, if a man has a moustache, the probability of having beard changes, as very few people with beard do not have moustache. In [30, p. 19-20], a new analysis of this problem attending to the conditional probabilities that links these premises among them is explained. For instance, PR2 entails PR1 and PR2; PR3 and PR4 entails PR5, which give us a probability of innocence around $1/3000$; which means that any couple that satisfies these properties is guilty. The next step is to calculate the *a priori* probability of a random couple is being guilty of robbery. Attending to the population of Los Angeles and being generous, there are $1,625,000$ possible couples and, hence, the probability of being a guilty couple is $1/1625000$. With this new information, we obtain the argument of Table 4.17.

PR1	25% people have moustache
PR2	10% people are black and had beard
PR3	10% people are women with a ponytail
PR4	33% people are blonde women
PR5	0.1% people are interracial couples
PR6	10% people are driving a yellow car
PR7	100% man with a mustache and black with beard and a woman with a ponytail and blonde and an interracial couple and drive and yellow car are guilty
C	Q man with a moustache and black with beard and a woman with a ponytail and blonde and an interracial couple and drive and yellow car are not guilty

Table 4.16: Syllogism for the case People vs Collins.

PR1	0.03% interracial couple are a man with a mustache and black with beard and a woman with a ponytail and blonde
PR2	0.000006% couple of Los Angeles are guilty
PR3	100% interracial couple are couple
C	Q man with a moustache and black with beard and a woman with a ponytail and blonde are guilty

Table 4.17: Syllogism with the conditional probabilities of People vs Collins.

The application of our model produces again $Q = [0,1]$; i.e., the result is still inconclusive. Although applying Bayesian reasoning a result of 0.002 chance of culpability is obtained [30, p. 20], which does not mean that 99.8% is the best probability of innocence either. The syllogistic argument reduces this probability to a total indetermination as it only considers the ends; i.e., with the available information, the set of couples who committed the crime. For a result where $Q = [0,0]$ means that no couple committed the crime; for $Q = [1,1]$ means that all couples committed the crime, given that the set under considerations has a single couple as member, the culpability is total.

Thus, we can conclude that the conclusions of a syllogistic argument are more careful, less committed than probabilistic inferences as they only consider the ends. This may be a virtue or a drawback, depending on the context and its purpose. For instance, in this type of use, it is more favourable for the accused.

The basic inferences shown in this chapter were executed using SEREA (the software library developed by ourselves referred to in chapter 2). Currently, it us under development and only BNs with less than five nodes can be managed by it. More complicated networks (for instance, the BN developed in the example of section 4.2.3) require a handmade formalization to be performed before its computational execution. In this case, we have developed a small programme in Octave for the optimization step.

4.4 Qualitative Version using Linguistic labels

We illustrated the behaviour of our syllogistic model for managing some examples of probabilistic reasoning in two fields where many of the probabilities can be expressed using linguistic quantifiers instead of exact probabilities. For instance, instead of saying “90% of the population is exposed to low air pollution” we say “the vast majority of population is exposed to low air pollution” or “0.000006% couple in Los Angeles are guilty” is substituted by “almost no couple in Los Angeles is guilty”. Thus, using our method, a linguistic version of Bayesian reasoning can be built, where the probabilities are defined through a set of linguistic labels and the results are also expressed and approximated to acceptable values for those linguistic quantifiers. The obtained results will not be totally precise but they will be more interpretable and therefore more use-oriented than precise ones from the point of view of their usability.

An initial extension to introduce linguistic labels is the proposal analysed in section 1.1.4, where a way of introducing linguistic quantifiers into fuzzy syllogism without using a version of the LSO is explained. As has also been mentioned, the main differences between both approaches is that the LSO is based on the truth properties of the quantified statements, while the use of lattices of linguistic labels is based on the properties of order and certainty of the quantifiers; there are no considerations about the truth value of the statements. Our proposal is fully compatible with the way linguistic quantifiers are introduced into Interval Syllogistics. Therefore, the procedure described in section 1.1.4.2 can be directly applied to our proposal.

It is true that all the examples described involve crisp values instead of fuzzy ones. This fact had two motivations: i) the examples are taken from a book about classic probabilistic reasoning [30]; ii) in this way, we illustrate how the crisp case can be perfectly managed with our model assuming the principle that the fuzzy case must include a crisp one as a particular case.

4.5 Summary

We have proposed three fields, where the information is usually expressed using natural language, for applying our framework for syllogistic reasoning: i) inferences using only conditional probabilities; ii) inferences using BNs; iii) arguments using other type of probabilistic reasoning. All the explained examples show a correct behaviour of our model, although the elaboration of the syllogistic argument needs certain considerations for its correct definition. There are summarized in Table 4.18.

We emphasize that the expression of probabilistic problems in syllogisms contributes to making the problem more interpretable. The premises and the conclusion have the same form (*Q A are B*) and new evidence can be directly added in the subject of the conclusion. On the other hand, it can deal simultaneously with precise and fuzzy quantifiers, which is a very typical situation in the real word. In addition, it is also relevant to note that the interval or fuzzy set obtained for the quantifier of the conclusion also denotes how complete and precise the information of the premises is: the smaller the size of the result, the higher the precision of information of the premises. Its compatibility with the use of lattice of linguistic quantifiers allows us to easily generate a kind of *linguistic Bayesian reasoning*; which, for the fields considered (medical diagnosis and legal argument) is simpler, since the building of a BN and its associated conditional probabilities usually demands the collaboration of an expert.

We have also illustrated how to express a problem by a syllogism and how it can help us to check the validity of our arguments. We show that a bad argument can have highly negative consequences, such as an wrongful conviction in the case of People vs. Collins shown in section 4.3.

By way of conclusion, it is also relevant to note that we have proposed new application fields for syllogisms different to its traditional use, showing in this way new possibilities for it. In addition, new types of inferences, different from the deductive ones, were included as a part of syllogistic inferences, something new in this type of argument.

Topology of pattern	Type of inference	Nodes between evidence and query		Additional premise	
		0	> 0		
Causal chain	Predictive	0	> 0	No additional premises	
	Predictive with evidences	0	> 0	No additional premises	Inverse of the conclusion
	Diagnostic	0	> 0	No additional premises	Inverse of the conclusion
Common causes	Predictive	0	> 0	No additional premises	
	Predictive with evidences	0	> 0	No additional premises	Inverse of the conclusion
	Diagnostic	0	> 0	No additional premises	Inverse of the conclusion
	Intercausal	0	> 0	No additional premises	No checked
Common effects	Predictive	0	> 0	No additional premises	
	Predictive with evidences	0	> 0	No additional premises	Inverse of the conclusion
	Diagnostic	0	> 0	No additional premises	Inverse of the conclusion

Table 4.18: Guidelines for building a syllogism.

Conclusions and Future Work

The starting hypothesis of this research was the development of a model for dealing with syllogistic arguments involving fuzzy quantifiers, more than three terms or more than two premises. To achieve this aim, several objectives have been defined which we analyse below.

The **first objective** was an in-depth review of different proposals of syllogistic models in the literature. Two main interpretations for the quantified statements that constitute syllogisms were identified: i) as assertions that define a particular quantity relationship between two sets; ii) as sentences whose depth form is a conditional proposition where the subject and the predicate respectively correspond with the antecedent and the consequent of the conditional and the quantifier denotes how strong is the link between them.

We called interpretation i) *set-based interpretation*, and it assumes an ontology based on sets. Thus, quantified statements are assertions whose purpose is to describe a fact denoting the link between two sets of elements, the subject and the predicate terms, and, the quantifier denotes the cardinality of that relationship. Interpretation ii), which we call *conditional interpretation*, assumes an ontology based on individuals and properties. Thus, quantified statements are complex propositions with a conditional form, where the subject is the antecedent and the predicate is the consequent. In this way, the fact is described in terms of its truth conditions and not as a description, and the role of the quantifier is to denote how strong is the link between the antecedent and the consequent, which can denote probability, belief or modality.

The analysis of both proposals led us to some relevant conclusions for the subsequent steps of this work. Conditional interpretation allows us to transform syllogisms into arguments of first order logic and to use well known reasoning patterns such as *Modus Ponens* (like Qualified Syllogistics proposes). On the other hand, it also has a direct link with probabilistic reasoning insomuch as quantifiers are interpreted in terms of probabilities. Nevertheless, from the point of view of the aim of this research, it shows two relevant drawbacks: i) a logical analysis of the statement to identify the conditional form under the natural language surface is required; and, ii), the most important one, it does not have a clear way for managing quantifiers different to the proportional ones ("most", "few", etc.) such us absolute ("five", "around ten", etc.) or comparative ("around double", "less than half", etc.), which are also very common in natural language. For both questions, set-based interpretation shows a better behaviour as it allows us to preserve the natural language form and manage the different types of quantifiers defined by TGQ. Despite this, the equivalence between proportional quantified statements and conditional probabilities is fruitful, as we shall explain, for using syllogisms to deal with some problems involving probabilities.

On the other hand, we also adopt Aristotelian Syllogistics as the classical referential framework, as it is the base of this type of reasoning and its compatibility with TGQ also has been proved. It is worth noting that in all but two of the analysed syllogistic systems based on *set-based interpretation* their compatibility with Aristotelian Syllogistics is explicitly analysed. The exceptions are Interval Syllogistics, which has been analysed in this thesis and is compatible, and Fuzzy Syllogistics, which is not compatible with Aristotelian Syllogistics and, in addition presents some other meaningful problems for its use. All these objectives were achieved in chapter 1.

The next step was the development of a syllogistic model based on the set-based interpretation incorporating the definitions of TGQ and making flexible the inference pattern. Firstly, we assumed Venn Diagrams (with more or fewer modifications) as the adequate knowledge representation tool for syllogisms, as all syllogistic models analysed except the fuzzy ones. They offer us a clear representation of the structure of the relationships involved in a syllogistic argument in terms of disjoint subsets and allow us to assign to each one the cardinality (fuzzy or not) denoted by the quantifier. At this step, we introduced algebraic methods to rewrite every premise of the syllogism as an inequation whose values are the bounds of the quantifier. If the quantifier is precise, the bounds match up; if it is imprecise, the bounds denote an interval; if it is fuzzy, the bounds are a four-tuple (the typical notation for a trapezoidal function) used together with α-cuts. In this way, we are interpreting syllogistic inferences as a mathematical optimization problem, which opened up the possibilities for using them not only as deductions but as a heuristic procedure for finding the conclusion that is most compatible with the restrictions or constraints defined as premises. Thus, we have divided the process into three phases: i) dividing the universe of discourse into disjoint sets; ii) defining sentences as systems of inequations; iii) selecting the optimization method that should be applied in each case in order to resolve the reasoning process. When the quantifier that must be optimized is an absolute one, the Simplex method can be used directly, while if it is a proportional one, linear fractional programming techniques must be adopted. These results satisfy the **second** and **third objectives** initially defined in this research.

The interpretation of this reasoning process in terms of a mathematical optimization problem led us to a syllogism that not only makes deductions but is also valid for heuristic procedures and non-deductive inferences, where the conclusion denotes the proposition more compatible with the premises, being or not a logical consequence of them *stricto sensu*. Furthermore, our proposal also overcomes most of the habitual limitations identified in the other syllogistic systems:

– It can deal with arguments with no limitations in the number of premises and terms.

– It admits the combination of different types of quantifiers in the same argument.

– It is not limited to a set of inference patterns where the position of the middle term(s) is predefined but that it can appear in any position.

As we have also stated, vagueness can also appear in terms. For this reason, once the system for dealing with vague quantifiers was developed, we addressed the problem of vagueness in terms. We mainly focus on the middle term, since any minimum change on it may generate a fallacy. Thus, we started exploring the literature on similarity measures between terms and we identified three approaches: i) similarity as fuzzy inclusion, ii) similarity as fuzzy overlapping and iii) similarity as synonymy; each one of them with its proper context of use. After analysing them, we concluded that the first two are developed sufficiently but the third one has some points to be improved. With these results, we satisfy the **fourth objective** of this work, the study of different alternatives to measure similarity between terms.

To improve the use of synonymy as a similarity measure, we adopted the following idea: the higher the number of synonyms two terms share, the higher the degree of similarity between them. In addition, we used a pragmatic definition of synonymy: two words are synonyms if they appear as synonyms in a thesaurus. As thesauruses, we chose WordNet for English and Galnet for Galician language, as they are very complete and provide additional information to synonymy, such us *glosses*, which are relevant for our goals. Thus, using the data supplied by them, we developed an approach divided into two steps: i) calculating the proportion of *glosses* shared by the terms and, ii) calculating the number of synonyms shared by these *glosses*. Step i) provides a criterion for choosing the similarity measure from among the standard ones (the Jaccard, Dice, Cosine, Mutual similarity, Overlap and Ratio model coefficients) following the principle that *the higher the proportion of* glosses *shared, the higher the similarity measure chosen*; step ii) to provide the synonyms that must be used in the calculation. Using this information, we built an extension of our syllogistic system, using the same principle as Mamdani's reasoning model, where a chaining based on similarity is directly translated into a degree of confidence in the conclusion. In this way, we satisfied the **fifth objective** of this thesis: we built a reasoning schema with approximate chaining where not all the arguments are fallacies and the approximated nature of synonymy (that is also very common in natural language) is translated into the conclusion.

Syllogism was traditionally devoted to arguments of a very specific type. Nonetheless, the equivalence between quantified statements and conditional probabilities opens up the possibility for studying an analogous link in the field of reasoning. This constitutes, precisely, the **sixth objective** of this work. Specifically, we explored how diverse examples of Bayesian reasoning can be modelled using syllogisms.

First, we analysed a probabilistic reasoning model that only uses conditional probabilities (the described one as the combination of Pattern I of Interval Syllogistics with Generalized Bayes Theorem). Our proposal achieved the same results as this model, employing fewer steps and using the standard form of syllogism without additional distinctions. This contributes to making the arguments more easily understandable for non-specialized users and to clearly ascertaining all the premises that a syllogism comprises. In addition, another relevant consideration is that a highly unspecific quantifier in the conclusion means that the information of the premises is not sufficient to infer a more informative conclusion.

Other case analysed is Bayesian Networks (BNs). We illustrated how a case of medical diagnosis can be expressed in syllogistic terms. As we have explained, the premises of the syllogism express the *a priori* probabilities and the corresponding conditional probability tables. In the necessary cases, this set of premises can be complemented with the corresponding additional premises according to the type of inference. The conclusion, in any case, is the information that we attempt to obtain, where the subject-term comprises the evidence of the BN and the predicate the query node.

The final case analysed is an example of Bayesian reasoning in legal argumentation. We focus on how an erroneous modelling through Bayesian reasoning of a situation can be discovered using a syllogistic formulation. While the Bayesian argument was incriminating and conclusive, the syllogism was inconclusive as the information of the premises, as is explained, was not sufficient to support the conclusion. Thus, our model of syllogism can be used as a tool (for non-specialized users) that contributes to avoiding unjust verdict, as several cases have been documented.

In addition, we believe that the use of syllogisms also clarifies the reasoning with probabilities in other aspects:

- Syllogism allows us to use quantified statements to express probabilities. We have explained how *a priori* probabilities, conditional probabilities and independence between events must be explained and preserving for all these cases the standard form: *Q S are P*.

- Our model not only deals with precise probabilities but also with imprecise ones. We have developed the examples using precise probabilities to illustrate that our model includes them as a particular case. The extension to imprecise or fuzzy quantifiers is easy: by substituting the precise values for intervals or fuzzy sets.

- Rendering implicit knowledge explicit. One of the critical points of BNs is the modelling of the problem, that is usually complex for many users due to its numerical nature. The use of the syllogism requires all the information used to be make explicit; the more precise the premises are, the more precise the quantifier of the conclusion will be.

- Syllogism is not only valid for deductions but also for other types of reasoning. We have shown how different types of inferences, such as abductive or intercausal ones, can be executed using our syllogistic framework. This breaks away the traditional conception of syllogism and opens up new possibilities for this type of argument since only a small part of human reasoning is purely deductive.

- It is true that the calculation procedure in the syllogism can be opaque, but it is very clear what the premises and the conclusion are. This fact facilitates the communication and the understanding of the problem for non-specialized users.

The computational implementation of our model forms part of the **seventh objective**. We have developed a software library, SEREA, which was used in the basic examples (with fewer than five terms). It is still under development to improve its behaviour. The most complicated examples (arguments with five terms or more) were implemented in Octave, also obtaining suitable results.

To conclude, we shall evaluate the satisfaction of the initial hypothesis according to the results obtained for the defined objectives. All of them have been satisfactorily achieved; therefore, the hypothesis of developing a syllogistic framework for dealing with arguments involving fuzzy terms and quantifiers as well as more flexible inference patterns has been accomplished.

Future work

The achieved results in this research have suggested us some interesting topics for the future work.

Our immediate challenge is to improve our model of syllogistics to deal with other types of quantifiers other than the binary ones. We have focused mainly on binary quantifiers as they are, perhaps, the most common ones in ordinary reasoning and an equivalence between binary quantified statements and probabilities can be established. However, there are other quantifiers defined in the TGQ, such us the ternary (for instance, "there twice as many dogs as cats and parrots"), that also appear in natural language and, therefore, also can be used for reasoning. This idea allows us to define a new type of syllogism comprising premises with different arity: binary, ternary, etc.

Another interesting task is to explore more complex inferences in BNs. In this research, the basic patterns of BNs (the three topological structures with the three basic types of inference) have been explored. Thus, our next step will be to analyse more complex patterns in order to verify whether the syllogism preserves a consistent behaviour; for instance, networks with more than five nodes, diagnostic or predictive inferences with more than one item of evidence, intercausal reasoning with more than two causes, etc.

Approximate chaining syllogism also allows subsequent developments. We have mainly explored the concept of synonymy as the support for a similarity measure between the chaining terms. Nevertheless, other possibilities can be contemplated, such us using the concepts of analogy or antonymy instead of synonymy, which may generate new models of approximated chaining. On the other hand, the study of the context deserves special attention in this type of inference, since it can determine, for instance, the gloss used in each argument. In addition, it is one of the major challenges for the further evolution of Computing with Words[6]. We have dealt with this problem giving a principal role to the user, who has to assume all the decisions of the construction in the arguments, but the challenge in computational intelligence is to manage this automatically.

We also identify computational tasks. The first step is to improve the current version of the SEREA library for dealing with all the linguistic quantifiers that we have addressed in this research, such as similarity or comparative ones, and syllogistic inferences involving more than five terms. A preliminary version has been implemented in Octave, which indicates that the migration to SEREA should be straightforward. As a medium-term task, taking SEREA as starting point, we shall aim to develop a software package similar to available ones for dealing with BNs, such us NETICA. Thus, we propose to create a graphical interface to build the knowledge representation. Over this graphic model, the corresponding syllogism will be generated, incorporating the corresponding premises according to Table 4.18, displaying it to the user. A further possibility is the opposite way; i.e., from a syllogism introduced by the user, to generate a graphical representation. In addition, the graphical interface also allow us to include the management of linguistic labels described in Interval Syllogistics (1.1.4), where the user, for instance, could select between different numbers of predefined linguistic labels and obtain results approximated to them or even group together some of these labels. Finally, the approximate chaining syllogism also will be broadly developed, using the graphical interface to inform the user about different similarity relationships and their results.

[6]This problem was analysed in the panel session *How to deal with Context in Computing with Words?* in the last edition 2013 IEEE International Conference on Fuzzy Systems (FuzzIEEE2013) held in Hyderabad in July 7-10.

Bibliography

[1] Princeton University "About WordNet." WordNet, 2010.

[2] SEREA, March 2012.

[3] S. Amarger, D. Dubois, and H. Prade. Constraint propagation with imprecise conditional probabilities. In *Proceedings of the 7th Conference on Uncertainty in Artificial Intelligence*, pages 26–34, Los Angeles, CA, July 13-15 1991.

[4] Aristotle. *Prior and posterior analytics*. Clarendom Press, Oxford, 1949.

[5] Aristotle. *Aristotelis Politica*. Oxford University Press, Oxford, 1957.

[6] Aristotle. *Categories and De Interpretatione*. Clarendon Press, 1975.

[7] Aristotle. *Prior Analytics*. Hackett, Indianapolis, 1989.

[8] J. Barwise and R. Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4:159–219, 1981.

[9] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, The Edinburgh Building, Cambridge, UK, 2004.

[10] Alberto J. Bugarin and S. Barro. Fuzzy reasoning supported by petri nets. *IEEE Transactions on Fuzzy Systems*, 2(2):135–150, 1994.

[11] Alberto J. Bugarin and S. Barro. Reasoning with truth values on compacted fuzzy chained rules. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 28(1):34–46, 1998.

[12] Irving M. Copi. *Introduction to logic*. MacMillan, New York, 1994.

[13] J. Corcoran. *Ancient Logic and Its Modern Interpretations*. Synthese Historical Library. Springer, 1974.

[14] O. Cordón, M. J. del Jesus, and F. Herrera. A proposal on reasoning methods in fuzzy rule-based classification systems. *International Journal of Approximate Reasoning*, 20:21–45, 1999.

[15] A. P. Dempster. A generalization of bayesian inference. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30(2):pp. 205–247, 1968.

[16] D. Dubois, L. Godo, R. López de Mántaras, and H. Prade. Qualitative Reasoning with Imprecise Probabilities. *Journal of Intelligent Information Systems*, 2:319–363, 1993.

[17] D. Dubois and H. Prade. On fuzzy syllogisms. *Computational Intelligence*, 4(2):171–179, 1988.

[18] D. Dubois, H. Prade, and J-M. Toucas. *Intelligent Systems. State of the Art and Future Directions*, chapter Inference with imprecise numerical quantifiers, pages 52–72. Ellis Horwood, Great Britain, 1990.

[19] L. Eciolaza, M. Pereira-Fariña, and G. Trivino. Automatic linguistic reporting in driving simulation environments. *Applied Soft Computing*, 13(9):3956 – 3967, 2013.

[20] S. Fernández-Lanza. *Una Contribución al Procesamiento Automático de la Sinonimia Utilizando Prolog*. PhD thesis, University of Santiago de Compostela, 2002.

[21] Tim Fernando and Hans Kamp. Expecting many, 1996.

[22] P. T. Geach. *Reason and Argument*. University of California Press, 1976.

[23] B. Geurts. Reasoning with quantifiers. *Cognition*, 86:223–251, 2003.

[24] G. Girard. *Synonymes françois et leur différents significations et le choix qu'il faut faire pur parler avec justesse*. Veuve d'Houry, Paris, 1749.

[25] I. Glöckner. *Fuzzy Quantifiers: A Computational Theory*. Springer, Berlin, 2006.

[26] X. Gómez Guinovart, X. M. Gómez Clemente, A. González Pereira, and V. Taboada Lorenzo. Galnet: Wordnet 3.0 do galego. *LinguaMÁTICA*, 3(1):61–67, 2011.

[27] J. Halliwell and Q. Shen. Linguistic probabilities: theory and application. *Soft Computing*, 13:169–183, 2009.

[28] H. Hoijer. *Language in Culture: Comparative studies of cultures and civilizations*, chapter The Sapir Whorf hypothesis, pages 92 – 105. Comparative studies of cultures and civilizations. University of Chicago Press, 1954.

[29] W. Kneale and M. Kneale. *The Development of Logic*. Clarendon Press, Oxford, 1968.

[30] K.B. Korb and A.E. Nicholson. *Bayesian Artificial Intelligence*. Chapman & Hall/CRC, 2004.

[31] I. Landau. Ways of being rude. Technical report, Ben Gurion University, 2006.

[32] Y. Liu and E. E. Kerre. An overview of fuzzy quantifiers (ii). reasoning and applications. *Fuzzy Sets and Systems*, 95:135–146, 1998.

[33] J. Lukasiewicz. *Aristotle's Syllogistic*. The Clarendon Press, Oxford, 1957.

[34] J. Malczewski. Ordered weighted averaging with fuzzy quantifiers: Gis-based multicriteria evaluation for land-use suitability analysis. *International Journal of Applied Earth Observation and Geoinformation*, 8:270–277, 2006.

[35] George A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995.

[36] P. Murinová and V. Novák. Analysis of generalized square of opposition with intermediate quantifiers. *Fuzzy Sets and Systems*, (0):In press, 2013.

[37] P. Murinová and V. Novák. A formal theory of generalized intermediate syllogisms. *Fuzzy Sets and Systems*, 186:47–80, 2012.

[38] Wallace A. Murphree. *Numerically Exceptive Logic: A Reduction of the Classical Syllogism*. Peter Lang, New York, 1991.

[39] Wallace A. Murphree. The numerical syllogism and existential presupposition. *Notre Dame Journal of Formal Logic*, 38(1):49 – 64, 1997.

[40] V. Novák. On fuzzy type theory. *Fuzzy Sets and Systems*, 149:235–273, 2005.

[41] V. Novák. A comprehensive theory of trichotomous evaluative lingusitic expressions. *Fuzzy Sets and Systems*, 159:2939–2969, 2008.

[42] V. Novák. A formal theory of intermediate quantifiers. *Fuzzy Sets and Systems*, 159:1229–1246, 2008.

[43] M. Oaksford and N. Chater. *Bayesian Rationality. The probabilistic approach to human reasoning*. Oxford University Press, Oxford, 2007.

[44] Barbara H. Partee. Many quantifiers. In J. Powers and K. de Jong, editors, *Proceedings of the 5th Eastern States Conference on Linguistics*, pages 383–402, Columbus, 1989. Ohio State University.

[45] J. Pearl. *Causality: Models, Reasoning and Inference*. Cambridge University Press, 2000.

[46] C. S. Peirce. *Collected Papers of Charles Sanders Peirce*. Harvard University Press, Cambridge MA, 1931 – 1958.

[47] C. S. Peirce. *Collected Papers. Cambridge*. Harvard University Press, Cambridge, MA, 1933.

[48] M. Pereira-Fariña. Una revisión del concepto de intencionalidad en un marco aproximado. In *Actas del XVI Congreso Español de Tecnologías y Lógica Fuzzy (ESTYLF2012)*, pages 435–441, 17-19 de Septiembre, Cuencas Mineras (Mieres - Langreo), Asturias, España, 2008.

[49] M. Pereira-Fariña. A revision about the concept of intentionality in an approximate framework. *Mathware and Soft Computing*, 16:5–15, 2009.

[50] M. Pereira-Fariña, F. Díaz-Hermida, and A. Bugarín. An analysis of reasoning with quantifiers within the Aristotelian syllogistic framework. In *Proc. 2010 International Fuzzy Conference on Fuzzy Systems (FUZZ-IEEE 2010)*, pages 635–642, July, 18-23, Barcelona, Spain, 2010.

[51] M. Pereira-Fariña, F. Díaz-Hermida, and A. Bugarín. On the analysis of set-based fuzzy quantified reasoning using classical syllogistics. *Fuzzy Sets and Systems*, 214:83 – 94, 2013.

[52] M. Pereira-Fariña, L. Eciolaza, and G. Trivino. Quality assessment of linguistic description of data. In *Actas del XVI Congreso Español de Tecnologías y Lógica Fuzzy (ESTYLF2012)*, pages 608 – 613, 1 - 3 de Febrero, Valladolid, España, 2012.

[53] M. Pereira-Fariña and A. Sobrino. Recursos lingüístico-formales para el silogismo aproximado. *Submitted to XVII Congreso Español de Tecnologías y Lógica Fuzzy (ESTYLF2014)*.

[54] M. Pereira-Fariña and A. Sobrino. Una propuesta de silogismo encadenado aproximado. In C. Martínez Vidal, J. L. Falguera, J. M. Sagüillo, Víctor M. Verdejo, and M. Pereira-Fariña, editors, *Actas del VII Congreso de la Sociedad de Lógica, Metodología y Filosofía de la Ciencia en España*, pages 705 –713, 18-20 de Julio, Santiago de Compostela, España, 2012.

[55] M. Pereira-Fariña, A. Sobrino, and A. Bugarín. A proposal of fuzzy chained syllogism based on the concept of synonymy. In *Proc. 2013 International Fuzzy Conference on Fuzzy Systems (FUZZ-IEEE 2013)*, pages 1–8, July, 7-10, 2013 - HICC, Hyderabad, India, 2013.

[56] M. Pereira-Fariña, Juan C. Vidal, F. Díaz-Hermida, and A. Bugarín. A fuzzy syllogistic reasoning schema for generalized quantifiers. *Fuzzy Sets and Systems*, 234:79 – 96, 2014.

[57] M. Pereira-Fariña, Juan C. Vidal-Aguiar, P. Montoto, F. Díaz-Hermida, and A. Bugarín. Razonamiento silogístico aproximado con cuantificadores generalizados. In *Actas del XVI Congreso Español de Tecnologías y Lógica Fuzzy (ESTYLF2012)*, pages 701–706, 1 - 3 de Febrero, Valladolid, España, 2012.

[58] S. Peters and D. Westersthål. *Quantifiers in language and logic*. Clarendom Press, Oxford, 2006.

[59] P. Peterson. Higher quantity syllogisms. *Notre Dame Journal of Formal Logic*, 26(4):348–360, 1985.

[60] P. L. Peterson. On the logic of “few”, “many”, and “most”. *Notre Dame Journal of Formal Logic*, 20(1):155–179, 1979.

[61] P. L. Peterson. Distribution and proportion. *Journal of Philosophical Logic*, 24(2):193–225, 1995.

[62] P. L. Peterson. How to use aristotelian logic to formalize reasonings expressed in ordinary language. Technical Report FS-96-04, AAAI, Menlo Park, California, 1996.

[63] P. L. Peterson. *Intermediate Quantities. Logic, linguistics, and Aristotelian semantics*. Ashgate, Alsershot. England, 2000.

[64] N. Pfeifer. *Argumentation in Theorie und Praxis: Philosophie und Didaktik des Argumentierens*, chapter Contemporary syllogistics: Comparative and quantitative syllogisms, pages 57–71. LIT, Wien, 2006.

[65] G. Polya. *Mathematics and Plausible Reasoning: Induction and Analogy in Mathematics V. 1*. Princenton University Press, 1990.

[66] W.V.O. Quine. *Word and Object*. MIT Press, Cambridge - Mass, 1960.

[67] W.V.O. Quine. *From a Logical Point of View*, chapter Two dogmas of empiricism, pages 20–46. Har, Cambridge - Mass, 1980.

[68] D. Kim Rossmo. *Criminal Investigative Failures*. CRC Press Taylor & Francis Group, 2009.

[69] B. Russell. *From Frege to Gödel*, chapter Letter to Frege, pages 124–125. Harvard University Press, 1967 edition, 1902.

[70] B. Russell. *The Problems of Philosophy*. Williams and Nogate, London - United Kingdom, 1957.

[71] Daniel G. Schwartz. Dynamic reasoning with qualified syllogisms. *Artifical Intelligence*, 93:103–167, 1997.

[72] G. Shafer. *A Mathematical Theory of Evidence*. Princenton University Press, Princteton, NJ, 1976.

[73] C. E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1979.

[74] A. Sobrino-Cerdeiriña and S. Fernández-Lanza. Synonymy from a computational and philosophical perspective. In *Proceedings eighth international conference IPMU*, volume II, Madrid, 2000. CSIC.

[75] A. Sobrino-Cerdeiriña, S. Fernández-Lanza, and J. Graña-Gil. *Soft Computing in Web Information Retrieval. Models and Applications*, chapter Access to a Large Dictionary of Spanish Synonyms: A Tool for Fuzzy Information Retrieval, pages 299–316. Springer-Verlag, 2006.

[76] S. Solt. *Understanding Vagueness. Logical, Philosophical and Linguistic Perspectives*, chapter Vagueness in quantity: Two case studies from a linguistic perspective, pages 157–174. Collegue Publications, 2011.

[77] F. Sommers. *The Logic of Natural Language*. Clarendon Press, Oxford, 1982.

[78] K. Sparck Jones. *Synonymy and Semantic Classification*. Edinbürg University Press, 1986.

[79] M. Spies. *Syllogistic Inference under Uncertainty*. Psychologie Verlags Union, 1989.

[80] A. Tarski. The semantic conception of truth: and the foundations of semantics. *Philosophy and Phenomenological Research*, 4(3):pp. 341–376, 1944.

[81] B. Thompson. Syllogisms using "few", "many", and "most". *Notre Dame Journal of Formal Logic*, 23(1):75–84, 1982.

[82] B. Thompson. Syllogisms with statistical quantifiers. *Notre Dame Journal of Formal Logic*, 27(1):93–103, 1986.

[83] William C. Thompson and Edward L. Schumann. Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior*, 11(11):167–187, 1987.

[84] A. Tversky. Features of similarity. *Psychological Review*, 84(4):327–352, 1977.

[85] J. Venn. On the diagrammatic and mechanical representation of propositions and reasonings. *Philosophical Magazine and Journal of Science*, 9(59):1–18, 1880.

[86] D. Westersthål. Aristotelian syllogisms and generalized quantifiers. *Studia Logica*, 48(4):577–585, 1989.

[87] A. N. Whitehead and B. Russell. *Principia Mathematica*, volume 3. Cambridge University Press, Cambridge, 1910-1913.

[88] Ronald R. Yager. Reasoning with fuzzy quantified statements: Part II. *Kybernetes*, 15:111–120, 1986.

[89] Ronald R. Yager. Quantifier guided aggregation using owa operators. *International Journal of Intelligent Systems*, 11(1):49–73, 1996.

[90] L. A. Zadeh. A Computational Approach to Fuzzy Quantifiers in Natural Language. *Computer and Mathematics with Applications*, 8:149–184, 1983.

[91] L. A. Zadeh. The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems. *Fuzzy Sets and Systems*, 11:199–227, 1983.

[92] L. A. Zadeh. *Aspects of Vagueness*, chapter A Theory of Commonsense Knowledge, pages 257–296. D. Reidel, Dodrecht, 1984.

[93] L. A. Zadeh. Syllogistic reasoning in fuzzy logic and its applications to usuality and reasoning with dispositions. *IEEE Transactions On Systems, Man and Cybernetics*, 15(6):754–765, 1985.

[94] L. A. Zadeh. Test-score semantics for natural languages and meaning representation via pruf. In George J. Klir and Bo Yuan, editors, *Fuzzy sets, fuzzy logic, and fuzzy systems*, pages 542–586. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 1996.

List of Figures

List of Tables